DESIGNING FOR CHANGE

DATA AND DISADVANTAGE PROJECT

FINAL REPORT

Australian National University
School of Cybernetics

Australian National University

Designing for Change: Data and Disadvantage Project Final Report

Cultivating reflexivity through systems visibility, for purposeful data decision making

The Australian National University

Funding

The Paul Ramsay Foundation's purpose is to help end cycles of disadvantage in Australia by enabling opportunities for people and communities to thrive.

The research was funded by the Paul Ramsay Foundation. Any opinions, findings, or conclusions expressed in this report are those of the author(s) and do not necessarily reflect the views of the Foundation.

The Foundation would like to thank its partners who were involved in this research for their contributions.

School of Cybernetics

ANU College of Engineering, Computing & Cybernetics

+61 2 6125 2790

admin.cybernetics@anu.edu.au

The Australian National University

Canberra ACT 2601 Australia

www.anu.edu.au

CRICOS Provider No. 00120C

TABLE OF CONTENTS

KEY TERMS

Capta

Some academics within the digital humanities have argued for the necessity to refer to data especially humanistic data as 'capta'. This idea draws from the insight that the word 'data' is derived from the Latin *dare* which means 'to give'. Rob Kitchin and others have argued that although the pervasive attitude to data is something that is simply given, the practical reality is that data often taken and therefore rarely neutral but embedded with decisions and values and cultural factors[1]. It is argued that the term 'capta' should be used instead since capta is derived from the Latin *capere*, meaning to take. So capta seeks to describe those data 'units that have been selected and harvested from the sum of all potential data.'[2] While 'data' can represent all that is knowable about things such as a person, Kitchin and others have argued that capta should frame what is selectively captured through measurement like surveys, polls etc[3]. We use it in this report in the sense that approaching socially relevant data as capta invites data professionals to embrace their role in giving meaning to datasets.

Circular causality in causally closed systems

Within system dynamics, a definition of a system entails that it is causally closed. This means all causal influences within a defined system will feedback on themselves leading to various forms of recursive action. This process of causal influences feeding back on themselves is referred to as a causal loop, causal feedback loop or an instance of circular causality.

Closed and open systems

The terms closed and open systems have been co-opted into systems research from thermodynamics concepts in physics. In this thermodynamics sense, closed and open refer to the ability of a given system to exchange energy and/or matter with its environment. If a given system can exchange both energy and matter with its environment, it is referred to as an open system while it is defined as a closed system if it only exchanges energy.

Data decision-makers and data workers and professionals

In this report, we have used the broad term "data decision makers" deliberately, to nominate individuals working with or within data assets, and who decide how, who, when, and why data gets used and for what purpose. Those decisions about data can happen at management level as well as at entry levels. Throughout the report, we also use the term 'data decision-makers' and 'data workers' and 'data professionals' interchangeably; we also make references to more specific data decision-maker roles, such as data custodian, researcher, analyst, and the like. All data decision-makers work within data systems.

Data driven and data informed

When organisations are using or planning to use data, they must decide on a data-driven or data-informed approach or both. A data-driven approach to decision making seeks guidance for actions from model results that are continuously updating and are adaptive to changes. A data-

[1] Melody Lynch, "Data Lives: How Data Are Made and Shape Our World," *The AAG Review of Books* 10, no. 2 (April 3, 2022): 20–21, https://doi.org/10.1080/2325548X.2022.2036546.
[2] Rob Kitchin and Martin Dodge, *Code/Space: Software and Everyday Life* (The MIT Press, 2011), 261, https://doi.org/10.7551/mitpress/9780262042482.001.0001.
[3] Kitchin and Dodge, 261.

informed approach to decision-making seeks to use the results of data models as a supplement to other types of advice that is seen to be critical to the decision-making process[4].

Data integration and data linking

The goal of data integration is 'to offer uniform access to a set of autonomous [i.e. independent] and heterogeneous data sources'.[5] That is, data integration seeks to provide ways in which disparate and diverse data sets can be usefully linked or cross-indexed, so that despite having different variables and parameters, they can be used to provide a richer picture of or answer to any given situation or question.

In 2010, government in Australian agreed upon principles for data integration across agencies for statistical and research projects. They also agreed on guidelines for putting in place governance and institutional arrangements. Appointed/authorised data integrators head efforts to integrate government data sets and are also responsible for data ethics.

Data integration remains fraught with challenges, with efforts affected by the following technical and legal considerations:

- Hardware platforms
- database implementation
- query languages and templates
- data structures and schemas
- complex and/or distributed data ownership and maintenance
- legislative requirements and protections

Data systems/assets

Within this report, we take the term *data system/asset* to describe the personnel, organisations, services, technical tools, infrastructures, and platforms that generate, collate, administer, and use various types of datasets including census, econometric and administrational datasets.

In the Australian context, public research or statistical institutes are key players in the data system, with demographic (census) and econometric data being a substantive backbone on and against which other data sets can be established and/or compared. Both kinds of data sets have been especially pertinent to disadvantage studies because of the way in which disadvantage has been historically understood and measured as synonymous with poverty (lack of resources, especially monetary income) among particular social groups. Bodies such as the Australian

[4] Anne Fleur van Veenstra and Bas Kotterink, "Data-Driven Policy Making: The Policy Lab Approach," in Electronic Participation, ed. Peter Parycek et al., Lecture Notes in Computer Science (Cham: Springer International Publishing, 2017), 100–111, https://doi.org/10.1007/978-3-319-64322-9_9.

[5] AnHai Doan, Principles of Data Integration, 1st edition (Waltham, Mass: Morgan Kaufmann, 2012), 6.

Institute of Health and Welfare,[6] the Australian Institute of Family Studies,[7] and the Australian Bureau of Statistics have primary responsibility for data gathering and reporting.[8][9]

Further, services administrators from federal government agencies as well as jurisdictions are often stewards of large administration data sets gathered in the process of delivering public services. The administrators collect, manage, and use this data to record/track and assess service provision.

More information can be found at https://www.oaic.gov.au/.

Human flourishing

By human flourishing, we mean that each person considers their life to be good when evaluated across "happiness and life satisfaction, health, both mental and physical, meaning and purpose, character and virtue and close social relationships."[10] We conceive of these indicators as being influenced by various systems through which a person navigates their life course. On this conception, flourishing is intimately connected to the well-functioning of technical, social and environmental systems for supporting life, such that disadvantage is largely understood as the *consequences* of an absence of human flourishing.

Information feedback

Information feedback refers to a pivotal component of systems where information resulting from some action in a given part of a system, travels through that system and eventually returns in some form to its point of origin, *potentially* influencing future action. This influence is potential rather than mandatory because system action does not necessarily depend on new information. Information feedback is a separate concept from causal feedback (see circular causality).

Linear and non-linear systems

The terms linear and non-linear systems are borrowed from systems engineering. In systems engineering the input of some systems lead to proportional outputs. If an input is increased by x-amount, the output will increase by x-amount and vice versa. In contrast some systems – typically complex social and environmental systems are non-linear. Inputs into systems often result in disproportional outputs. If an input is increased by an x-amount, the output might be much larger than x or much lower than x or sometimes alternates between larger, same or lower.

---

[6] Alan Hayes and Andrew Hacker, "Persistent Disadvantage in Australia: Extent, Complexity and Some Key Implications," Australia's Welfare Series (Canberra: Australian Institute of Health and Welfare, 2017), https://www.aihw.gov.au/getmedia/9592571c-801c-46be-9c9d-75d0faffbb5b/aihw-australias-welfare-2017-chapter1-6.pdf.aspx.

[7] "Homepage | Australian Institute of Family Studies," Australian Institute of Family Studies, 2023, https://aifs.gov.au/.

[8] David Gruen, "Realising the Potential of Data in Government," Australian Bureau of Statistics, 2022, https://www.abs.gov.au/about/our-organisation/australian-statistician/speeches/realising-potential-data-government.

[9] Australia: Funding the Australian bureau of statistics to better collect data on disadvantage. (2022, Sep 08). MENA Report Retrieved from https://virtual.anu.edu.au/login/?url=https://www.proquest.com/wire-feeds/australia-funding-australian-bureau-statistics/docview/2711767530/se-2

[10] Tyler J. VanderWeele, "On the Promotion of Human Flourishing," Proceedings of the National Academy of Sciences 114, no. 31 (August 2017): 8148–56, https://doi.org/10.1073/pnas.1702996114.

Lived experience experts

We use the expression lived experience experts to denote people who have had lived experience navigating support systems as a vulnerable person or a person who has experienced or is experiencing disadvantage.

Rich data

In this report, we have used rich data to describe datasets that are useful or potentially useful for social intervention purposes. Highly complex and sensitive data sets that are always expanding in both breadth and depth are rich in the sense that they contain many and increasing variety of attributes to continue to add breadth and depth.

EXECUTIVE SUMMARY

CSIRO's 2022 'Our Future' report reemphasised that Australia and the world are at a pivotal point in the ongoing complex and interrelated systemic challenges of climate change, public health, automation, labour, education and others[11]. With the prevalence of machine learning applications, one sustained effort towards managing these complexities has been the deliberate and concerted effort to utilise complex and linked datasets in solutions to these grand challenges which, for Australia, includes entrenched disadvantage. The 2023 Commonwealth Government's 'Data and Digital Strategy' agrees that extreme system dynamics 'have supercharged the adoption of data and digital technologies across Australia' and that data presents a 'wealth of opportunities' for delivering 'services to provide better outcomes for all people'[12]. In 2015, the Committee for Economic Development of Australia (CEDA) made a similar case when it called disadvantage "a wicked problem for any society" and wrote in their 'Entrenched Disadvantage' report that "poverty and disadvantage are experienced by many people at some point throughout their lives, but the issue of entrenched disadvantage has not had the focus it deserves."[13] Close to a decade later, entrenched disadvantage continues to be a key challenge for the Australian Commonwealth Government, which announced in the 2023 Budget, a '$200 million package to target entrenched community disadvantage.'

In agreement with the 2015 CEDA report, the 2023 budget announcement centred data as the key to a comprehensive solution to entrenched disadvantage and lamented that community change-makers were 'hampered by a lack of relevant, available data' or 'difficulties in accessing suitable data across health, education, employment and security.'[14]

The literature shows that it's common for grand data initiatives to centre or focus on the data often without a detailed consideration of how these data projects and products might interact with other systems not considered in the design process. Considering this, we are proposing that decentring data helps make explicit the ways that datasets and their related processes and resources affect and are affected by many other systems. By drawing the boundary more broadly around what is being analysed in data projects, it is possible to see the ways that the design and implementation of data and computation technologies reveal themselves as being contingent on human values, on time and circumstances, on locations, interactions, and the events experienced by the individuals who create, maintain and use it.

Through our project, we are presenting another approach to responsible data practice that counterintuitively decentres data as we explore other productive ways for data professionals to account for the dynamic nature and consequences of systemic influences on data projects that are geared towards disadvantage interventions. Our focus on decentring data means that data, as crucial as it is, cannot be our prime focus. This decentring approach invites data professionals to become aware of interactions between a set of systems where data systems are only one of those systems. It does not devalue the relevance of data but places it within the context of larger conversations about other systems and concepts that also matter.

Recognising the ever-evolving ethical, technical, societal, and environmental systems surrounding these data assets, we firmly believe in the importance of adopting a systems

---

[11] CSIRO, "Our Future World" (CSIRO), accessed October 4, 2023, https://www.csiro.au/en/research/technology-space/data/our-future-world.

[12] "Introducing the Data and Digital Government Strategy | DDGS," accessed October 4, 2023, https://www.dataanddigital.gov.au/.

[13] "CEDA - Addressing Entrenched Disadvantage in Australia," Kentico, 2015, https://www.ceda.com.au/ResearchAndPolicies/Research/Population/Addressing-entrenched-disadvantage-in-Australia.

[14] "Introducing the Data and Digital Government Strategy | DDGS."

approach and cultivating a high level of reflexive disposition among data professionals. This approach is essential in not only mitigating or minimising harm to vulnerable individuals but in supporting those individuals on their journey towards human flourishing ends.

Our project was founded on the view that data professionals working in very dynamic and complex data environments can help us identify current best practices in collecting, managing, and using data for social good, especially in situations with unclear or competing policies and regulations. The complex data systems we are focusing on in this project are the Multi-Agency Data Integration Project (MADIP)[15] - an effort on the part of federal government to better leverage the administrative data sets collected in the course of providing public services; and Generation Victoria (GenV), a longitudinal, life-course health study of Victorians born in a particular pre-defined 'generation' (birth window), following a cohort from birth to death.

We have been focusing on these two data systems because they are highly complex and are visibly interacting with various other complex systems across a variety of domains. They offer a rich linked-data system within which to explore different kinds of data use and data decision-making that has greater complexity than traditional longitudinal studies. As they both collect data on dynamic, changing subjects, in dynamic, changing contexts, they are inherently high-stakes and therefore also incredibly valuable and useful in a range of settings while also needing safeguards against the potential for misuse or inadvertent harm over time. While examples of data systems like MADIP are many worldwide, GenV is unique and a trail blazer in terms of the extent to which it seeks to provide a holistic data ecosystem for health and life outcomes[16]. Lastly, the analysis of data assets like MADIP and GenV are particularly well suited to cybernetic concepts and approaches especially given they have been purposefully designed with higher levels of data sophistication and maturity.

In this research, we are using cybernetics in two ways. Firstly, we are using it as a tool for articulating the variety of systems and their interactions with MADIP and GenV. Secondly, we are conceptualising a cybernetic reflexivity tool to understand and explore opportunities to improve data decision-making in the context of wicked problems such as disadvantage interventions. Our project has been scoped over several phases, with this initial phase dedicated to understanding best practices in decision-making within complex, data-rich environments within uncertain legal and policy settings.

In this initial phase, we have been assessing the presence of a culture of cybernetic reflexivity within the rich-data environment as demonstrated by data professionals from various sectors and positions. These sectors are the Commonwealth Government departments (ABS, Department of Education, Department of Services), research institutes and universities. We have been looking for tangible evidence of this reflexivity through the data decision makers' awareness of key system components that are crucial in the realm of complex data systems. We have achieved this, by applying our novel PAFCARSS method to identify signs of cybernetic reflexivity through activities such as boundary selection, recognition of causal loops, and the identification of opportunities for information feedback. These activities provide us with

---

[15] In May 2023, the Australian Statistician, Dr David Gruen announced that in recognition of what the MADIP data asset represents, the asset's name would transition to PLIDA (Person-Level Integrated Data Asset). There is a phased transition towards PLIDA branding and MADIP will continue to be used in tandem.
[16] Melissa Wake, Sharon Goldfeld, and Andrew Davidson, "Embedding Life Course Interventions in Longitudinal Cohort Studies: Australia's GenV Opportunity," *Pediatrics* 149, no. Suppl 5 (May 2022): e2021053509R, https://doi.org/10.1542/peds.2021-053509R; Melissa Wake et al., "Integrating Trials into a Whole-Population Cohort of Children and Parents: Statement of Intent (Trials) for the Generation Victoria (GenV) Cohort," *BMC Medical Research Methodology* 20 (September 24, 2020): 238, https://doi.org/10.1186/s12874-020-01111-x; Yanhong Jessika Hu et al., "1059Innovative Epidemiological Methods in a Whole-of-State Cohort of Children and Parents: Generation Victoria (GenV)," *International Journal of Epidemiology* 50, no. Supplement_1 (September 1, 2021): dyab168.290, https://doi.org/10.1093/ije/dyab168.290.

insights into the extent to which data itself was decentred in the context of rich data approaches to interventions addressing disadvantage. Additionally, they reveal opportunities for enhancing system-wide improvements through systems mapping workshops, education and training.

In examining the reflexive attitudes of data professionals, their views broadly fall in to one of three thematic areas. We explore attitudes towards disadvantage, legitimacy and trust and towards their experiences navigating data access and data asset resource allocation. In each thematic area, evidence of system visibility and cybernetic reflexivity have readily been observable. These observations are helping us to develop a series of practical lessons for each thematic area. Our main research findings for each thematic area and their associated practical lessons are detailed in full in Appendix 1 and in the discussion section. In the following sections, we provide our notable research findings and their related practical steps forward.

Data professionals' attitudes towards disadvantage

We found that definitions of disadvantage were often pragmatic, focusing on quantifiable factors and minimal examples of cybernetic reflexivity were observed. Even so, reflexivity was more prevalent during the initial project stage in which data options were being considered but much less evidence for reflexivity in the final stages and post-production.

Data professionals often admitted openly that they lacked in-depth knowledge of the disadvantage literature and there was general agreement that involving lived-experience experts[17] in modelling disadvantage was valuable. We propose that lived-experience experts could play a pivotal role in establishing information feedback mechanisms that could improve data interventions. Moreover, it was clear that disadvantage data projects that involve a community of lived-experience experts throughout the duration of those interventions were likely to create mutually beneficial data-driven or data-informed interventions.

Our interview analysis revealed that asset-framing or strength-based data approaches were not widely practised even though data professionals expressed a strong desire to learn asset-framing and strength-based data approaches to disadvantage.

Practical lessons for addressing attitudes to disadvantage

1. The data community will benefit from creating widely agreed definitions of disadvantage for each domain while making sure to canvass diverse perspectives.
2. If senior data decision-makers align themselves with high-level system purposes, they can increase the uptake of reflexive practice by other data workers.
3. Identifying and aligning relevant datasets to asset-framed objectives will support data professionals to cultivate reflexive data approaches to disadvantage.
4. Charities and support systems could shift toward human flourishing goals when adopting a systems view.
5. Funding should support frontline workers' data skills to enhance the deployment of interventions that have been approved by lived-experience experts.
6. Data professionals expressed strong desire for development programs that give them opportunities to interact with systems involving disadvantaged communities.

Data professionals' attitudes towards legitimacy and trust

We found that data professionals' attitudes toward legitimacy and trust manifested most strongly through data collection methods and during the development of consent tools.

[17] See our 'key-terms' section for greater detail on 'lived experience experts'.

The format of consent forms for rich data projects were strongly dependent on whether data collection was to be largely automated or through physical interaction.

Consent instruments that embraced systems approaches tended to consider long-term viability of those instruments.

Our research analysis revealed lower than expected instances for revising consent forms. In a dynamic environment with changing privacy regulations and changing data subjects, interviewees agreed that greater opportunities for revising consent forms is a good idea.

Risk orientation and risk communication heavily influenced decision-making and trust within all rich data systems. We found that the need to satisfy and strengthen safeguarding measures were front of mind.

Effective data sharing was hampered by differing risk orientations based on mismatches between foundations and missions of each data donor or asset manager but this was often not explicitly acknowledged within data sharing discussions.

Practical lessons for addressing legitimacy and trust

1. Dynamic consent management was recommended to ensure long-term viability.
2. Active consent management could result in the establishment of ongoing participant support services; such as consent counselling as young research participants mature to consent age.
3. Our analysis revealed that it was critical that data decision makers consider data asset management approaches that are adaptive to shifting norms around data ownership in both private and public spheres.
4. Transparency and collaboration could improve legitimacy and trust.
5. Cybernetics can suggest some reflexive strategies in support of triage processes surrounding the disclosure of medical results while planning bespoke responses to the needs of vulnerable data subjects.
6. Our meta-analysis suggests data donors and asset managers could engage in collaborative data systems mapping exercises to increase transparency between themselves.
7. Workshops that illuminate causal pathways and risk communication channels can open opportunities for risk averse decision-makers to consider economic and service repercussions if their organisations do not share critical public data.

In summary, our reflexivity through systems visibility approach reveals that data risk manifests as the responsibility not only to safeguard but also as a responsibility to make data useful for social good. This finding suggests opportunities for the deployment of cybernetic tools that can supplement data professionals' data pipelines with opportunities for mapping these future facing systems dynamics.

Attitudes towards data system resourcing and access rights

We found that the identification of diverse data user groups and their objectives by asset managers improved access to datasets held by those data assets.

Sustained long-term funding for rich-data projects was a shared concern with some data professionals suggesting that the situation could be improved if funders were continuously made aware or reminded about the interconnected funding needs of all components of complex data assets.

Practical lessons for addressing data system resourcing and access

1. Consolidating data asset information and involving user groups in the process can enhance understanding of the complexities of data assets for new data-asset users.
2. Collaboration among data assets could lead to standardized formats and faster access processes.
3. Training and awareness-raising around complex AI assisted data techniques are essential for data asset managers' ability to manage effectively and efficiently, the access to data assets.

Conclusion

We have relied on our novel cybernetic framework to assess data professionals' reflexive decision-making within data environments aiming to build interventions for vulnerable communities. The findings suggest opportunities for conducting training and awareness workshops to show case how cybernetic reflexivity can promote better definitions of disadvantage, adoption of asset-framed approaches, enhancement of legitimacy and trust, and the optimisation of data system resourcing and access rights.

Our findings show that there is willingness among those working with data for disadvantage to embrace reflexive improvements aimed at fostering safe, responsible, and sustainable practices. The broad lessons outlined in this report can serve as guiding principles for the design of comprehensive tools spanning the variety of disadvantage data domains. These tools can aid in enhancing both systems awareness and the practical implementation and expansion of reflexive practices in the data-driven and data-informed initiatives. Our ultimate goal continues to be to equip data professionals with the necessary tools to reduce immediate and future harm to data subjects and steer data projects toward outcomes that promote human flourishing.

While our findings revealed limited evidence for decentred approaches to data projects and nascent reflexivity among data professionals associated with data assets and disadvantage-focused projects, we did uncover evidence of their strong desire to cultivate this reflexivity. This desire emerged especially within environments characterised by political, legislative, and regulatory opacity. Our research shows there is support for programs designed to support data professionals to cultivate reflexive approaches to navigating the intricate dynamics of highly complex systems, where various systems and agents interact within technological environments.

INTRODUCTION

Data solutions to wicked social problems and its limitations

"The 21st century is being defined by complex, interrelated, systemic challenges: climate change, public health, automation, labour, education and radicalisation."[18] Indeed we live in volatile times that are being exacerbated by rapid evolution and ubiquity of digital technologies. Invariability, these evolving digital technologies are also ushering in novel and increasingly dynamic ways of capturing, storing, transforming, sharing and using data. Within our current complex and dynamic data environments that are being impacted with other complex systems undergoing change, there is inevitably a greater risk of data mishandling and misuse. This is occurring at a time when there are growing calls for Australian society to realise the value of data-driven and data informed approaches across the for-purpose sector[19].

In 2015, the Committee for Economic Development of Australia (CEDA) called disadvantage "a wicked problem for any society" and wrote in their 'Entrenched Disadvantage' report that "poverty and disadvantage are experienced by many people at some point throughout their lives, but the issue of entrenched disadvantage has not had the focus it deserves."[20] According to CEDA, the key to pulling people out of entrenched disadvantage was "early intervention programs that can break the cycle of poverty."[21]

In the report, CEDA proposed that one key to discovering and developing viable solutions for disrupting disadvantage was to collect more data. They recommended that understanding the dynamics of disadvantage better required longitudinal data, founded on admin data held by government departments, and more in-depth research. The report recommended that a strong data approach would help identify those suffering disadvantage as they navigate "multiple touchpoints with government services and service providers, including social security, housing, justice and child protection."[22] They proposed that collecting more data could both improve the visibility of disadvantaged people at these touchpoints, and help to significantly improve the navigation of government services by those experiencing disadvantage.

CEDA's report series reflects a broader agenda of government and research institutions to increasingly invest in the collection of new forms of rich, purposefully representative data about people in pursuit of better health and social outcomes.[23] These rich and granular data sources are frequently considered a trustworthy input for new kinds of research and analysis, including for machine learning. These rich data sources become a spine for a range of research, policy proposals and targeted interventions designed to mitigate disadvantage, improve wellbeing and lift people out of poverty and other expressions of disadvantage. As big data has become more prevalent, so has the desire to harness data for social good.

---

[18] Ellen Broad, "We Need to Let Go of Regulating 'Artificial Intelligence,'" *InnovationAus.Com* (blog), August 8, 2022, https://www.innovationaus.com/we-need-to-let-go-of-regulating-artificial-intelligence/.

[19] Australian Government, "Australian Data Strategy: The Australian Government's Whole-of-Economy Vision for Data," October 1, 2022, https://www.finance.gov.au/sites/default/files/2022-10/australian-data-strategy.pdf.

[20] "CEDA - Addressing Entrenched Disadvantage in Australia."

[21] "CEDA - Addressing Entrenched Disadvantage in Australia."

[22] CEDA, "Disrupting Disadvantage: Setting the Scene.," Trove, 2019, 9, https://nla.gov.au/nla.obj-2056065312.

[23] "Australia's Welfare Indicators - Australia's Welfare Indicators," Australian Institute of Health and Welfare, 2022, https://www.aihw.gov.au/reports-data/indicators/australias-welfare-indicators.

Particularly since data-driven responses were adopted widely by governments during the COVID19 pandemic, using complex linked data approaches as a means to drive real change has become the goal for an increasing number of stakeholders in policymaking and service design in Australia. This has increased interest and investment in rich data projects like those discussed in this report, placing data at the centre of the current debate around disadvantage. Best practice data use has the potential to be an important tool in disrupting disadvantage, but as we seek to demonstrate in this report, it must be contextualised within broader considerations and action.

Many data workers would agree with the broad agenda of government and institutions to invest in rich data collections and they are often passionate about making a real difference through data, to the lives of those experiencing economic and social disadvantages[24]. Simultaneously, many authors have given accounts of the messy and unpredictable nature of data and the influential role data professionals play in what data is collected and how that data is translated and made meaningful[25]. Secondly, data professionals recognise that applications of data for various uses is fraught with risks of causing harm especially when the data is about vulnerable populations or those populations at a disadvantage. Ensuring these risks are minimised or mitigated are top priority partly because data privacy and protection are fundamental human rights and data decision makers have the social responsibility to make sure this right is protected particularly for individuals experiencing disadvantage[26]. Some of these data harms may include identity theft, discrimination, or targeted scams while exploitative risks can include data profiling, discriminatory algorithms, or invasive data collection[27].

The repercussions of harms and risks to those experiencing disadvantage might further result in limited opportunities for those experiencing disadvantage which could result in further marginalisation from economic and social benefits. Moreover, partly due to often negative experiences those experiencing disadvantage have with government agencies and law enforcement, trust of these agencies or people associated with these agencies is delicately balanced[28]. If these members of community perceive that their data is not handled responsibly, they may be less likely to cooperate with and give consent to the collection and use of their data with these agencies and organisation[29].

In this very dynamic time in history and in rapidly evolving data ecosystems, often data professionals are required to interact with various visible and invisible systems simultaneously as they navigate ways to utilise new forms of data in safe, responsible and sustainable ways. We have also outlined how, concerted effort in safe handling, collecting, storing and using data from

---

[24] data.org, "Workforce Wanted: Data Talent for Social Impact," *Data.Org* (blog), accessed November 22, 2022, https://data.org/reports/workforce-wanted/.

[25] Lyria Bennett Moses, Kylie Valentine, and Janet Chan, "Data Practices in a Web of Values: Reflections on the Gap between Ethical Principles and Data-Driven Social Policy," SSRN Scholarly Paper (Rochester, NY, October 1, 2021), https://papers.ssrn.com/abstract=4054085.

[26] OAIC, "What Is Privacy?," OAIC, March 10, 2023, https://www.oaic.gov.au/privacy/your-privacy-rights/your-personal-information/what-is-privacy.

[27] Zoe Staines et al., "Big Data and Poverty Governance under Australia and Aotearoa/New Zealand's 'Social Investment' Policies," *Australian Journal of Social Issues* 56, no. 2 (2021): 157–72, https://doi.org/10.1002/ajs4.129.

[28] Loïc Wacquant, *Punishing the Poor: The Neoliberal Government of Social Insecurity* (duke university Press, 2009); "To Create Fairness and Accountability in the Use of Government Decision Making Algorithms - Churchill Trust," accessed August 29, 2022, https://www.churchilltrust.com.au/fellow/dr-owen-churches-sa-2018.

[29] S. McKay, "Poverty or Preference: What Do 'Consensual Deprivation Indicators' Really Measure?," *Fiscal Studies* 25, no. 2 (2004): 201–23, https://doi.org/10.1111/j.1475-5890.2004.tb00102.x.

lived experience experts contributes to societal gains of building trust, reducing harm, and advancing social and ethical values in data-driven and data-informed products and decision-making processes. This requirement to interact with various systems simultaneously has necessitated data decision makers and data workers to explore a systems approach to their data projects. One of the motivating factors for this is, a systems view has the potential to ensure a holistic approach to data ethics and responsible data practices that consider the data environment as a whole while creating space for the sensitivities required for utilising the data of those people experiencing various disadvantages. Through these systems views, the datasets or data assets themselves are decentred as the key agenda becomes how to best maintain the reliability of data systems now and in to the future in a way that is safe, responsible and sustainable.

In these times of rapidly evolving data systems and processes, flexibility, adaptability, and a strong commitment to safety of data subjects are essential. Data professionals engaged in the for-purpose sector should be proactive in identifying and addressing risks as they arise and should be prepared to adjust strategies and practices accordingly to maintain trust and compliance even in the current policy and regulatory environment that is sparse, opaque and not customised and sometimes not customisable. "Change will come unevenly, iteratively and distributed across a broader system of regulation and expertise."[30] The challenge is how to (or what approaches to use to) identify and address data risks individually or as a small collective in such a way that it will be effective across the systems that interact with a given set of data projects or data assets over time. There have been numerous suggestions for reducing risky approaches for utilising data for disadvantage intervention[31]. In all these approaches, we have identified that they often centre data as they try to account for the impacts of systems on data projects. By centring data, we are not meaning the technical procedure of correcting datasets around the mean of a given data set. In saying 'approaches often centre the data', we mean it in the social science sense of the term that points to making datasets the central focus of projects[32].

Through our project, we are presenting another approach to responsible data practice that counterintuitively decentres data as we explore other productive ways for data professionals to account for the dynamic nature and consequences of systemic influences on data projects that are geared towards disadvantage interventions. Our focus on decentring data means that data, as crucial as it is, cannot be our prime focus. This decentring approach invites data professionals to be aware of interactions between a set of systems where data systems are only one those systems. It does not devalue the relevance of data but places it within the context of larger conversations about other systems and concepts that also matter. For instance, when a decentring attitude to data is embraced, it can open an opportunity to explore in greater focus, the data concerns and risks that emanate from challenges associated with identifying and defining who is disadvantaged or vulnerable. This consideration is important because as we outline below, focussing on the use of broader definitions of and indicators for disadvantage is a

---

[30] Broad, "We Need to Let Go of Regulating 'Artificial Intelligence.'"

[31] Bennett Moses, Valentine, and Chan, "Data Practices in a Web of Values"; Abeba Birhane, "Algorithmic Injustice: A Relational Ethics Approach," *Patterns* 2, no. 2 (February 12, 2021): 100205, https://doi.org/10.1016/j.patter.2021.100205; Joanne Luke et al., "Questioning the Ethics of Evidence-Based Practice for Indigenous Health and Social Settings in Australia," *BMJ Global Health* 7, no. 6 (June 1, 2022): e009167, https://doi.org/10.1136/bmjgh-2022-009167; Kord Davis, *Ethics of Big Data*, 1st edition (Sebastopol, California: O'Reilly, 2012); Daniel W. Tigard, "Big Data and the Threat to Moral Responsibility in Healthcare," in *Datenreiche Medizin und das Problem der Einwilligung: Ethische, rechtliche und sozialwissenschaftliche Perspektiven*, ed. Gesine Richter et al. (Berlin, Heidelberg: Springer, 2022), 11–25, https://doi.org/10.1007/978-3-662-62987-1_2.

[32] See this paper by John Horton who speaks about centring reflexivity in creative research. John Horton, "Centring Reflexivity, Positionality and Autoethnographic Practices in Creative Research," in *Creative Methods for Human Geographers*, ed. Nadia von Benzon et al. (SAGE Publications Ltd, 2021).

critical opportunity in data projects addressing disadvantage, because definitions tend to shape ways of measuring, analysing and evaluating people and interventions for people deemed disadvantaged.

Challenges with identifying and defining the disadvantaged

In reviewing the literature on opportunities and challenges for using data to address disadvantage, we found that approaches to using data for disrupting disadvantage can simplify disadvantage as a sub-set of identities or indicators, which are not helpfully representative of the individuals, peoples, and experiences of disadvantage that we would like to serve and support out of disadvantage[33]. Pragmatically-oriented research on disadvantage tends to continue to emphasise disadvantage as a disabling form of 'lack'. For example, CEDA's stance on disadvantage as an economic status, equivalent to "sitting below the poverty line", has long been questioned in academic discourse, challenging the framing of disadvantage as predominantly being economically defined. Academics and those with lived experience of disadvantage have pointed out the way in which this kind of definition can pejoratively label those who are experiencing or have experienced economic or social disadvantage. Key academic studies have helped shed light on ways in which disadvantage is far more complex than being classified within an economic bracket in a particular time. To a large extent, CEDA's reports give evidence for this complexity of disadvantage citing for example that "only a small subset experiences poverty spells lasting more than a decade, some are never at risk of long-term disadvantage, while others are at risk of falling in and out of disadvantage through their life course."[34] Leading academics in the field have suggested that the parameters defining a state of disadvantage is too narrow, and should extend to joblessness, which has been acknowledged to involve a variety of complex causal factors too[35].

There is a growing movement to redefine disadvantage in ways that contest 'deficit framing' – this focus on economic and social lack. Some authors advocate that deficit framing can be harmful due to a given language's "constitutive effects"[36] – they help form identities and other key parts of society[37]. The label of 'disadvantaged' can potentially exacerbate social disconnection rather than reconnection and social cohesion. The words we use matter not only because it could affect people's feeling but because they influence the formation of the type of people they become in society. As an example, McDonald and Chenoweth[38] showed that by Centrelink moving to using "New Public Management" language of referring to clients as Customers rather than clients, they were inadvertently removing rights of those people whether intentional or unintentionally. According to McDonald and Chenoweth, "… the social workers in Centrelink are participating in the reconstruction of services users' identity – from that of rights – bearing citizens with legitimate claims on the state, to an enfeebled form of 'customer'. Centrelink 'customers' are not sovereign consumers with sufficient resources to purchase

---

[33] "Developing longitudinal qualitative research into the dynamics of poverty could enable exploration of the complex factors influencing individual and household transitions, and it may provide insight into the perspectives of those experiencing poverty as an everyday reality" "CEDA - Addressing Entrenched Disadvantage in Australia," 47.

[34] "CEDA - Addressing Entrenched Disadvantage in Australia," 19.

[35] "The term 'jobless' simply recognises a common factor among poor households but it tells us little about the complex causal factors underpinning these experiences…" "CEDA - Addressing Entrenched Disadvantage in Australia," 46.

[36] Catherine McDonald and Lesley Chenoweth, "(Re) Shaping Social Work: An Australian Case Study," *The British Journal of Social Work* 39, no. 1 (2009): 154.

[37] McDonald and Chenoweth, 154.

[38] McDonald and Chenoweth, 154–55.

goods and services…they are, for the most part, impoverished people in an imbalanced power relation with a system which has the ability and the will to force compliance." As this quote exemplifies, terms like disadvantage, vulnerable and even excluded are necessarily negatively loaded and help to hide power imbalances between those defined as disadvantaged and those making laws, policies and service delivery guidelines. Moreover, the use of references like 'the disadvantaged' not only refers to a particular person but also contributes to creating an identity of the disadvantaged. This identity framing then leads to many other problematic terminology such as educational disadvantage[39], social mobility disadvantage[40], disadvantage as social exclusion[41], socio-economic disadvantage[42] and many others.

More recently, an influential set of ideas on disadvantage is identifiable, especially building on the work of economist Amartya Sen and political philosopher Martha Nussbaum. Their work — and that of their colleagues — define disadvantage in terms of not having and/or being able to exercise the following:[43]

i. Life (normal span)
ii. Bodily health (incl. food, shelter)
iii. Bodily integrity (incl. freedoms from and to)
iv. Sense, imagination, and thought (incl. self-expression, faith)
v. Emotions (emotional health)
vi. Practical reason (critical thinking)
vii. Affiliation (human relationships)
viii. Other species (live in relation to and concern for nature)
ix. Play (incl. enjoyment)
x. Environmental control (incl. political participation, work)

While these dimensions present a more complete and holistic account of what it is like to experience disadvantage, the tendency in data approaches remains to focus on an absence of these conditions rather than to focus on what capabilities those facing disadvantaged already have, and what can be done to support those capabilities, and equally what conditions are endemic in the systems in which so-called disadvantaged people find themselves that are influencing their outcomes.

Minimal support for data workers and decision makers

the literature shows that data professionals have few options for training in decentring approaches to disadvantage data projects that can help them combat potential risks that accompany the use of data sets related to the disadvantaged. Additionally, the persistent and dynamic nature of the data environment has made it difficult if not impossible to create broad and fixed principles for data professionals to apply to disadvantage projects across varieties of

[39] Laura Perry, "Educational Disadvantage Is a Huge Problem in Australia – We Can't Just Carry on the Same," The Conversation, 2017, http://theconversation.com/educational-disadvantage-is-a-huge-problem-in-australia-we-cant-just-carry-on-the-same-74530.

[40] "Paul Ramsay Foundation Supports ANU to End Disadvantage," ANU (The Australian National University, April 29, 2022), https://www.anu.edu.au/giving/impact-stories/paul-ramsay-foundation-supports-anu-to-end-disadvantage.

[41] "Our Social Exclusion Measure," Brotherhood of St. Laurence, 2020, https://www.bsl.org.au/research/our-research-and-policy-work/social-exclusion-monitor/social-exclusion-measure/.

[42] Rosanna Scutella, Michael Horn, and Roger Wilkins, "Measuring Poverty and Social Exclusion in Australia: A Proposed Multidimensional Framework for Identifying Socio-Economic Disadvantage," *SSRN Electronic Journal*, 2009, 3, https://doi.org/10.2139/ssrn.1639867.

[43] Jonathan Wolff and Avner De-Shalit, *Disadvantage* (Oxford: Oxford University Press, 2007), 33–39.

sectors such as health, law enforcement, legislature, policy and the for-purpose sector. Even so, data professionals working with disadvantage data try to mitigate negative repercussions through continuous upskilling in data ethics and guidelines around responsible data practices[44].

The data decision makers and data workers themselves are sometimes encouraged to embrace an agile mindset that allows them to quickly adapt to a variety of changing ethical and regulatory landscapes[45]. Indeed there has been a rise in data approaches that foreground values and histories[46]. These approaches attempt to make explicit the messy relational and dynamic aspects of data, as useful elements to support data decision makers understand the connections between the datasets they use and broader systems like academia, government, services and policy[47]. These practises aim to reveal how, why, and by whom data was compiled – as part of making clear its limitations. Many of these approaches focus on what relational and dynamic aspects can be gained from static data set properties (records of provenance, for example, included in metadata). However, they still centre the properties of a data set to be used, and not the actions of a data decision maker, who also needs to account for how their own actions (from data compilation and manipulation, data design, through to data use and the contexts of that use) are shaping its effects. Ideally, data decision makers and workers would also centre themselves and their actions as part of the data system – not external to it - within which they are trying to effect change. We are interested in outlining a decentred approach to data and disadvantage projects. That is, in this project we have sought to understand how clearly rendering a broader context, of which a data set is only a part and only a partial representation of, may support more reflexive practice as a core element of a responsible data practice.

What is reflexivity?

Linda Finlay is a well-known reflexivity researcher who has written numerous guides which seek to clarify and make reflexivity accessible. According to her, "[r]eflexivity in research can be described as the use of a critical, self-aware lens to interrogate both the research process and our interpretation or representation of participants' lives in our social world. It's a vehicle that acknowledges the complexity and messiness of our qualitative project."[48]

Reflexivity has given mostly social science researchers a way of examining and evaluating how their "background, assumptions, positioning, behaviour, and subjectivity might impact on the research process and vice versa."[49] Another key reason for why supporters think reflexivity is necessary and useful is that it's thought to encourage all researchers to embrace ethical

---

[44] Tess Johnson, Konrad Kollnig, and Pierre Dewitte, "Towards Responsible, Lawful and Ethical Data Processing: Patient Data in the UK," *Internet Policy Review* 11, no. 1 (March 18, 2022), https://policyreview.info/articles/analysis/towards-responsible-lawful-and-ethical-data-processing-patient-data-uk; Matthew Zook et al., "Ten Simple Rules for Responsible Big Data Research," *PLOS Computational Biology* 13, no. 3 (March 30, 2017): e1005399, https://doi.org/10.1371/journal.pcbi.1005399.

[45] Niina Zuber et al., "Empowered and Embedded: Ethics and Agile Processes," *Humanities and Social Sciences Communications* 9, no. 1 (June 6, 2022): 1–13, https://doi.org/10.1057/s41599-022-01206-4.

[46] Rosanna Scutella, Michael Horn, and Roger Wilkins, "Measuring Poverty and Social Exclusion in Australia: A Proposed Multidimensional Framework for Identifying Socio-Economic Disadvantage," *SSRN Electronic Journal*, 2009, https://doi.org/10.2139/ssrn.1639867.

[47] Birhane, "Algorithmic Injustice"; Timnit Gebru et al., "Datasheets for Datasets" (arXiv, December 1, 2021), http://arxiv.org/abs/1803.09010.

[48] Linda Finlay, "Championing 'Reflexivities,'" *Qualitative Psychology (Washington, D.C.)* 4, no. 2 (2017): 1, https://doi.org/10.1037/qup0000075.

[49] Finlay, 1.

challenges that exist in most research projects especially within those projects involving people and livelihoods[50].

If applied as intended, reflexivity has the potential to encourage self-awareness and critical self-reflection about one's data profession and their decisions all through the data pipeline. Within the literature addressing reflexivity in the data ecosystem, unfortunately, most sources tend to treat reflexivity synonymously with adherence to most often, formulaic or checklist of ethical considerations[51]. Reflexivity differs to these approaches in that it places responsibility on the researcher to see themselves and acknowledge their role in the data production process. We are of the view that data professionals who work with and make decisions about sensitive human data need to engage in reflexivity if their data projects are to be beneficial to promoting human flourishing now and in the future.

A reflexive approach that can be applied within the data science practice is found in Jamieson et al (2023)[52]. They define reflexivity as a process that allows researchers to question who they are as researchers and how their research shapes their worldview and how their worldview shapes their research (see page 2). Supporters have argued that if reflexivity is to be taken seriously within quantitative studies, 'they need to be included in the project protocol and research design from the outset' that doing so would 'add a depth of understanding about how, where, when, and by whom data were collected.'[53] Jamieson et al provide a thorough list of suggested reflexive ideas from pre-research, during data collection, analysis, results and through to post research[54]. Agreeably, they suggest their work is timely as reflexivity across disciplinary areas continue to gain acceptance.

It is within the context of growing acceptance of cross disciplinary reflexivity that our research applied a version of reflexivity that was developed in the 1940s-50s as a result of multidisciplinary debates on the importance of reflexivity in computing, artificial intelligence and information and data theory. These debates occurred during a series of conferences which were later referred to as the Macy Conferences on Cybernetics, after the inaugural book on cybernetics was written by Norbert Wiener and titled 'Cybernetics – or control and communication in the animal and the machine'[55]. As we discuss in the following sections, for these early pioneers of thought around how computing and humanity would interact, reflexivity was not simply a supplementary novelty but a crucial element in sense-making. We see this approach as one of the most appropriate conceptualisations through which to analyse the reflexive practise of today's data professionals engaged in rich complex data projects within uncertain and dynamic legislative and policy environments that have a bearing on the wellbeing of vulnerable people and people at a disadvantage.

---

[50] Finlay, "Championing 'Reflexivities.'"

[51] Foster Provost and Tom Fawcett, *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*, 1st, 1st ed. ed. (Sebastopol: O'Reilly Media, Incorporated, 2013); Davis, *Ethics of Big Data*; Peter C. Bruce, *Responsible Data Science.* (Newark: John Wiley & Sons, Incorporated, 2021); K. J. Whitehair, "Hasselbalch, Gry. Data Ethics of Power: A Human Approach in the Big Data and Al Era," *CHOICE: Current Reviews for Academic Libraries* 60, no. 2 (2022): 182-.

[52] Michelle K. Jamieson, Gisela H. Govaart, and Madeleine Pownall, "Reflexivity in Quantitative Research: A Rationale and Beginner's Guide," *Social and Personality Psychology Compass* 17, no. 4 (2023): e12735, https://doi.org/10.1111/spc3.12735.

[53] Louise Ryan and Anne Golden, "'Tick the Box Please': A Reflexive Approach to Doing Quantitative Social Research," *Sociology* 40, no. 6 (December 2006): 1198, https://doi.org/10.1177/0038038506072287.

[54] See Jamieson, Govaart, and Pownall, "Reflexivity in Quantitative Research."

[55] Norbert Wiener, *Cybernetics or Control and Communication in the Animal and the Machine* (Technology Press, 1948).

Cybernetics

Wiener's book introduced simultaneously a concept, a research field and a transdisciplinary research group. As a concept, cybernetics was developed by Wiener and his team at MIT as an applied science focused on understanding and producing interacting systems comprised of machines and agents that could communicate through information feedback recursively in order to achieve a given objective. The *continuous updating of action* due to information feedback is what is now referred to as *circular causation*. For Wiener and his cybernetics group, this discovery was to lead to a novel approach to mathematical systems engineering of guided missile technology where the most optimum solutions were to be found *only* by treating these processes as "circular processes"[56] or equivalently as exhibiting circular causation. Importantly, these circular processes necessitated that systems had to be considered as a whole since the whole is greater than the sum of its parts. This holistic and circular approach was immediately applicable to a wide variety of systems across various disciplines and it was this discovery that they published in their paper 'Behavior, Purpose and Teleology'[57]. As it turned out, this paper foreshadowed the formation of cybernetics as a discipline and the ideas it contained set the foundation for the cybernetics conferences in the 1940s-1950s that concentrated on the transdisciplinary experimentation with the importance of information feedback (communication) to the task of effective control in machines and animals (including people).

The cybernetic reflexive turn

Although cybernetics (the study and creation of information feedback mechanisms and circular causation in adaptive closed systems) was widely applicable across disciplines, each of those systems studied had to be defined in terms of strict empirical, verifiable and repeatable processes."[58]

Due to the interdisciplinary nature of cybernetics, almost immediately, some of the cybernetic practitioners voiced concern over the strict mathematical and observational approaches to cybernetics because it was increasingly overwhelming to apply it to nonlinear, complex and dynamic systems such as social systems like business corporations and human population growth and the economy. These difficulties culminated in some cyberneticians developing an approach that conceived themselves as a system interacting with the systems they are observing. Those cyberneticians supporting this reflexive approach saw it necessary to consider circular causal interactions resulting on them and emanating from them to the larger systems via information feedback mechanisms.

It was only after the Macy Conferences had ended that Heinz von Foerster, who also participated in the conferences, dedicated himself to developing ways of addressing reflexivity[59]. He developed this new approach in his essays which were collectively called "Observing systems"[60]. The title was a deliberate play on words in which "Observing systems" was making the dual points that observers are systems that are also engaging in observing other systems. According to

---

[56] Wiener, 8.

[57] Rosenblueth, A, N Wiener, and J. Bigelow. "Behaviour, Purpose, and Teleology," Philosophy of Science, 10, 18-24 (1943).

[58] Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*, 10.

[59] Alan Hayes and Andrew Hacker, "Persistent Disadvantage in Australia: Extent, Complexity and Some Key Implications," Australia's Welfare Series (Canberra: Australian Institute of Health and Welfare, 2017), 133

[60] N. Katherine Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* (Chicago, IL: University of Chicago Press, 1999), 133, https://press.uchicago.edu/ucp/books/book/chicago/H/bo3769963.html.

Hayles, by defining the observer as a system, "...it reduced the problem of the observer to a problem of communication among systems."[61]

Thus for this participative approach to cybernetics, reflexivity amounts to the articulation of the circular causal acts the agent sees themselves performing via information feedback mechanisms between themselves as observing systems and other systems of interest[62]. These ideas were developed by von Foerster (1972)[63] where he views the observer as experiencing and interacting with and within their systems environment rather than observing their environment. As each observer experiences and interacts, they are also performing 'computations'[64] that transform, modify, re-arrange, order the very entities in consideration[65].

Since cybernetics sees each person as self-regulating and adaptive closed systems, autonomy and individuality is respected and protected and partly due to these reasons, "cybernetic boundary questions often involve deep ethical and psychological issues"[66] At the same time, the autonomy articulated in cybernetics acknowledges that people are situated in communities and that individual actions can "affect the lives of other human beings and, hence, have ethical significance."[67] On this, Krippendorff acknowledges through Gregory Bateson[68] that, "we always participate in the circuitry of the world. Acknowledging one's participation in a larger system, whether as explorers, designers, or constituents of social formations, is a reflexive turn that reveals reality, its parts, and the self as interactively or dialogically constructed, and admits that individual knowledge is necessarily incomplete, expandable with efforts."[6970]

In this reflexive view of cybernetics which we are calling cybernetic reflexivity, constructive participation is based on an understanding that each person is able to conceive of ideas independently from others in the team (autonomy). Secondly, that constructive and participative reflexivity requires the ability to make choices about data and be accountable for those choices (agency). Within this reflexive framework, it is the agents' recursive application of new information via feedback mechanisms that enables agents to affect change through circular causality[7172].

Decentring data and centring cybernetic reflexivity through systems visibility

In summary, we live in a world that is undergoing complex, interrelated, systemic challenges that include climate change, public health, automation, labour, education among others. One approach

---

[61] N. Katherine Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* (Chicago, IL: University of Chicago Press, 1999), 134, https://press.uchicago.edu/ucp/books/book/chicago/H/bo3769963.html.

[62] See H. von Foerster, "On Self-Organizing Systems and Their Environments," in *Understanding: Essays on Cybernetics and Cognition*, ed. Heinz von Foerster (New York, NY: Springer, 2003), 1–19, https://doi.org/10.1007/0-387-21722-3_1.

[63] Heinz Von Foerster, "Notes On An Epistemology For Living Things," *BCL Report #9.3*, 1972, https://semiorganized.com/resources/other/foerster/epistemology.pdf.

[64] Computation is used here broadly in the sense used by von Foerster to denote abstract or physical changes caused by an observer as they act implicitly or explicitly on a system.

[65] Foerster, "Notes On An Epistemology For Living Things," 7.

[66] Hayles, *How We Became Posthuman*, 141.

[67] Hayles, 142.

[68] Daniel Foxman and Gregory Bateson, "Steps to an Ecology of Mind," *The Western Political Quarterly* 26, no. 2 (June 1973): 345, https://doi.org/10.2307/446833.

[69] Klaus Krippendorff, "Cybernetics's Reflexive Turns," January 1, 2008, 178.

[70] This participative and reflexive approach to developing knowledge products was outlined by William Ross Ashby, *An introduction to cybernetics*, 1961.

[71] See Krippendorff, "Cybernetics's Reflexive Turns," 182–83.182-183.

[72] William Ross Ashby, *An introduction to cybernetics*, 1961.

to managing this complexity has been the deliberate and concerted effort to utilise complex and linked datasets to address wicked challenges such as entrenched disadvantage. We discovered that it was common for these big data approaches to centre or focus on the data often times without acknowledging and/or designing for the inclusion of equally important systems that are interacting with those data systems. In light of this discovery, we proposed that decentring data helps make explicit the ways that datasets and their related processes and resources affect and are affected by many other systems. By drawing the boundary more broadly around what is being analysed in data projects, it is possible to see the ways that the design and implementation of data and computation technologies reveal themselves as being contingent on human values, on time and circumstances, on locations, interactions, and the events experienced by the individuals who create, maintain and use it.

As Genevieve Bell writes, 'When focused on the system and what it is trying to achieve, we can improve our understanding of what additional data are needed and where datasets can be linked to help us determine points where we can intervene in social support systems and leverage groups from disadvantage. […] We must understand data as part of a system that has more complex inputs and outputs and interactions than simply those captured in a dataset'.[73]

Our approach to decentring data is what we have called cybernetic based reflexive data use and data decision making. We outlined how one reason for anchoring our analysis on cybernetic reflexivity is based on its origins as a concept developed and used by those who understood and developed the very concepts of computer information and data – they understood both the technical importance of information and data systems as well as the impact the data professional's agency and autonomy can have on interacting systems comprised of machines, data, agents and the environment. We decided that cybernetic reflexivity's historical and theoretical connection to the data concept together with its reliance on the identification of the features of systems is a practical way of thinking about decentring of data and enabling system visibility. So in this project, we have reimagined cybernetic reflexivity as a tool for analysing the reflexive practise of data professionals working with rich and complex datasets and situated within uncertain and dynamic legislative and policy systems that have some bearing on the wellbeing of people experiencing disadvantage.

RESEARCH FOCUS

For this research, we are taking a cybernetic reflexivity approach by decentring data and looking at a broader system of analysis. We look at this system in two key ways. Firstly, the system has a purpose, or multiple purposes, which shape how the system behaves and the ways we see the system. Secondly, we have drawn the systems' boundaries in such a manner as to enable the observation of circular causal processes and the locations and instances of information feedback mechanisms within our system of analysis. We have sought to understand the ways that data-workers and decision-makers interface with government, academia, and non-government actors. We believe that looking at data assets in a broader context, requires thinking about how these actors interact, and in some instances coordinate efforts to collect and use data to inform their behaviours and approaches to positively impacting disadvantage. As we summarised in the introduction, cybernetics has identified feedback mechanisms as critical to communication between and within system components, and they help to reveal circular causal processes that steer the system toward its purpose/s. As we analysed the extent to which data professionals already approach their work in a reflexive manner in line with

[73] Genevieve Bell et al., "Do More Data Equal More Truth? Toward a Cybernetic Approach to Data," *Australian Journal of Social Issues* 56, no. 2 (2021): 7.

cybernetics, our analysis drew on these fundamental understandings of cybernetic systems, and on specific elements of cybernetic theories developed in the works of Donald Mackay [74], Heinz von Foerster [75], Jay Forrester[76], W. Ross Ashby[77] and Stafford Beer.[78] We drew on these understandings to analyse the extent to which data professionals shape a system's purpose, how they make decisions under limited ethical guidance and how they conceive the import and flow of information within and across interacting systems. In the following section paragraphs, we outline some detail around how we applied these understandings.

How observers in the system shape a system's purpose

As we summarised in the introduction, Heinz von Foerster developed some cybernetic conceptions of how an observer of a system plays a critical role in enacting system regulation. He says that observers interpret and understand data systems through the meanings that they attribute to the system and/or the data coming from those systems.[79]

In our analysis, we have used von Foerster's focus on the observer of the system to narrow and direct our analytical focus to the insights provided by the data workers and decision makers. We explored the ways that these observers see themselves as participants in the systems and how they understand the goal of using data for breaking cycles of disadvantage. Moreover, we evaluated the recursive relationship between the ways that definitions of disadvantage, (including who we label as disadvantaged), are understood, and the ways that these definitions shape how data is collected and analysed.

Decision-making under conditions of uncertainty

We have also focused our research around data professionals involved with complex dynamic data projects working with often limited information. These conditions offer opportunities for data workers to engage in overt reflexivity as they are nudged towards making decisions based on their own judgement rather than on policies or regulations. Donald Mackay's work on decision-making processes under conditions of limited information and/or increased uncertainty has helped frame our thinking. He utilised information theory from Claude Shannon[80] among other cyberneticians in order to theorise how information is packaged, processed, transmitted and optimised by people and ultimately by computers and artificial intelligence.[81] His work when combined with Ashby's law of requisite variety[82], has supported the identification of instances of information feedback and helped to conceptualise possible locations where more feedback mechanisms can be placed for improved system viability.

Increasing information flows across a system

Finally, our cybernetic approach to investigating the interactions and behaviour of data assets is guided in part by Stafford Beer's Viable Systems Model. In the VSM, Beer made allowances for data limitations and potential data biases in his model.[83] The VSM is designed to enable optimum management of organisations by relying on the efficient flow of data and/or

---

[74] *Information, Mechanism and Meaning* (Cambridge, Massachusetts: MIT Press, 1969).

[75] "On Self-Organizing Systems and Their Environments."

[76] Jay W. Forrester, *Principles of Systems* (Productivity Press, 1990).

[77] Ashby, *An introduction to cybernetics*.

[78] "Brain of the Firm; a Development in Management Cybernetics." (New York: Herder and Herder, 1972).

[79] von Foerster, "On Self-Organizing Systems and Their Environments."

[80] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal* 27, no. 4 (1948): 623–56, https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.

[81] MacKay, *Information, Mechanism and Meaning*.

[82] Ashby, *An introduction to cybernetics*, 211.

[83] Beer, "Brain of the Firm; a Development in Management Cybernetics."

information within a given organisation. In optimising information flows, one solution Stafford Beer had for limiting the negative consequences of incorrect or limited data was to build in more than the required number of data sources. For example, in collecting data from a factory floor and transmitting to a leadership team, there was scope to always have more than one channel from which to receive that information. The redundancies while making the communication systems costlier, helped to create greater visibility of the system(s) of interest and more resilient against data misuse or misappropriation.

The VSM can be envisaged as an early instantiation of a data-driven system rather than a data-informed system. Through a data driven mentality, attitudes of continuous learning or of remaining adaptive become part of the culture of such systems. In the works of both Ashby and Forrester, we are guided in identifying some of the mechanics within and between systems such as instances of circular causality.[84]

While our rich data case study projects were not purely data driven, both aspire towards being data driven in the future.

Our rich-data case studies – MADIP and GenV

Supported by the Paul Ramsay Foundation, our research sought to analyse the extent of reflexivity and systems approaches as demonstrated by data professionals from two world-class data assets: the Multi-Agency Data Integration Project (MADIP AKA PLIDA)[85], an effort on the part of federal government to better leverage the administrative data sets collected in the course of providing public services; and Generation Victoria (GenV), a longitudinal, life-course health study of Victorians born in a particular pre-defined 'generation' (birth window), following a cohort from birth to death.

We chose these two data assets because:

- They offer a rich system within which to explore different kinds of data use and data decision making (e.g. complex longitudinal data collections are increasingly a basis for a range of data science projects, increasingly using AI and machine learning, that seek to address aspects of disadvantage).
- They collect data on dynamic, changing subjects, in dynamic, changing contexts, for which cybernetic concepts and approaches are particularly well suited.
- They are inherently high-stakes, because they are both incredibly valuable and useful in a range of settings while also needing safeguards against the potential for misuse or inadvertent harm over time.
- Their approaches to legal and privacy considerations differ and offer us a window into varying approaches.
- Purposefully-designed data assets often have higher levels of data sophistication and maturity. The sophistication and maturity stand to offer us a good indication of current best practice and awareness of systems dynamics.

In analysing the complex relationships surrounding MADIP and GenV, we identified a set of five overlapping and interfacing systems aligned to disrupting disadvantage. These systems are the:

1. Research system (academia, public institutes)
2. Data system (public services and institutions, infrastructure providers)

[84] See Ashby, *An introduction to cybernetics*; Forrester, *Principles of Systems*.

[85] In May 2023, the Australian Statistician, Dr David Gruen announced that in recognition of what the MADIP data asset represents, the asset's name would transition to PLIDA (Person-Level Integrated Data Asset). There is a phased transition towards PLIDA branding and MADIP will continue to be used in tandem.

3. Policy system (government, regulators)
4. Charity system (NGOs, philanthropist foundations, social enterprises)
5. Legislative and jurisdictional systems (as it applies to data sharing and use)

In drawing this broader analytical boundary, we seek to make explicit the dynamics between reflexive data decision making, data assets, and broader technical, infrastructural, cultural, social, economic, and political systems in which they operate.

Reflexive data decisions in these dynamic environments often span the boundary between the data assets and the broader five systems of research, data, policy, charity, and legislative and jurisdictional. This group is highly attuned to the ways in which data analysis, using MADIP and GenV, is or intends to be used in support of designing policy and service delivery. At this stage, we have not explicitly sought representatives from disadvantaged communities impacted by data decision-making; however, we perceive this as crucial to future stages of development of this work in the future.

Together, these two assets offer an examination of how they are and could be used to address disadvantage and in so doing, illustrate some of the potentials and pitfalls data professionals must manage while using complex datasets to solve for disadvantage. They provide timely insight into what current and future data work for interventions into social challenges like disadvantage can and might look like. These two assets exemplify our primary focus on how data assets are contextualised in a broader set of systems and how and why data professionals need to engage in reflexive cybernetic practises in environments where ethical judgements are critical.

While they share the commonalities above, MADIP and GenV are substantively different data assets in terms of their organisation, operation, data sets, and remit. Moreover, their highly complex and differing social and political remit offers the opportunity to explore highly complex multi-system interactions and how data professionals who are themselves cybernetic systems are positioning themselves within these dynamic set of systems. Here we summarise key aspects of each data asset.

Multi-Agent Data Integration Project (MADIP)

Purpose

MADIP, part of the Data Integration Partnership for Australia (DIPA) scheme, was initiated in 2017 by the Department of Prime Minister and Cabinet with a $130.8 million investment. It aims to enhance the utility of government-held data by consolidating information on health, education, government payments, income, taxation, employment, and demographics to create a comprehensive longitudinal view of Australia.

Legislative Framework

MADIP operates within the Census and Statistics Act (1905) and the Privacy Act (1988), which safeguard individual identities, prohibiting use for regulatory or compliance purposes. Microdata access is further anonymized under the Census and Statistics (Information Release and Access) Determination 2018. The "5 SAFES" framework ensures data privacy, and MADIP risks and operations are frequently assessed, including via the MADIP Privacy Impact Assessment. The Data Availability and Transparency Act (DATA) of 2022 is tipped to enable improved data sharing between federal government data custodians and jurisdictional data custodians because the widening of the range of ‘accredited data service providers’ is widely seen as paving the way for greater and fairer involvement by state and territory government bodies, and Australian universities in data sharing initiatives.

Organisational Structure

Enduring data custodian agencies, including Australian Bureau of Statistics (ABS), Australian Taxation Office (ATO), Department of Health and Aged Care (DHAC), Department of Social Services (DSS), Department of Education (DE), Services Australia (SA) and Department of Home Affairs (DHA), supply data for linkage. Jurisdictions can also contribute data either through these agencies or directly upon request.

Operations

Access requests are submitted to ABS by researchers or government agencies. Researchers must meet ABS safe people criteria, including proficiency in statistical languages, research experience, location in Australia, organizational affiliation, and project approval. Although most researchers are located in Australia, a small number of international organisations (considered on a case by case basis), are approved to access data from overseas. The ABS provides the myData Portal for request management and access tracking. The primary access method is through ABS's DataLab, hosted on Microsoft Azure, meeting "Protected" level security standards. Data encryption, closed network virtual machines, and multi-factor authentication enhance security.

Data Sets

MADIP encompasses standard detailed microdata, which includes ABS surveys, external data with custodian approval, and integrated data like MADIP modular products, BLADE, IPLORD, and merchandise imports/exports. This data is designed for use within DataLab, with direct identifiers removed, and additional confidentiality measures applied. It covers diverse topics such as health, education, labour, demographics, and more. Limited release detailed microdata is available to government employees, contractors, academics, and researchers for approved projects.

Future Projects

Several new data integration projects and data asset partnerships are underway to improve MADIP:

1. **Justice Asset**: This national data asset links criminal offenders and prisoners, enabling analysis of interactions within the justice system and potentially linking with other datasets for deeper analysis. For example, a pilot project aims to link crime and justice data with social datasets, providing insights into topics like domestic violence and perpetrator outcomes.
2. **National Disability Data Asset (NDDA)**: Under development, NDDA integrates de-identified data from various service systems to provide insights into people with disabilities' pathways through services.
3. **Vocational Education and Training (VET) National Data project**: Integrating VET data with MADIP and BLADE, this project will offer insights into the outcomes of VET students, including employment, further study, and government assistance needs.

MADIP represents a significant effort to harness government data for research, policy development, and public benefit while upholding strict privacy and security measures and this makes it one of the most dynamic and complex data systems in the Australian linked-data

ecosystem[86]. In May 2023, the Australian Statistician, Dr David Gruen announced that in recognition of what the MADIP data asset represents, the asset's name would transition to PLIDA (Person-Level Integrated Data Asset).[87]

Generation Victoria (GenV)

Purpose

GenV is an ambitious research project led by the Murdoch Children's Research Institute (MCRI). It stands as the world's most extensive early and midlife cohort study, spanning a decade and uniquely dubbed the "COVID cohort." GenV's primary objective is to design, deploy, and execute a comprehensive longitudinal life course study across Victoria, Australia. This study simultaneously involves two key cohorts: babies and parents, each organized to participate in health discovery and intervention trials throughout their lives, with the intention to extend beyond a century. The pivotal concept behind GenV is the fusion of observational research (discovery) and proactive intervention trials, aiming to expedite progress in solving complex health and societal issues.

Focus Areas

GenV has identified six focus areas that significantly impact children, parents, families, and the community. These areas have been selected due to their burden, cost, slow progress, and the overarching themes of inequity and vulnerability. The focus areas encompass:

1. Development and Learning
2. Infection, Immunity, and Allergy
3. Organ Health
4. Healthy Environments
5. Obesity and Diabetes
6. Mental Health and Wellbeing

Organisational Structure

GenV is headquartered at the Murdoch Children's Research Institute (MCRI) and is financially supported by grants from various sources, including the Paul Ramsay Foundation (PRF), the Victorian Government (Department of Health & Human Services and Department of Jobs, Precincts and Regions), the Royal Children's Hospital Foundation, and collaborations with the University of Melbourne's Department of Paediatrics. The PRF is a philanthropic foundation that is partly focused on health, education, and early childhood. The Victorian Government (Department of Health & Human Services) is responsible for policies, programs, and services to enhance the health and wellbeing of Victorians.

Operations

[86] Visit the DataLab website for greater detail - https://www.abs.gov.au/statistics/microdata-tablebuilder/datalab#applying-for-and-using-datalab

[87] "Supporting Analysis of The Life Course | Australian Bureau of Statistics," August 16, 2023, https://www.abs.gov.au/about/our-organisation/australian-statistician/speeches/supporting-analysis-life-course.

The recruitment of participants into GenV is managed by MCRI teams and is facilitated by hospitals and community health services. Various partnerships, processes, and collaborators ensure that all eligible Victorians have the opportunity to participate with informed consent. The recruitment approach involves MCRI-employed recruiters personally approaching parents of newborns for consent, with interpretation and translation support if needed. Children are given the opportunity to decide on continued participation as they reach the age of legal consent.

Technological Infrastructure

GenV's "Solution Hub" will serve as the engine driving the project's impact. It seeks to play a pivotal role in advancing GenV's scientific research, building human capacity, translating knowledge, and securing broad-based funding for analysis and research. The Solutions Hub is divided into two main arms: one focusing on epidemiology and methodological design, and the other concentrating on partnerships, knowledge translation, and collaborative research. Both arms collaborate closely to fulfil the Solution Hub's multifaceted objectives.

Data Sets

GenV adopts a three-tiered approach to data collection, blending study-collected, study-enhanced, and linked data. All GenV data and biosamples are utilized exclusively for research aimed at enhancing human health. The cohort consists of a representative sample drawn from the Victorian population, totalling between 150,000 to 170,000 participants. Cohort members remain part of the study even if they move out of Victoria, with an opt-out mechanism in place. GenV's data collection strategy includes consent soon after birth, retrospective and prospective data linkage, banking of biosamples, and GenV-collected biosamples and data.

Future Projects

GenV has a series of future projects on the horizon:

1. **Trials and Interventional Capabilities**: GenV is planning an "Intervention Hub" for clinical trials, uniquely targeting children and younger adults who are often underrepresented in research, especially in large-scale trials.
2. **Policy Agenda and Mapping**: GenV aims to investigate the impact of various policies (federal, state, local) on individuals and communities, leveraging its diverse and large cohort, and the unique context of its establishment during the COVID-19 pandemic.
3. **GenV's School Entry Wave**: By 2028, GenV's cohort will be of school age, providing an opportunity for technology-driven assessments of various aspects of child development, utilizing tools like wearable devices and apps.
4. **Scaling Direct Participant Measurement**: GenV is exploring the use of its customized 'GenV and Me' app to remotely deliver assessments and measurements, aiming to collect rich phenotype data at a universal scale over time and across geographic areas.

As it is evident, the GenV project is ambitious and it represents a monumental effort to enhance our understanding of health and wellbeing across the lifespan, leveraging cutting-edge technology and a diverse cohort to address complex health challenges and improve the lives of Victorians and beyond[88]. The successes achieved through this asset thus far exemplifies decision making in which high stakes reflexivity continues to be pivotal across all its processes.

[88] Generation Victoria, “GenV,” Generation Victoria, accessed October 4, 2023, https://www.genv.org.au/.

Interview questions

Our research sought to discover the extent to which the decisions data professionals make when utilising complex datasets for disadvantage interventions in dynamic environments, can be described as exhibiting or utilising cybernetic reflexivity. Motivated by this research agenda, we developed questions that sought to illuminate whether data workers and decision makers were explicitly aware of acting within a series of complex interconnected and dynamic systems in which they can intervene, whether they were cognisant of circular causal processes within the systems in which they are situated and whether they were aware of the pivotal role information feedback loops play within those complex data projects in support of safe, responsible and sustainable data use, whether they saw their roles within the data systems as observers and users of data or whether they saw themselves as participants in the data creation process, whether data professionals reflected on the nature of disadvantage and whether they often reflected on the repercussions of definitions to the human rights of those they imagine to be vulnerable or at a disadvantage. We developed and asked the following interview questions to professionals who are using data from or are situated within, our two case studies: The Multi-Agency Data Integration Project (MADIP) and Generation Victoria (GenV).

1. What do data professionals understand by the phrase rich-data and how to they envisage its utility for disadvantage intervention?
2. How are data decision-makers using this data for disadvantage research and/or social innovation (to improve any number of wellbeing, health and education outcomes and/or equivalently support human flourishing ideals)?
3. What challenges do data assets and data decision-makers face when seeking to address disadvantage?
4. What involvement in data assets and data projects do data decision makers and data workers envisage for those experiencing disadvantage?
5. How can data assets serve as pivotal components of multi-perspective approach to interventions addressing disadvantage?
6. What are the broader (pre-)conditions that would enable the continuous best use of data for disadvantage initiatives?
7. What kinds of change planning and future-proofing would enable data assets to be of ongoing value in disadvantage intervention?

METHODOLOGY

Qualitative interviews

We initiated this research by seeking ethics approval to interview, analyse and use data obtained from research interviews. Our ethics protocol was approved and the permission to use quotes from interviewees within this report have also been affirmed.[89]

We then undertook qualitative interviews with data professionals working with two data assets - GenV and MADIP, across all sectors that touch these data assets. Nine expert witnesses were consulted formally, with supplemental informal conversations held both before and during the research process with a larger cohort of participants (n~30).

We then used open/grounded coding to discern common themes discussed in the interviews and conversations. Findings from this qualitative enquiry are synthesised in our research findings section, and supplemented with key debates in the literature. As the data assets we examined are fairly new, the academic literature we draw on often reflects only generalised and/or recent historical situations of the data, policy, and disadvantage landscape in Australia.

Meta-analysis

After research findings were synthesised, we needed to discern the extent to which our research participants exhibited cybernetic reflexivity. We achieved this meta-analysis by developing a novel methodology that complimented our abductive reasoning about the cybernetic reflexivity of our research participants. Our novel methodology is called PAFCARSS and it was founded upon few key insights about the way in which systems function as units of analysis. We will present the PAFCARSS method here then describe the key insights that led to the development of this method.

The PAFCARSS method as a systems visibility inspired reflexivity indicator

PAFCARSS is an acronym that stands for Purpose, Agency, Feedback, Causality, Autonomy, Reflexivity, System and Steering. In our methodology, each of these ideas act as checkpoints the data professionals go through in discovering and practising cybernetic reflexivity. Our key insight in developing this method was that systems visibility through a cybernetics lens necessitates the agent to be reflexive. This insight will be explained in greater detail after presenting the method itself. We used the PAFCARSS checkpoints to analyse our research findings abductively and doing so enabled us to tell a story about the extent to which the interviewed data professionals practised reflexivity.

The steps we went through in applying our new methodology are outlined below. In each step, our use of 'agent/s' refers to data workers and decision makers but the methodology can be applied to other agents in the future.

1. Agents acknowledge their perspectives in identifying the purposes, objectives and goals that are evident within the systems they have identified.
2. Agents utilise purposes in 1 to surface other autonomous systems by selecting appropriate system boundaries (boundaries that will allow for the identification of circular causality and information feedback points).
3. Agents identify information feedback points by considering 1 and 2.
4. Agents identify causal loops by considering how 3 flows through 2 and impacts 1.
5. Through reflexive awareness of the systems' impact on them and vice versa, agents utilise 3 and 4 to consider how they themselves are one of the systems in 2.

[89] ANU Human Ethics Protocol 2022/485

6. Having become aware of themselves as systems that are acting and being acted on by other systems and within causal loop structures, agents work out their next steps of action of steering the systems discovered in 2 and 5, towards a preferred purpose or goal. In our analysis, the preferred goal is the development of data-driven or data informed products that support vulnerable people and those experiencing disadvantages to be safe from data harms while being supported towards human flourishing ends.

In summary, we tested our data and disadvantage literature and interview analysis using PAFCARSS and the result of this analysis have been detailed in our discussion section.

Key insights to developing the PAFCARSS methodology

From its earliest instantiation, cybernetics has been concerned with the role human *agency* plays in achieving any given human-machine system's *purpose*, objective or goal. At this early stage of its development, taking the *system* as a unit of analysis and the role of human agency were key components of cybernetics. These early stages of cybernetics give us the PAS in PAFCARSS.

Secondly, it was by reflecting on the role of human agency (voluntary action) that Wiener and other cyberneticians co-opted the concept of feedback from engineering in order to illuminate the pivotal role information *feedback played in* supporting the agent to guide their repetitive (circular) actions (*causes*) towards the system's desired purpose/s, objective/s or goal/s. This interaction between agency, information feedback and circular causality were to play a pivotal role in the field of cybernetics as research later revealed various types of these system structures across, biological, technical and environmental systems. Adding these concepts gave us the PAFCS in PAFCARSS.

Thirdly, these systems exhibiting information feedback and circular causality were causally closed systems. They were causally closed systems in the sense that regardless of the causes upon them, these systems maintained their unity - they were autonomous. These systems concept of *autonomy* under causal closure also enables us to think of systems as having objective/s, purpose/s or goal/s. Adding our concept of autonomy gave us the PAFCAS in PAFCARSS

The fourth insight centres on humans as systems. When a critical mass of cyberneticians finally accepted the idea that all human beings like all biological entities are autonomous and hence systems, it was inevitable that they had to be conceived as systems that are acting and being acted upon within a larger set of interacting systems. The realisation that each observer is a system, necessitated the observer to consider and factor in the effects of their agency on other systems and the consequence of other systemic actions on them as systems – they were obligated to partake in reflexive practise. In short, *reflexivity* became an important concept and practice as these cyberneticians relied on cultivating it for gaining an understanding of large complex social systems and the impact of their causal processes on segments of the community such as the vulnerable and those at a disadvantage.

When combined with the ability to notice circular causality and information feedback points within the systems in which they interact, the fifth insight was that reflexivity gave those cyberneticians a way to contemplate where and in what manner they can exercise their agency in a system, in order to *steer* it towards identified purposes. Adding reflexivity and steering completed our acronym PAFCARSS.

In extrapolating these cybernetics insights and applying them to the data context, we concluded that, as data are products of complex social systems in which data professionals who are autonomous systems are situated, each data professional reacts to each dataset in ways that are dependent on their own characteristics as autonomous systems with agency. Through this cybernetic understanding, it becomes obvious that regardless of the extent to which ethical guidelines are followed, each data professional's agency impacts significantly on how datasets

are handled, made sense of and communicated and in turn how they impact the vulnerable and those experiencing disadvantage.

The methodology treats system components as indicators of reflexivity and it has allowed us to consider the degree to which those data professionals we interviewed could be said to be engaging in cybernetic reflexivity. Through meta-analysis of our interview findings using PAFCARRS, we sought to ascertain the degree to which data professionals are cognisant of the impact of their agency on datasets in general and in particular, on datasets from or about vulnerable citizens or those considered to be at a disadvantage.

Cybernetic praxis as our broader research approach

Heinz von Foerster was also one of the early proponents of an approach to cybernetics that was a praxis - simultaneously theory and practise. "If we wish to maintain our scientific credibility", stated von Foerster, then we need to show that the science of regulation also applies to the global society of cyberneticians. Von Foerster saw that cybernetics endowed practitioners with strong skills (which he called competences[90]) across disciplines, and that these skills mandated cyberneticians with the responsibility to create systemic change where they see the need. Von Foerster's emphasis on cybernetics praxis has significantly informed our broader research approach that goes beyond literature reviews, interviews and analysis and in to approaches that were always open to new ideas from different voices and perspectives or via recursive processes of learning by observing, reflecting, proposing, testing, iterating and evaluating.

We have intentionally engaged with diverse experts within the School of Cybernetics and externally to consult on data, poverty, social deprivation and social exclusion which all come together to define disadvantage.

These principles have informed the design of this project through the following forums:

Firstly, we created an advisory group of staff and students at the School of Cybernetics to consult on this project. The advisory group comprised of cybernetics experts from within the ANU School of Cybernetics and other data and social intervention professionals from Commonwealth Government and the Paul Ramsay Foundation. The role of the Advisory group was to provide support and guidance to the project team throughout the development until the completion of the first phase of this project.

Secondly, we engaged with a range of experts at various stages in our research over this first phase of our project, as part of building a community around this work and improving our own research approach.

We spoke to a number of experts as part of a cybernetic approach to a literature review that is cognisant of the vital contextual information experts have that is not published. Our acknowledgements section details some of these experts.

We also had various experts with whom to check technical claims arising from research interviews. These have also been detailed in the acknowledgements section.

After our research data analysis, we engaged experts across MADIP and MADIP users group across government, academia and the private/non-for-profit sector to explore key themes that will guide our systems mapping design workshops. These expert group is also listed in the acknowledgements.

---

[90] Heinz von Foerster, "Cybernetics of Cybernetics," in *Understanding Understanding: Essays on Cybernetics and Cognition*, ed. Heinz von Foerster (New York, NY: Springer, 2003), 197, https://doi.org/10.1007/0-387-21722-3_13.

Our research interview participants were also an expert group comprised of various other data decision makers and data workers sourced with input from advisory group, PRF and other research networks. We thank them for their generosity with their time and openness.

DISCUSSION

Our data and disadvantage project had been scoped broadly around the possibility that cybernetics can offer a valuable perspective on how to think about the effectiveness of data-driven and data informed disadvantage interventions. We were intrigued by the question of whether data professionals' visibility of and an appreciation of the complex systems interacting with the data and disadvantage ecosystem, could result in changes to data decision making practices, and perhaps unearth new interventions that not only minimise or mitigate future possible harms but could also lead towards human flourishing. The project has been scoped as a multi-year project with this initial year dedicated to understanding the current best practise in systems' approaches to data decision making within complex and dynamic rich-data environments

Our research hypothesised that data workers and data decision makers (data professionals) within rich data contexts like GenV and MADIP can support us to formulate a sense of current best practice decision making in collecting, caring for and using data for social good, within highly dynamic environments with unclear or competing policies, laws and legislations. In our introduction, we had given reasons for why a focus on examining reflexive and purposeful data decision making within these environments, can be fruitful and preferred when it is founded upon;

1. our cybernetic conceptions of data professionals as autonomous systems interacting with a variety of other technical, environmental and social systems.
2. our use of the PAFCARSS method to evaluate the extent to which these data professionals were aware of various systems' complexities.

For reference, a summary of our PAFCARSS methodology is as follows:

1. Agents identify purposes within systems of interest.
2. Agents identify systems by selecting appropriate system boundaries.
3. Agents identify information feedback points by considering 1 and 2.
4. Agents identify causal loops by considering how 3 flows through 2 and impacts 1.
5. Agents utilise 3 and 4 to consider how they themselves are one of the systems in 2.
6. Agents work out their next steps of action of steering the systems discovered in 2 and 5, towards a preferred purpose or goal.

PAFCARSS has enabled us to analyse and articulate the extent to which data decision makers and data workers exhibited reflexive attitudes as they went about using data for social intervention. In this discussion, we explore how cybernetic reflexivity was demonstrated across key systems complexities that combine to decentre data in rich-data environments. We visualised system complexities by focusing on data professional's attitudes and approaches to disadvantage, on legitimacy and trust and on concerns relating to data system resourcing and data access rights. Our in-depth research and interview findings and detailed practical lessons learned from the findings can be found in appendix 1 but here we analyse few of the standout findings and practical lessons.

Data professional's attitudes and data approaches to disadvantage

In the literature, through our expert panel focus groups and in our research interviews, we sought evidence for reflexivity and decentring of data during project phases in which 'vulnerability' or 'disadvantage' were defined.

We found approaches to defining disadvantage were largely pragmatic and privileged quantifiable options and this has resulted in various ways of modelling disadvantage using a variety of variables. These varieties have resulted in two types of identified inconsistencies. Firstly, there were inconsistencies in the general use of the concepts of disadvantage. Secondly, there were notable absences of consistent data definitions of disadvantage within specific

domains like education, health, and economics. While in most cases this was done with significant sensitivity and respect for data subjects, it also revealed significant scope for developing reflexive approaches to defining vulnerability and disadvantage within data contexts. Moreover, it was evident that reflexivity was either practised or welcomed by data professionals within these definitions stage of the data pipeline but there was less awareness of its appropriateness across the rest of the data-pipeline. In particular, there was little evidence to suggest reflexivity featured in post-production where questions were likely to be asked concerning the immediate, medium term and in the long term impact on vulnerable data subjects.

We found that data professionals largely acknowledged their poor knowledge of the disadvantage literature and have often expressed this as the reason for working with disadvantage domain experts. From a reflexivity perspective, this is valuable but not a substitute for data workers' own comprehension and appreciation of disadvantage, which can be enhanced by involving lived-experience experts when deciding upon how to model disadvantage.

Our focus groups have made it clear that defining vulnerability and disadvantage is complex and messy when data is involved. When data is centred and reflexive practise is mooted, some key datasets representing experiences of those at a disadvantage, can be curated out of datasets as outliers or can be overlooked in other statistical techniques. In contrast, cases in which data professionals exhibited reflexivity, manifested in their willingness to dwell in the complexity of the challenges associated with definitions and adequate representation and that eventually they sometimes discovered ways to include often overlooked datasets that did not align with expected dimensions. One example of how key data sets were included was where key lived-experience experts were incentivised to participate in data collection and validation resulting in greater data robustness.

Here, we see a significant opportunity to offer a sector-wide cybernetic tool that can support reflexive practice on the impacts of data uses on data subjects. For example, the collaborative development of information feedback mechanisms that include both data decision makers and workers, and lived-experience experts would not only support the identification of relevant datasets but would also enrich those datasets making them truly representative. Ultimately, these feedback mechanisms would support the creation of data products that are potentially better aligned to lived-experience experts' aspiration towards human flourishing.

Most of our research participants also identified the need to shift towards asset-framing or strength-based approaches to disadvantage. This identification was strongest where data was decentred and broader systemic purposes or objectives were the focus. A consistent goal that was identified for the disadvantage data ecosystem was to meaningfully measure and evaluate how well it was addressing disadvantage. This identification of system goals is symptomatic of data professionals embracing steps towards reflexive practice. One of the reasons for this is, it opens the opportunity to reflexively align the ecosystem's broad objectives with asset framed data and disadvantage approaches. In this way, reflexive approaches have the potential to offer different intervention opportunities, which are intended to be positively reinforcing.

This reflexive focus that is anchored in system purposes suggests the formation of working groups that are inclusive of lived-experience experts and that aims to agree upon functional definitions of disadvantage for each domain, considering a diversity of viewpoints and quantifiable, strength-based data variables to bring clarity for intervention and policy impact. As an example, strength-based definitions of disadvantage in education could include measures such as student educational aspiration, self- and community-identification with scholarship and achievement, teacher dedication, school models, and societal attitudes towards schooling.

Moreover, there was evidence to suggest that when senior data decision makers engage in a reflexive focus on the data systems' purposes, they are more likely to initiate actions that will expose other data workers more removed from the 'coal-face' of policy impact, to opportunities that are supportive of their alignment with these high level system purposes.

Another high level system goal we identified among disadvantage data professionals was the importance of having a clear understanding of the strategic impacts and outcomes of data projects utilising linked-data from rich data ecosystems. A suggestion from the data workers is a need for linked-data custodians to readily identify and highlight data that is relevant to disadvantage. An exemplar of this multisystem synergy that is possible when data itself is decentred is the latest ABS Life Course Data Initiative that seeks to compile data sets that are readily available to support data-driven efforts to help those experiencing severe disadvantage[91]. On this point, it was agreed by interviewees that support systems such as charities themselves, which are largely operating under a crisis 'safety net' paradigm, are likely to embrace a redesign towards human flourishing ends when there is a pivot towards a systems view that seeks to align with the overall goal of the data and disadvantage ecosystem.

We see a significant opportunity here to deliver training workshops that can support organisations and data teams to identify a variety of system goals that are conducive for multisystem synergy.

We also unearthed evidence relating to consequences of non-reflexive data centred approaches. Data professionals not engaged reflexively tended to see the dignified individuals experiencing complex disadvantages objectively as monolithic data points. Others have identified that such objective approaches limit the efficacy of data-driven initiatives aiming to build support services for vulnerable citizens. This is partly because they stifle possibilities of seeing the usefulness of meaningfully incorporating lived-experience experts as active co-creators of the service. Compounding these consequences were the inflexible nature of the allocated project time and funding. The importance of system purpose identification together with an embracing and preservation of multiple perspectives and agencies recommend that data should not be used to simply identify and track the disadvantaged where they struggle to navigate social services and/or health systems, but also use aggregates of these types of data to understand how the systems themselves may be preventing or hindering the vulnerable or those at a disadvantage from assessing support.

As part of the larger reorientation of support systems towards enabling flourishing, funding should be made available to support charities and organisations at the frontline of working with those experiencing disadvantage to upskill towards best uses of data for those they serve. It is widely acknowledged and documented that the for-purpose sector is generally underserved when it comes to data skills and literacy. By offering frontline workers, who often notice interacting systems and competing objectives, an appreciation of and competency in data and data analysis will enable them to identify opportunities for further involving and co-designing data-led interventions with their communities. This would be an effort to supplement their existing systems approaches with data practises rather than seeking to replace or reduce them in preference for data-driven or data-informed approaches.

Through our PAFCARRS methodology, we identified that data workers were aware of the importance of expanding their interactions with systems that interface directly with those experiencing disadvantage but where often at a loss of how this could be operationalised. Our practical advice here is for the data and disadvantage community to invest in building a community of people who self-identify as experiencing disadvantage. The literature shows that, such a community would help the broader research and data community create more effective and potentially less harmful data products that are cognisant of the differing priorities and needs of those experiencing disadvantage[92]. These synergies can empower lived-experience

[91] "Supporting Analysis of The Life Course | Australian Bureau of Statistics."

[92] See for examples the communities established as part of the 100 families project Shawn Phillips et al., "Insights into Hardship and Disadvantage in Perth, Western Australia: The 100 Families WA Report" (100 Families WA, 2021), https://doi.org/10.25916/B914-1J34.

communities to ensure that data projects have their best interests at heart and they can beta-test intervention solutions intended to be helpful.

Through our cybernetic approach, we envisage that a community of lived experience experts would play a pivotal conduit role in the data and disadvantage ecosystem's effort to establish information feedback mechanisms or iterative approaches that aim to communicate with lived-experience experts beyond an identified community of lived experience experts. We see robust opportunities and have begun developing system mapping resources to support rich data projects to identify the salient points at which consultation with lived experience experts can be a useful step in intervention co-design.

Overall, our meta-analysis of data professionals' attitudes to disadvantage has revealed that attitudes and data approaches to disadvantage can benefit from a cybernetic assessment of how the experience of disadvantage is typically defined and treated within a data context. This assessment can then be supplemented with hands-on workshops and training in identifying ways to build impact and agency more strongly into data and disadvantage initiatives that we see as human flourishing initiatives.

Data professional's attitudes and approaches regarding legitimacy and trust

In investigating and analysing data professionals' attitudes and approaches surrounding legitimacy and trust, we assessed the broader cultural, social, and political systems that enable the creation, scaling and maintenance of rich data assets as enduring social goods.

Our research revealed that rich data assets utilise either one or combination of passive and active data collection methods and that each approach necessitates differing forms of consent and sensitivities around social license. Interviews revealed that consent tools that embrace a systems approach that is contingent on the reflexivity of designers and accounts for long term evolution is more likely to ensure its long term viability. Part of the reason for why long term viability is an important consideration in consent design is that often very complex data assets like GenV face dynamic ethical challenges related to the disclosure of sensitive medical results belonging to participants. As cultural norms and social and political approaches shift overtime, a reflexive approach necessitates ensuring that consent forms are open to continual revision. In considering the information feedback mechanisms and causal paths associated with dynamic forms of consent, our interviews revealed that data professionals had low awareness of such mechanisms and causal pathways. Upon reflection on possible feedback and cyclic causal paths impacting on consent, we discovered that any form of dynamic consent signed by research participants would benefit from active consent management – a set and forget approach would not be ideal. Active consent management could result in the establishment of ongoing participant support services; such as consent counselling as young cohorts mature to consent age. These active management approaches would contribute to fostering higher participant retention rates and emphasizes the benefits of participation.

When a cybernetic systems mapping of causal paths and information feedback mechanisms are considered, it promotes data asset management approaches that are adaptive to shifting norms around data ownership in both private and public spheres. Different data ownership models (private vs. civic) have implications for data assets, including challenges like handling deletion requests and compensating participants for commercial data uses. We found low evidence of deliberate consideration of approaches that are actively accounting for shifting norms around data ownership. This finding suggests opportunities for the deployment of cybernetic tools that can supplement data professionals' data pipelines with opportunities for mapping these future facing systems dynamics. This is especially the case for data assets like GenV which are likely to produce new knowledge about certain participants wellbeing over time through participation in medical trials or through routine health tests. In these cases, those data assets that have in place timely information feedback mechanisms can triage processes surrounding the disclosure of

medical results while planning bespoke responses to the needs of research cohorts. Over time and due to their cyclic nature, these feedback mechanisms offer research cohorts the opportunity to revisit disclosure and consent discussions and decisions as those cohorts' journey through the life course.

Trust and relationship building concerns were evident not only between data custodians and research cohorts but it also existed within groups of data custodians as well as between custodians and data assets. Through our literature search and in interviews across both our case studies, we found little to no formal mechanisms to compel the processes around data linkage and sharing. In most cases, data assets operated largely on the goodwill of data custodians. At the same time, these structures of goodwill were sensitive to perceived conflicts of interest, power asymmetries, and legacy relationships. These sensitivities resulted in either improving or diminishing the closeness and timeliness of collaboration. Secondly, our analysis uncovered implicit or informal relationships between institutions with data custody responsibilities. Depending on the nature of these relationships, data linkage processes were either stifled or expedited. In both these cases of goodwill structures and implicit connections, our meta-analysis suggests that opportunities exist for all parties to engage in a collaborative data systems mapping exercise to increase transparency between themselves by articulating the autonomy and agency of each of the custodians and by surfacing shared goals and mutual benefits.

Lastly, risk and appetite for risk tolerance emerged as a substantive influence on legitimacy and trust surrounding data decision making across the junior to senior leadership spectrum. We found that the need to satisfy and strengthen safeguarding measures were front of mind. This manifested in various demonstrations of implicit and explicit risk management practises. Data systems demonstrated a variety of orientations to risk from extreme risk aversion to overt risk tolerance that was to an extent dictated by legislative frameworks and operational missions. This mismatch between foundations and missions are often not explicitly acknowledged within data sharing discussions. Our reflexive approach suggests opportunities for identifying causal pathways connected to risk communication mechanisms. In turn, this knowledge will support reflexive considerations about how each data decision maker considers their risk orientation in reference to organisational risk orientation. Illuminating causal pathways and risk communication channels open opportunities for risk averse actors to consider the impact on impactful data utility and other counterfactuals against data control. These reflexive sense making both internally and externally to existing data systems has the potential to improve legitimacy and trust.

In summary, our research has revealed that when our reflexive approach is founded on systems visibility, data risk manifests as the responsibility not only to safeguard but also as a responsibility to make data useful for social good. Within a wider social dialogue in which all public data sets are implicated, especially where the asset is comprised of public data that could be used for disadvantage work, those assets need to be seen as contributing to the betterment of society while being handled appropriately and participants treated ethically. A cybernetics systems approach can showcase various approaches for data professionals to maintain legitimacy and trust in their data institutions.

Data professionals' concerns about data system resourcing and data access rights

Our PAFCARRS analysis of data professionals' attitudes to data system resourcing and data access rights revealed that data custodians are more likely to expedite their processes for granting access to data whenever data assets clearly identified data groups and their purposes for wanting data access. This indicates that there is great value in dedicating resources to articulating and understanding the diversity of data worker groups and their purposes for wanting access to data assets. As trends show that complex data assets will seek to provide data access at scale to a diverse and broad-based user group, a cybernetic mapping of various data user groups and their objectives for wanting to access datasets can surface some fundamental

guidelines that will improve public utility of those data assets. We found immediate opportunity here for data assets to participate in reflexive awareness of their permission protocols as it relates to various user groups such as disadvantage-research groups. If this is not attended to, it is likely that future challenges may arise as those data assets aim to cater to heterogeneous and non-expert groups.

Secondly, existing complex data assets like MADIP necessarily have many processes and procedures that can bewilder data user groups. Our analysis found that one approach assets employ to guard against confusing user groups is to have distributed packets of information about asset products, processes and procedures. While this may be effective for efficiency purposes, our cybernetic approach reveals that it can hide the data asset's systemic structure and thereby hinder new users' ability to understand the complexities and justified reasons for those complexities. One immediate opportunity to remedy this would be to consolidate data asset information online by employing information feedback mechanisms and by designing asset information in partnership with some key data user groups. Doing so promises to better sensitise new data users to the processes and timescales of working with data assets. In turn, this would encourage data users to have appropriate expectations on services and communication from data asset managers. Our analysis has revealed a variety of these rich data assets are being developed or are maturing across Australian states and territories and the potential collaboration between GenV and MADIP point to the benefits of all rich data assets coming together to explore system level organisation that will be mutually beneficial. For example, given many of these data assets will eventually hope to share or link data, our analysis shows that an agreement on a national data linkage spine that can be used across all data assets would support immediate and future data access and utility across Australia.

Approaching data asset design at national systems thinking level can improve whole-of-system operations in ways that are yet to be realised. For example, over time, greater collaboration among data assets may lead to majority of those assets aligning their software and providing data in automated and standardised formats. Secondly, data access processes require that a data user's intended technique for manipulating data be vetted prior to granting access to the data. Research participants have reported that this venting process can sometimes take much longer than expected. Our systems analysis has shown that these processes can be a significant bottleneck especially where machine learning techniques are involved. Through our analysis of circular causal pathways that are exhibited in our data asset case studies, we see that training and awareness-raising of state-of-the-art techniques and tools are crucial for effective management at all levels of rich data assets like MADIP and GenV.

Lastly, funding was consistently identified as a critical priority by our case study data assets. In combining this finding with trends from literature and via expert panels discussions, we found that securing sustained funding for long-term asset development and research is a shared goal among various data assets even when their specific funding profiles differ. Even so, our cybernetic analysis shows that much can be done to future-proof data assets beyond funding concerns. For example, we found little evidence to suggest that funders have been made aware of the interconnected nature of funding strands that are required for data assets. This is an opportunity for a systems mapping education and awareness campaign that could promote a funding case if successfully delivered to funders.

In summary, when we examined the extend of the reflexive attitude of data professionals regarding the infrastructural and resourcing systems that underpin the data assets themselves, we found ample evidence suggesting that data professionals are aware and concerned of how operational and technical systems impact the effectiveness of datasets. However, because existing systems analysis don't appear to consider all key technical, environmental and social systems that impact data asset utility, the challenge remains of how to provide data to diverse stakeholders interested in the asset. A cybernetics systems mapping support can readily surface key systems components across technical, environment and social and it can suggest reflexive

approaches to monitor and steer those data assets towards responsible, sustainable and safe futures.

CONCLUSION

This project was initiated with the motivation to understand best practices of data professionals and to develop ways to support them to apply better data practises to services that seek to pivot vulnerable members of society and those experiencing disadvantage, towards human flourishing ends.

Given the dynamic nature of the ethical, technical, societal and environmental systems in which complex data sets are stored, linked, shared and used, we gave reasons for why a systems approach to these datasets and assets and why a high level of data professional reflexive disposition is pivotal if harm to the vulnerable are to be mitigated or minimised and impactful data interventions are to be created.

In this initial phase of the research project, we sought to ascertain the extent to which a culture of cybernetic reflexivity exists within the rich data systems environment as evidenced by awareness of systems components that are important to complex data systems. Through our PAFCARSS method, we sought signs of cybernetic reflexivity through evidence indicating data professionals engaging in purpose identification, boundary selection, causal loop recognition and identification of information feedback opportunities. Together, these activities indicated to us the extent to which data itself was decentred in rich data approaches to disadvantage interventions and where opportunities exist for systems improvement – where data systems can be designed with positive change in mind.

So in considering reflexive attitudes to disadvantage, in cases involving legitimacy and trust and as data systems are navigated in search of resources or access to data, we saw some evidence for cybernetic reflexivity and a willingness for those utilising data for disadvantage to embrace reflexive improvements to safe, responsible and sustainable practice. We have uncovered broad lessons that can help guide the design of various data sector-wide tools that can support both systems awareness and practical introduction and extensions of reflexive practises that seek to equip data professionals with tools to reduce immediate and future harm to data subjects and pivot their data projects towards asset-framed ends that can be steered towards promoting human flourishing.

Although we found limited evidence for wide spread decentring of data, as evidence by equally minimal evidence for cybernetic reflexivity from data professionals connected to data assets and utilising data for disadvantage, we did find evidence suggesting a strong appetite towards engaging directly with those experiencing disadvantage and towards cultivating appropriate reflexivity to enhance ethical decision making within environments high in political, legislative and regulative opacity. As demonstrated through PAFCARSS, our reflexive approach to data analysis can be cultivated individually and collectively through systems mapping and analysis of key system elements that are pivotal to the dynamics within extremely complex systems involving a variety of systems and agents interacting within technological environments.

NEXT STEPS

To illustrate and extend our findings and recommendations surrounding the necessity to raise awareness of greater data systems and processes visibility, feedback and what cybernetics suggests are potential interventions, we are planning a series of design workshops and training programs targeting various stakeholders as outlined below.

Workshops

In making use of our research finding that data professionals are seeking more support with systems approaches to data analysis, we hope to develop and deliver in-person systems mapping and design workshops to test some our discoveries. In particular, with extra resources, we hope to support members of the DCN[93] network and others working in the for-purpose sector to design robust data interventions that are future facing and strengths based.

Secondly, extra resourcing will enable us to work with government stakeholders to carry out systems mapping exercises that seek to illuminate other ways in which these institutions can bring greater numbers of jurisdictions to share data. These mappings will be undertaken collaboratively within a cybernetic stakeholder engagement framework that is cognisant of the need for data asset management to anticipate shifting norms around data ownership in both private and public spheres.

Training programmes

We hope to design a training and awareness-raising output together with key stakeholders from DCN and other government bodies, to communicate and demonstrate why data workers and data decision makers' actions are as pivotal as the state-of-the-art tools when it comes to data use towards social interventions.

2024 Student Engagement

The learnings from this initial phase of the Data and Disadvantage project will be beta-tested in postgraduate teaching engagement within the School of Cybernetics' 2024 Masters' coursework exploring data and networks.

[93] DCN is a new cross-sector collaboration formed by PRF and Infoxchange, to strengthen the data capabilities of the not-for-profit community sector to better disrupt cycles of disadvantage See "Introducing the Data Catalyst Network," Infoxchange, April 18, 2023, https://www.infoxchange.org/au/news/2023/04/introducing-data-catalyst-network.

ACKNOWLEDGEMENTS

Under the guidance of Distinguished Professor Genevieve Bell, this project was initially led by Associate Professor Ellen Broad then by Maia Gould. Their leadership were both instrumental in co-designing, research data gathering and steering this report to its final instantiation. Through its entire lifecycle, this project has been co-led, managed, coordinated, co-designed, co-conducted and co-written by Christopher Ijjo Edema Mesiku who also developed the PAFCARSS methodology.

It has been expertly supported in research analysis, report framing and writing by Dr. Louisa Shen and guided on report structuring and writing by Ellen O'Brien. The team has also been supported at various times by various research assistants from The School of Cybernetics, Juliette Parsons, Isabel Richards and Holly Hill. Their support in sourcing and summarising key literature and organising references and bibliography has been invaluable and deeply appreciated.

We would like to also acknowledge our many expert panel members and the School's Advisory Group members who have seen and guided the development of this project from an idea, through the human ethics approval processes to this report.

Advisory group members (excluding project team members)

Professor Katherine Daniell
Dr Elizabeth Williams
Dr Alex Fischer
Dr Alex Zafiroglu
Dr Brendan Traw
Brenda Martin
Nick Davis
Dr Ben Swift
Ellen O'Brien
Sharon Lopez
Professor Alexandra Zafiroglu
Associate Professor Andrew Meares
Dr Hannah Feldman
Lorenn Ruster

Expert panel members (not exhaustive)

Professor Barry Sandison
Jenifer Tennison OBE (UK)
Professor Matthew Gray
Associate Professor William Fogarty
Professor Matthew Fuller-Tyszkiewicz
Professor Roger Wilkins

Funding

This project was made possible by many people across a variety of data-interested organisations and partner organisation, and we would like to acknowledge the financial support of the Paul Ramsay Foundation in making this initial phase of the project possible.

APPENDIX 1:

PRACTICAL LESSONS & INTERVIEW FINDINGS

Focusing on MADIP and GenV, we identify and assess the structural and systemic opportunities and challenges for disadvantage work that have emerged from these initiatives. In our findings, we synthesise and set out the key themes commonly discussed by our expert interviewees, all of whom are intimately involved in MADIP or GenV. By closely attending to how each data asset has been set up, administered, and used, our research analyses the innovations, dependencies, tensions, frictions, and potentialities inherent in each system. Comparing and contrasting their approaches, our research teases out the impact and implications of systems organisation, legislation, operational procedures, technological infrastructures and tools, as well as future ambitions on the available data sets. Further, we examine how these data sets have been actively leveraged by researchers and other public agencies in the service of addressing disadvantage to date, and the attendant sensitivities and complexities around working with and on such difficult social problems. We provide various practical lessons as starting points that could be used to rethink these hurdles and hindrances.

The following section summarises the practical lessons we have made from the analysis of interviews we conducted with data workers and decision-makers linked to both MADIP and GenV. We present these practical lessons alongside findings from the interviews. These practical lessons can be understood as ways to better leverage how we use and think about data to address disadvantage. Both the practical lessons from interview findings and the findings themselves are organised around three focus areas. They are:

- Disadvantage
- Legitimacy and Trust
- Data systems

Disadvantage

1. Approaches to disadvantage within data

Different approaches to and understandings of disadvantage were discussed by the interviewees. A wide range of variables for capturing, representing, and modelling disadvantage were suggested, informed by substantively different research and personal experiences. All interviewees were conscious of disadvantage issues, despite the fact that the job roles they occupied were not necessarily oriented to disadvantage. At the same time, disadvantage was often defined in pragmatic, functional, and/or rationalised ways within a data context.

Practical lessons for addressing data and disadvantage as concluded from the interview findings

- All data workers should know about the formal definitions of disadvantage as a combination of poverty, social deprivation, and social exclusion. Awareness of debates around 'disadvantage' is imperative because they (a) involve ethics of harm minimisation and (b) enable data workers to position themselves within a wider discourse and understand the advantages and limitations of their own perspective on disadvantage. Working with domain experts (e.g. in child poverty, homelessness), while critical, is not a substitute for the data worker's own comprehension and appreciation of disadvantage. To this end, lived-experience experts ought to be treated as integral to generating data definitions of disadvantage.

- Projects aiming at disadvantage intervention using data assets like GenV and MADIP must be careful in thinking about applying those datasets towards disadvantage interventions. We recommend a move from deficit to asset framing in data approaches to disadvantage. Currently, asset framing is the dominant paradigm within research and initiatives aimed at Aboriginal and Torres Strait Islander peoples, and can be regarded as a touchstone for ways to implement this approach elsewhere. For addressing disadvantage within Indigenous communities, forming an indigenous task force or advisory group around big data can play a pivotal role in articulating strength-based data uses and promote data sovereignty. For addressing disadvantage more broadly using an asset-framed approach, the lived-experience experts mentioned above could also form a panel to serve an advisory function.
- How we include lived-experience experts in big data projects is a question that needs very careful consideration. Up to now, data interventions in disadvantage have been limited in their efficacy in terms of improving lives. This lack of efficacy can be attributed to a number of data factors; in particular, large-scale data might not adequately capture or represent the disadvantage in enough detail, across enough time, in enough numbers (within a cross-section of the population). To properly capture those experiencing disadvantage in datasets, there needs to be fit-for-purpose incentives to enable them to participate in ongoing data collection, data intervention, and data validation. Data projects should invest greater time and resources to develop and test the robustness of data collection instruments and/or mechanisms of outreach, recruitment, and retention. A potential avenue might be working with charities/service providers, incentivising and training them to reach out to self-identified disadvantaged people, for the aims of enriching existing admin or sampled data with lived-experience data. The test of whether a project is truly representative and equitable is whether these data sets can be directly used to address disadvantage through research, policy or service provision.
- It is recognised that there are no internally-consistent or agreed-upon definitions of disadvantage within specific domains (education, health, economics) that operate data projects. We need to establish working groups to define functional definitions of disadvantage in each domain, as doing so will likely bring about greater clarity for intervention and policy impact. Disagreement around disadvantage in the academic literature should inform functional definitions to ensure a diversity of viewpoints; we suggest that functional definitions state disadvantage in terms of quantifiable, strength-based data variables. For instance, disadvantage in education might be functionally defined as asset-framed parameters, including but not limited to, student educational aspiration; self- and community- identification with scholarship and achievement; teacher dedication and passion; school models and aspirational modelling; pathway plurality for school leavers; cultural and societal mindsets and attitudes towards schooling.

Interview findings on approaches to disadvantage within data

1. Given that 'all data is capta', a significant challenge described by a number of interviewees centred on how to adequately capture and represent disadvantaged groups in data sets. Administrative data sets included in MADIP had the advantage of almost whole-of-population coverage (in excess of 90%), which means that many marginalised or underserved communities are included by default. On the other hand, sampled data sets like GenV do not have the benefit of broad-based representativeness, and therefore must make deliberate efforts to ensure inclusion and equity. A health research director noted that GenV attempts to improve uptake by having 'cultural connectors' and

recruiters in health settings and translated recruitment material for minority groups[94]. Moreover, that once the data collected is made available for use, efforts must be made to ensure that the data of the disadvantaged is treated in the same way as data from other non-disadvantaged groups (equity). Data has to be used 'ethically and with the highest utility', especially around ensuring that disadvantaged participants are not excluded from research.

2. In the interviews, we specifically enquired into the richness of data, by which we mean the (a) varieties of variables, and (b) varieties of use to which these variables can be put (that is, a rich data set is one that could be used across the spectrum from summary statistics to predictive analytics). It is not only the spread of variables that matters in data, but also the variety of variables, and their stability over time. Interviewees noted that where admin data (for example from MADIP) provides breadth, often it lacks depth. Conversely, depth of data can be found in sampled data sets. A data research director noted that GenV is designed as an asset that captures intersecting social science indicators and medical science indicators of health and welfare, allowing social researchers and clinicians to model and investigate and intervene in physical and mental health alongside socio-economic participation (work) and support.[95] GenV itself includes a focus on inequity and vulnerability as part of its remit.

3. Given the need for spread and variety, data linkage is a way to offset the inadequacies in different data sets. Data enrichment also allows 'gaps' in existing data sets to be 'backfilled' from other pre-existing data to give a potentially richer and more detailed picture of disadvantage. Further, the spread and richness of variables in a data set underwrites the possibility of creating more flexible data structures that allow a higher degree of manipulation during analysis (known as 'degrees of freedom'). This flexibility makes better insights possible and in turn this stands to benefit data approaches to supporting those experiencing combinations of poverty, social exclusion and social deprivation.

4. To give a robust, high-dimensional representation of disadvantage, the kinds of variables that disadvantage is being 'translated' into is an important consideration. The interviewees provided a range of variables which they considered to be indicators of disadvantage. An algorithm developer[96] described how within a machine learning context, a probabilistic measure of success is used, with disadvantage being defined as a lack of this probabilistic success: "[T]hey're disadvantaged if their environment, if their characteristics or the reality that they're experiencing results in less chances of [them] being successful". Disadvantage is thus framed as opposite of success and life satisfaction, lacking one or more of adequate income, family time, leisure time, positive self-reflection. Other interviewees accounted for disadvantage in terms of economic (income, prosperity, work), minority status, education, and health (including life expectancy) indicators, with these being the most common measures cited. Less common indicators included provision/availability of services, geospatial analysis of neighbourhoods e.g. LGAs (Local Government Areas), family environment, psychological and social competence (including temperament). These are all examples of structural deprivation and/or cultural exclusion which are inhibitors to participation, and which have been identified in the disadvantage literature as two of three drivers of disadvantage (the other being poverty).

[94] Project interviewee PRF9

[95] Interview transcript PRF3

[96] Project interviewee PRF5

5. How the relationship among variables is then mapped also has a significant bearing on how disadvantage appears in the data. Variables are often cross-indexed and understood to be proxies for each other as well as for disadvantage in general. However, an algorithm developer also noted that quantifying disadvantage accurately and appropriately is very difficult[97]. For instance, although interviewees agreed that it is possible to measure socioeconomic disadvantage 'using education as a marker'[98], an algorithm developer also noted that lack of higher education attainment is not always an indicator of disadvantage[99]. Although the correlation between them is strong, it does not hold for many (many people are 'successful' in life without degree qualifications). A health analytics leader describes the way in which variables of disadvantage should be inter-related and often exist as trade-offs: for instance, having enough to eat means not having enough medication.[100] Some interviewees also noted that data use in disadvantage can be based on assumptions (hence the need to cross-reference). For instance, an assumption might be made that bulk billing doctors are concentrated in the most deprived communities (in many cases, this is true)[101]. However, due to historical reasons, it is sometimes the case that bulk billing GPs are instead concentrated in areas that are now affluent. In this case, the data needs to be carefully cross-referenced to ensure right service delivery. Assumptions like these need to be tested when disadvantage is accounted for in data variables.

6. The choice of variables for articulating disadvantage also has the effect of opening up or closing down possibilities for intervention. The default, dominant approach has been to choose variables that highlight and amplify the ways in which disadvantaged persons are deficient in key dimensions of life (e.g. inability to complete education, inconsistency in holding down jobs, chronic substance abuse, poor relationship skills). Part of this default approach is also to focus on risk factors of these deficiencies, which presupposes that decreasing or eliminating these risks will automatically entail an improvement in life outcomes. The academic literature has also shown that a focus on risk mitigation can compound rather than reduce harm. Variables that focus on supporting capacity to improve life circumstances offers empowerment, in contrast to the disabling measurements of the deficit approach. PRF 8 underscored that how data is collected and analysed — that is, what kinds of variables are surfaced and considered — has a direct effect on how that data can be used for improvement. A clear example cited was opting to use the measure for infant birth weights instead of the measure of infant mortality, because tracking birth weight is an indicator that still provides a window in which health interventions can be delivered. In contrast, infant mortality precludes many of the possibilities for preventative care. In the academic literature, this approach to data reporting is known as 'asset framing' or 'strength-based reporting'. [102]

7. When representing disadvantage in data, analysis techniques are also critical. PRF 4 expressed wariness of statistical techniques that could further 'disenfranchise' those who are at a disadvantage, when data sets are aggregated and averaged. Statistical analysis based on averages or broad inferences are more likely to pose this risk; disadvantage data points are hard-to-get and fewer in number, therefore more likely to

[97] Project interviewee PRF5

[98] Interview transcript PRF1

[99] Project Interviewee PRF5

[100] Interview transcript PRF8

[101] Interview transcript PRF4

[102] Kerry McCallum, Tess Ryan, and Jo Caffery, "Deficit Metrics in Australian Indigenous Education: Through a Media Studies Lens," Discourse: Studies in the Cultural Politics of Education 43, no. 2 (March 4, 2022): 266–81, https://doi.org/10.1080/01596306.2020.1828285.

appear as outliers and therefore increasing the possibility of being discounted during analysis or conversely being focussed without adequate contextuality. For example, within basic classification problems outliers in datasets can sometimes lead to classifiers with large error margins. This underscores the importance of taking extra care in choosing statistical or machine learning techniques. Instead of utilising a Support Vector Machine (SVM) for example, techniques such as k-means clustering might better locate a specific subset of people who, with appropriate context, could be identified as being or in risk of being disadvantaged.

8. Consideration also needs to be given to those who fall outside services and data systems as a direct result of the poverty, exclusion and deprivation experienced by them. Those who are defined as experiencing cycles of entrenched disadvantage can often be difficult to capture through data collection approaches. Thus, the admin/census datasets held about them are either missing or incomplete. A health analytics leader gave an example of a health intervention aiming to increase uptake of medical care among ATSI (Aboriginal and Torres Strait Islander) communities, where doctors are incentivised to enrol them on registers that capture medical visits but the burden of reaching GP services still rest on community members themselves, who are not incentivised in the same way because they are largely 'out of reach' of the health/services infrastructure[103]. (It is therefore much easier to incentivise doctors who are 'in the system'.) This Indigenous Health Incentive which is intended to facilitate indigenous access to GPs and enable data to be captured about their health, is hampered by putting the onus on them to make the effort to participate. Thresholds for participation cannot be set high as it has an unintentional deterrent or exclusionary effect on those experiencing disadvantage. Thresholds for participation include factors like opportunity costs (travel time and cost, forfeited pay or work), expenses incurred, access to telecommunications and technological devices, and language and cultural barriers.

2. Data's impact on disadvantage

Different data workers and data decision makers in different roles within GenV and MADIP have substantively different senses of the way data can impact disadvantage. A spectrum of responses was apparent, with interviewees assuming various degrees of responsibility for and involvement in making data impactful for disadvantage.

Practical lessons for data's impact on disadvantage as concluded from the interview findings

- Those who straddle the technical, strategic, and research orientations connected to GenV and MADIP show that it is valuable for all data workers within a team to have a line of sight to the strategic ends/impacts/outcomes of data projects. We recommend that GenV and MADIP leaders explore options for technical workers to have community-of-practice gatherings with those 'downstream' of their roles. The purpose of these discussions would be to share understandings of what consequences technical solutions have for making social interventions and how these technical solutions may or may not enable appropriate action to be taken on the ground.
- It is neither necessary nor feasible to require the data sets to re-orient to impact for disadvantage. Instead, ways to identify and surface data that would be relevant to

[103] Interview transcript PRF8

disadvantage, given the massive scale of both assets, should be prioritised. For instance, given that MADIP is whole-of-population, ABS could be funded to organize data sets to specifically focus on the samples or subset of the population that the indicators or variables capture as experiencing entrenched, persistent disadvantage. We suggest that both MADIP and GenV could work in partnership with more disadvantage-focused institutions and networks (such as the newly established Data Catalyst Network[104]) to uncover the data that is most salient for disadvantage.

Interview findings on data's impact on disadvantage

1. Those working in technically-oriented roles (architects, analysts, data-integrators) are more likely to see themselves as removed from the final outcomes or good that data can do. They consider their role to be (i) furnishing the platforms and tools that enable data to be used to derive insights, and/or (ii) to generate the statistical analyses/pattern interpretations from the data itself. Although they tend to be aware that their work can lead to real outcomes (for example, changes in GP service provision), they are not necessarily invested in these results that are much further 'downstream' to their work, and seen as having a separate remit to their sphere of responsibility.

2. Those working in more strategy- or research-oriented positions demonstrate much stronger awareness of the impact pathways (policy systems, intervention systems) in which data gets taken up. Many express forceful and compelling views that data work must have an impactful mission and demonstrable outcomes for the disadvantaged to be meaningful. Impactful outcomes would be very concrete improvements to health, community development, social participation, educational and work opportunities, and service access (indicators of human flourishing).

3. A handful of interviewees have experience and roles that straddle both the technical and strategic orientations that are connected with/to GenV and MADIP. They are able to articulate how technical solutions shape impact outcomes, and the ways in which these elements are closely imbricated. For instance, PRF 3 notes how the technical design of GenV, which deliberately employs a 'birth window' for sampling, has a direct impact on the kinds of participants and health parameters that can be accounted for and understood within the study. The birth window 'means that people can come into GenV at any time over the next 80 or 90 years' provided they fit the age and location requirements, where this enables migrants, refugees, and other groups newly arrived in Victoria to be included. Further, other vulnerable groups can be included by GenV 'design that allows [...] for us to have enriched cohort', leveraging collaborators at registries for additional data. It might be that through cross-referencing with a hip dysplasia registry, '80% of children with hip dysplasia' in the birth window are captured within the GenV data. Ultimately, these technical choices and solution design will determine whether meaningful conclusions for those experiencing disadvantage can be drawn from the data.

4. Moreover, differing research orientations to data also has a direct influence on how data can be effectively used for disadvantage initiatives. A data research director underscores how social scientists' and health scientists' diverging approach data – both in the kinds of data collected and how that collection occurs – shapes the kinds of claims they can

[104] "Introducing the Data Catalyst Network."

make.[105] The research director also notes that 'the real difference' is social scientists are 'really about exploring workforce and mental health and welfare', and are averse to biological specimens and physical assessment. Conversely, 'health and medical scientists think [...] it is really harmful to ask about things like bullying and how much you get paid and whether you've got a job and those sorts of things'.[106] These differing 'pain points' produce different accounts of disadvantage, and thus the design of data assets like GenV should try and 'encompass' these diverging research orientations by sampling in high dimensions and developing a multiplicity of indicators (rich data).

5. Both MADIP and GenV are at different stages of data asset development, and also have different impact assessment considerations. Although both data assets can be used for disadvantage work, neither are explicitly geared towards addressing challenges experienced by those who experience multiple forms of disadvantage.

6. MADIP's metrics for success are based on the number of projects that are using the data asset, and not on measuring whether impactful outcomes have been achieved as a result. Accountability is reported as data access and services offered to the research community (e.g. data linkages, data enrichment).

7. In contrast, GenV has an overt orientation to pragmatic impact. The aim is to deliver intervention and observational studies that lead ultimately to new or improved medications and healthcare provision. A marked difference in health outcomes for the cohorts studied is the anticipated or expected outcome. In this sense, GenV has impact assessment 'baked in' to its project design over its lifecycle.

3. Negotiating agency within rich data assets

When using data for disadvantage initiatives, the need to preserve the agency of those experiencing disadvantage and therefore prevent them from merely becoming 'subjects' in the data or study, is ethically critical. To this end, we can use data to build support services that work for them, and include them as active co-creators of these support initiatives.

Practical lessons for negotiating agency as concluded from the interview findings

- When discussing data use for disadvantage, we seldom conceive of using data for improving systems serving the disadvantaged. Data that exists around or tracks the efficacy of disadvantage interventions can be put to use in improving the effectiveness of those interventions. Data should not simply identify and track the disadvantaged where they struggle navigate social services and/or health systems, but use aggregates of this kind of data to understand how the systems themselves may prevent or hinder them from seeking or receiving help. Support systems themselves, which are largely operating under a crisis 'safety net' paradigm, are likely to need redesign towards human flourishing ends.
- As part of the larger reorientation of support systems towards enabling flourishing, funding should be made available to support charities and organisations at the frontline of working with those experiencing disadvantage to upskill towards best uses of data for those they serve. It is widely acknowledged that and documented that the for-purpose

[105] Interview transcript PRF3

[106] Interview transcript PRF3

sector is generally underserved when it comes to data skills and literacy. Giving frontline workers an appreciation of and competency in data and data analysis will enable them to identify opportunities for further involving and co-designing data-led interventions with their communities.

- Efforts to build a community of people who self-identify as experiencing disadvantage would help the broader research and data community create more effective and potentially less harmful data products. Proactively including lived experience experts in data projects will sensitise researchers further to how to treat disadvantage in a responsible and ethical way. In particular, the differing priorities and needs of those experiencing disadvantage – who are not at all a monolithic group – could be more effectively brought to the fore. Since those experiencing disadvantage are diverse, this multiplicity represents and compounds challenges for those intending to use data to support those who are living in poverty or suffering from social exclusion and/or deprivation (entrenched disadvantage). It also compounds possible risks of harm to those who are vulnerable when their voices fail to be taken into account. A lived-experience community can make sure that data projects have their best interests at heart and also beta test intervention solutions intended to be helpful.
- Data work for and with those experiencing disadvantage should also integrate opportunities for lived-experience consultation in the data use processes. Feedback mechanisms or iterative approaches ought to be established for communicating with lived-experience experts. Mapping systems and processes around current or future rich data projects to identify the salient points at which consultation is needed can be a useful step in intervention co-design.
- Given the complexity of the GenV and MADIP data assets, every interviewee expressed a reflexive awareness of how they themselves as individuals with agency were situated in the broader systems underpinning their work. In particular, the interviewees discussed (a) their philosophical understandings of the data assets and the kinds of good it might produce, as well as (b) their positions and responsibilities in relation to the data subjects included in the asset, and (c) the attendant political and cultural sensitivities arising from these complexities.

Interview findings on agency within data assets

1. Views of and dispositions towards the data and data assets as a dynamic resource with significant potential to better life outcomes were common. A data research director commented on 'the passion' behind their data work, where the aim is to figure out 'how do we create a system that enables impactful research to happen on a diversity and at a scale match to burden of disease that isn't currently possible'.[107] The overriding motivation is not just joy from making a difference, but the prospect of making a substantial difference by 'get[ting] faster, more meaningful impact on a more diverse scale'. The potential of data is in enabling knowledge – to build an asset that unlocks our ability to 'ask lots of very diverse questions very, very efficiently'.

2. Data was also seen as a way to address problems we have not yet measured or defined carefully, where it to be leveraged properly. A health research director stated that 'the use of data to drive system change is totally underestimated', where 'we don't use data to drive human capital systems in the way that businesses use data to drive their systems'

[107] Interview transcript PRF3

especially ‘for the things that we care the most about, i.e. people’. For instance, ‘we use no data to look at whether or not the multi-billion-dollar investment we make, particularly in services, makes any difference whatsoever’. Programmes might ‘increas[e] inequity rather than decreasing it’; data could be used to surface these as-yet ‘invisible’ disadvantage issues.[108]

3. In relation to the use of data for disadvantage, there was consensus that this could be effectively achieved, coupled with acknowledgement of their own privileged positions in the data system in relation to the disadvantaged. A digital humanities researcher, in the course of using data for disadvantage research, is careful to distinguish disadvantage as a systemic rather than personal problem, having regard to the researcher’s position in the systems (data, academia, social, educational) as 'the actual antithesis’ of those experiencing disadvantage[109]. Often, socio-economic research does not receive the benefit from lived experience experts, leaving the privileged researcher to ‘wonder’ in the process of talking to the underprivileged, whether they are ‘qualified to speak for these people?’ and whether 'they want my help?’, and even further, ‘what do they think about me?' In this sense, there is strong awareness and even significant worry around the agency of the relatively privileged being used to serve the underserved, without stripping away the agency of the latter.

4. A number of interviewees shared strong convictions that the relationship between themselves and data subjects should not be a top-down or hierarchical one, where the subject is powerless to have agency or understanding of the data. A health research director outlines the need to actively ‘feed back’ community data to communities, so they can both understand and use it for their own benefit.[110] In order to achieve this goal, it is necessarily to consider questions such as 'how do you get data out to communities so they can see what's happening in their communities for their children?’ and ‘[w]hat sort of visualization, accessibility to data do they need?’. Catering to ‘a whole series of end users […], which is everything from parents and children themselves, right through to service providers, communities, professionals, policymakers, researchers’ moves beyond the outdated paternalistic paradigm of more traditional forms of social science or health science enquiry.

5. This sense of responsibility to data subjects was especially heightened for those researchers working with Aboriginal and Torres Strait Islander groups, where active co-design and input from the communities were essential for the legitimacy and robustness of the data project. A health analytics leader notes that it is important to be attentive to whether subjects feel safe[111]. This entails more than the basics of ‘go[ing] through ethics for media’, ‘go[ing] through MADIP training’, and ensuring that ‘team members know what to do’. Instead, it extends to co-developing the project – for instance around ‘data governance structures’ in consultation and accord with the community who are embedded in ‘a data committee in the Department of Health where people talk about Indigenous health’. These community leaders ‘can become part of my project team’ and help think through how data can be best harnessed to ‘help [them] make better decisions using the data [they] have access to.’ The primacy of the focus is around helping communities and achieving community buy-in for the project. Ways of working that do not

[108] Project interviewee PRF9

[109] Interview transcript PRF1

[110] Project interviewee PRF9

[111] Interview transcript PRF8

serve that goal can and should be sidelined. For instance, if conventional outputs or pathways such as publication may not necessarily be the right avenue for sensitive data, publication does not necessarily need to happen. Cooperation with Aboriginal and Torres Strait Islanders communities should enable them to hold meaningful sovereignty over their data and derive meaningful use from it.

6. Part of putting community data into the hands of the community also entails enabling those experiencing disadvantage to have agency over their own story in data. A health analytics leader underscores the need for not just consent in the form of 'a letter of support from communities', but that this support can and should only come when the use of data is for tackling issues that are a 'priority for communities', rather than simply what the privileged researcher thinks is a priority.[112] When data is used for addressing disadvantage in a way that is divorced from the realities of lived experience, the narrative that arises can often be alienating, feeling 'like somebody else is doing the talking about us and it just doesn't resonate'. For those experiencing disadvantage, the pertinent question becomes, 'Why are you talking about this other thing that doesn't really matter much to me, and not talking about these other things that I want you to talk about?'. More data does not compensate for working at cross-purposes with mismatched outlooks on the help that is required. Those experiencing disadvantage must have meaningful participation in any efforts to use data to help them.

Legitimacy and trust

1. Consent and social license to operate

Receiving and maintaining consent and social license to operate emerges as an ongoing concern for the viability of rich data projects. Whereas MADIP can rely on passive uptake of data, GenV has to actively recruit from the public. Consent and social license therefore function differently for the two data assets.

Practical lessons for consent and social license as concluded from the interview findings

- We suggest continual ethics reviews of dynamic long-term consent at set intervals, in order to continue to maintain strong social license to operate. Over the long term, individuals' (and families') circumstances change, as does the legal landscape of parental and individual rights. Data projects over the life course face substantial risk of controversy and harm if these changing attitudes, preferences, accountabilities, and norms around sensitive personal data over a lifetime are not accounted for within the study. Disclosure of medical findings to participants or participants' legal guardians (when they are minors) is already a fraught area in medical ethics, and are further complicated by time and age, and thus needs to be managed over the life course. Longitudinal life course data projects require reviews of whether the consent continues to be fit-for-purpose over time.

- In the same way that dynamic consent enables the data project to update data types or data collection parameters, so too should it enable the participant to receive ongoing participant support services, so that truly informed consent can be maintained going forward. Means of delivering this support might include consent counselling and consent management, so that the complexities of medical study and its implications can be managed. (This support would need to include language support.) We suggest that active

[112] Interview transcript PRF8

consent management will enable, in the long run, higher rates of participant retention than simply passive, 'hands-off', or 'one-off' consent because it offers a stronger ethic of care towards participants and provides continual opportunities to emphasise the impact and benefit of participation.

- Given that the public interest test is integral to social license, we suggest that government and public institutes (in the broadest sense) continue to carefully manage the ringfences around data gathered from the public, especially around establishing strong guidelines and public principles around managing data release to, use by, or funding from for-profit, commercial entities. Where data assets are operating under divergent regulations or commitments to commercial data sharing, incompatibilities at this political level will have direct implications on data linkage possibilities. GenV's linkage to MADIP will likely face hurdles of this kind.

- Part of data asset management is forecasting and futureproofing around likely contingencies that arise in the future. We suggest that anticipating shifting norms around data ownership is critical, as the data ownership conversation is already changing in the industry space and in the civic space. For instance, commercial products already exist to enable the individual to sell their own data to search engines, while standards like Solid[113] are being pioneered for the non-profit storage and access control of personal data. That is, data is increasingly seen as the private property of the individual (with commercial value). The civic model of data ownership, on the other hand, emphasises data is a collective commons, owned by the public at large rather than belonging to any government department or public entity. These normative perspectives have implications for data assets. MADIP currently does not require direct consent or enable opt-out from the data set (that is, an individual cannot request to be deleted from admin data). If data-as-personal-property becomes a widely-accepted social and legal principle, it raises the question for MADIP of how to cope with deletion requests. If GenV accepts involvement of privatised medical and pharmaceutical companies, it is likely the project will have to consider whether under the private data ownership model, participants need to be able to be compensated for providing their data, or to opt to withdraw themselves from for-profit endeavours where they have initially consented to provide data under a charitable framework. Under the public data ownership model, the implications are different again, in particular around the duty and obligations of government agencies to handle public data as public property serving the public interest, which may demand, in principle, that they are less able to guard and administer data access.

Interview findings on consent and social license

1. Consent is critical in particular for longitudinal studies, given that obtaining useful and robust results (especially trends over time) hinges on participants' long-term willingness to stay within the cohort. GenV is designing dynamic consent so that participants do not have to re-consent as the study's parameters and investigations change across the decades (the choice to opt-out is retained). GenV's consent process relies on active recruitment 'at every single birth', and therefore is increasingly set to scale with intent to bundle GenV consent processes with the consent processes for universal hearing screening and neonatal screening going forward.

[113]"Solid is a specification that lets people store their data securely in decentralized data stores called Pods. Pods are like secure personal web servers for data. When data is stored in someone's Pod, they control which people and applications can access it." https://solidproject.org/

2. Appeal to altruism and idealism underwrites GenV's recruitment, with many participants expressing a willingness to contribute their data for the advancement of science and medical knowledge. Some interviewees like the health research director reported that parents hearing about the study are excited and "feel like they are contributing to something really important"[114]. The same researcher noted that the ultimate social contract is for members of society to trust and contribute data to GenV as a project in the service of the "good of humanity"[115]. Researchers also acknowledge that reaching that level of social license requires immense work and is "the ultimate social contract with society."[116] This social contract has to be actively built and contrasted with other data practices that harvest data in ways that perpetuate social ills. (Examples of the latter might include for-profit, commercial data enterprises.)

3. Part of maintaining social license is also to be seen as accounting for and including those most in need in the population and correcting for inherent biases. A health data researcher noted that working with "cultural antenatal services" was important to recruit and retain indigenous parents as well as other minority background parents, even as "low-literacy, low-income families" are less likely to be captured[117].

4. In contrast to GenV's active consent, MADIP relies on passive consent – that is, consent that has already been given in the process of collecting admin data. Further consent for linkage and use is not arriving from the population level, but rather becomes approval from the jurisdictions and partner agencies, who closely control data access and use. In this sense, data project legitimacy derives less from collective population consent and more from general confidence and trust that approval will be correctly administered on behalf of the population by public agencies and services. That is, the social contract between the public and the government is critical to MADIP legitimacy.

5. MADIP leverages ABS's strong social contract with the public, entrusting data management and oversight to the agency as the accredited integrating authority. The creation of MADIP centred around finding the right ways to address and overcome concerns over data privacy, security, and integrity. Describing the events that surrounded the creation of MADIP, Data services personnel[118] stated that "as each of those changes rolled through, there were a lot of obstacles and challenges that had to be overcome. So, there's certainly always, in the back of our mind, the question of social license."

6. Public trust that the government and public agencies are strict, secure data handlers is foundational for GenV and MADIP to continue operating. A data access manager cited research that indicated the public at large trusted government with data more than it trusted private companies.[119]

7. It may be necessary for government to also build on a sense of public data as a public good to maintain social trust in their operation of data assets. A health research director

[114] Project interviewee PRF9
[115] Project interviewee PRF9
[116] Project interviewee PRF9
[117] Interview transcript PRF9
[118] Interview transcript PRF7
[119] Interview Transcript PRF2

speaks about the importance of promoting the efficacy of data across society, where it is necessary to be able to 'talk about data across different professional paradigms so that other people see the point of data, so that it doesn't get relegated into geeksville or nerdsville depending on how you think about things'. Beyond researchers, the general public 'see[s] data as very relevant', and therefore it is good to see 'data in the hands of a range of end users, including services and communities and populations'[120]. Satisfying the public interest test is a widely-held expectation for government rich data projects.

2. Building trust and relationships

Relationships among researchers, data service providers, data custodians, and jurisdictions are all based on goodwill and building very genuine relationships and understanding between collaborators and partner agencies. It is apparent from the interviews that the data system runs on trust and implicit knowledge of institutional networks.

Practical lessons for building trust and relationships as concluded from the interview findings

- Perceived conflict of interests and power asymmetries determine the level of trust and the closeness of cooperation between stakeholders in rich data assets. Where needed, we suggest that efforts can be made to overcome erroneous beliefs that the parties are working at odds with each other. This effort might include foregrounding more strongly shared goals and shared wins, and highlighting the benefits of collaboration for all involved. (Refer to the exceptional cooperation during COVID as an exemplar of potential.) For instance, efforts to demystify the consequences of sharing data with MADIP is likely to be crucial, especially for jurisdictions. Clarifying how little jurisdictional authority or autonomy is actually ceded to federal agencies, where data is not simply 'subsumed' into a larger Commonwealth pool, could be one message that might serve to allay fears, especially as such a perception carries burden of legacy relationships between state and federal governments.
- Given that data assets' operations rely overwhelmingly on trust relationships; we suggest that 'connection transparency' would be beneficial to all parties. By 'connection transparency', we mean making visible the key contacts at different institutions who are critical to facilitating data processes. Currently, knowing who to contact in what situation appears to be largely implicit knowledge held by key staff members at different stages of a data project. Relationships on which institutions depend may be largely sustained by the presence of a few key data decision-makers at each place. This 'hidden' information could be turned into more formalised institutional knowledge, so that relationships can be cultivated more broadly by the institution, and key data decision-makers could be better supported in building and strengthening these relationships.

Interview findings on trust and relationships

1. Managing relationships with stakeholders with different levels of investment in the data assets requires different strategies. A data services personnel[121] identified that within the 'Commonwealth space, most of the agencies that are sharing data with us are also really big users of the data themselves'. This means there is a mutual win derived from collaboration, as these agencies are 'getting that benefit back and that firsthand

[120] Interview transcript PRF9

[121] Project interviewee PRF7

experience back'. These more established relationships are also built on the basis of 'established funding arrangements'. However, where there is a lack of shared benefit, relationships are much harder to sustain. When it comes to asking jurisdictions for data access, the 'value proposition for them is quite different' and that necessarily means that 'the trust situation with them is newer'. Jurisdictional data can already enable states to gain insights into their own administration, and moreover many states have their own data integration infrastructures (e.g. Centre for Victorian Data Linkage - CVDL). Therefore, states may feel they have no reason for them to share that data with similar federal infrastructures, and perceived risk an erosion of their authority and autonomy around state data. This sense is exacerbated by the fact that MADIP was not originally conceived as a jurisdictional initiative but rather a federal one. Thus, for MADIP, building relationships with 'people we have not worked before' requires more effort.

2. Relationships are also built without any legal mechanisms to compel collaboration or to override or resolve conflicts or friction. While some MoUs exist for both MADIP and GenV, which shapes expectations around co-operation, the relationships overwhelmingly remain a 'trust and partnership kind of proposition' (Data services personnel[122]). If there is insufficient trust in place for any reason, then data use can be substantially hindered. An algorithm developer describes how a mismatch in technical capability or knowledge amongst stakeholders can lead to a need to verify each step of the data use process[123]. Where a new algorithm is used, it must be checked by the relevant team, who may or may not be 'familiar with the analytic methods'. Subsequently, another check 'that the outputs are good enough [...] to incorporate in [...] research outputs in papers or technical reports' is required. Within an already controlled environment, 'either [an] automated or trustful approach' is seen as the key to reduce relationship friction.

3. The kinds of the relationships among data asset entities also underwrites the integrity and viability of the data asset. For GenV in particular, the mutualistic, collegial nature of the relationships between MCRI, Royal Children's Hospital, University of Melbourne Paediatrics, and partner hospitals and labs means that there is overwhelming alignment of goals. These institutions also have long histories of collaboration. Partner hospitals are also critical for building relationships with potential cohorts, as 'each hospital sends a letter to every prospective parent'. As a data research director notes, this means that GenV is effectively 'authorized by the people that they already trust', at a 'supra-level' and 'also at the service level'.[124] Smaller rural hospitals also derive benefit from belonging to an ambitious study.

4. In the MADIP system, there are asymmetries in the relationships between jurisdictions, data custodians, ABS data services, and researchers because jurisdictions and custodians can refuse access to the data, which puts pressure on ABS data services to maintain good relations with them. Managing these relationships is largely informal and reliant on assurance and negotiation, either without basis (with the jurisdictions directly) or based on MoUs where they exist with MADIP data custodians. In contrast, ABS's relationship with researchers, centres on performing due diligence, and is also more formalised, proceduralised, and transactional. Under ANDII, these asymmetries could shift as researchers become data donors under their institutions.

[122] Project interviewee PRF7

[123] Project interviewee PRF5

[124] Interview transcript PRF3

3. Risk management and risk tolerance

Risk management was identified as a critical factor in developing and/or enabling data projects. In MADIP, risk management centres on the granting of data access to researchers and relevant civil servants. GenV, being at an early stage, has yet to encounter such problems and is aware of these difficulties within MADIP as it undertakes initial scoping around data linkage with MADIP. For GenV, risk management is concerned with the scope and ambition of the asset and the logistics of sustaining data capture.

Practical lessons for risk management and tolerance as concluded from the interview findings

- We suggest that differing outlooks on risk must be acknowledged so that data assets can manage their expectations of each other as they explore better ways of working together. Attitudes towards risk can be found implicit in the way data assets operate. Although strong in-principle agreement between data assets, data custodians and linkage projects around the benefit of collaboration exists, fundamental differences in the founding purposes and agendas between stakeholders can hinder efforts to realise the benefits of rich datasets. For instance, GenV has been founded upon a vision of improving the health and wellbeing of a whole generation of Victorians through observational and interventional studies. GenV's mode of operation is therefore about exploring realistic possibilities of realising that ambition. MADIP emerged as a data integration and sharing initiative strictly determined by several policies and legislative requirements (typified by the Australian Privacy Act and the Census Act). MADIP's mode of operation is therefore about maintaining a high level of public trust and social license which obligates MADIP to have a lower risk tolerance relative to GenV.

- To alleviate often-legitimate concerns data custodians express during data access discussions, we suggest that a formal risk rubric of each access request is completed, where this could be better achieved by facilitating risk information flows between stakeholders. Both data custodians ('providers') and researchers ('requestors') should exchange risk information.

  For instance, the current form for accessing ABS's DataLab (ABS DataLab Project Proposal for Detailed and Integrated Microdata) requires researchers to provide project details in order for ABS to conduct their own risk assessment. However, researchers themselves have little or no visibility of the risk considerations that worry each data custodian; therefore, researchers are writing proposals that may not clearly address such concerns. As such, it becomes very onerous for the ABS Data Services team to undertake due diligence as the mediator. We suggest therefore that each data custodian provides clear guidance on the risks that are most serious for them, referring to precedents set by them in previously-approved MADIP projects. Researchers, through the Data Services team, should also be able to clarify data custodians' threat perceptions on a case-by-case basis.

  Indeed, researcher's perceptions of risk are likely to differ from those of data custodians. We therefore further recommend that the ABS microdata application form makes provisions for discussing data risks from the researchers' perspective (prior to ABS Data Services conducting their own risk analysis). Having the researchers' own assessment of their risks could also help data custodians better understand the extent of risk and in turn this can lead to streamlined data access discussions with the relevant data donors. These changes, if implemented, would make risk assessment more of a shared responsibility in the MADIP system.

- The duties and cultures of data custodianship are a hidden factor in risk assessment. Currently, there is a perception that custodianship operates in one dimension – as the custodians exerting control over and ensuring security of the data for which they are responsible. To change this perception, data custodians would need to strengthen best practice of managing the data asset in a way that allows the maximum realisation of its potential to address problems of public interest. This is another dimension of the stewardship function of data custodians, where the custodians' duties and responsibilities to public data can readily be seen as the responsibility to act as an enabler of data-for-good initiatives.

- Data custodians can be legitimately worried about risks of releasing data, but counterfactuals around the risks or drawbacks of not granting access to data, which has knock-on effects for effective intervention into disadvantage, should also be factored into decision-making. For example, we suggest a recurring, periodic discussion around best practice would benefit all MADIP data custodians.

Interview findings on risk and risk tolerance

1. In MADIP, risk centres in particular around data access, and its assessment is contingent on human sensibilities within organisations that are data custodians, rather than on formal risk measures [The ABS data access application form and eventual acceptance processes and procedures are considered as a risk assessment undertaking by the ABS Data Services Team]. Given that much of the data is very sensitive, this kind of conservative attitude is prevalent and aimed at avoiding high-risk breaches (such as re-identification (Data access manager[125]), even if the incidence of those breaches would likely be low or negligible. Some interviewees identified risk aversion as one of three major roadblocks to efficient data use in MADIP (the other two being time and money). Roadblocks caused by risk aversion result in substantial time wasted waiting for decisions from the data custodians via the ABS Data Services. A digital humanities researcher also noted how 'normally, [...] every single project needs every single data custodian's approval', which can be 'enormously time-consuming'.[126] This process means 'all that it takes is one data custodian to be slightly risk averse or to not quite understand the project and the whole thing can just come to a grinding halt'. The implication is that data custodians effectively have veto power over data projects.

2. Risk management in MADIP tends to be based on providing reassurance around threat perceptions. A data services personnel stated that with data projects, it is 'never a question of can you do it', but more about 'should we do it'[127]. Where a data project has satisfied all other criteria around 5 SAFES, and where it can be supported by the available data and tools, getting approval for data access from data custodians can remain a challenge. Projects have to be undertaken in a way that is 'consistent with what the custodians of this data are comfortable with us doing'. If a data custodian is 'far less comfortable or they feel they have fewer protections', and 'feel like they're opening themselves up to a really unnecessary risk or that the ABS or one of these researchers are going to do something really unexpected with their data', then further negotiations will be required (see relationship building). Perception management is central to getting data projects across the line. As a data research director rightly suggests, a generalizable potential drawback of this kind of risk aversion is a reduction in the quality of research agenda making use of data.[128]

3. Risk management in GenV internally expresses itself in the degree to which GenV staff are cautious about the vision or aspirations of the leadership. The same research director notes that internal staff can be averse to ambitious scoping, and can insist upon reining the aims back in. If feasibility assessments come back positively on pragmatic execution, reviving the larger vision can mean 'a lot of work to then re-expand'.[129]

4. In contrast, GenV sees external risk as involving the management of partner institutions, so that a lack of timely collaboration does not result in GenV losing the opportunity to collect (time-sensitive and time-contingent) data. This type of data loss can be the result of protracted efforts to coordinate and agree on partnerships and recruiting cohorts.

[125] Interview transcript PRF2

[126] Interview transcript PRF1

[127] Project interviewee PRF7

[128] Interview transcript PRF3

[129] Interview transcript PRF3

Often birthing hospitals and health providers require several meetings and discussions before agreements are reached. A data research director describes how GenV has to onboard and maintain '58 agreements with 58 birthing hospitals across Victoria', plus 'ten pathology providers across the state'. This entails 'endless agreements, endless ethics' and where if that is not kept up, data is lost due to not capturing health changes or events that happen to the participants. This external risk affects the 'value of the ultimate product'.[130]

5. Another aspect of external risk to GenV is tied to various potential funders – both government and private sector bodies often requiring numerous meetings and negotiations prior to allocating funds; GenV has so far secured the initial tranche of funding for recruiting cohorts and setting up the foundational structures of the asset (data ingestion, linkage, use).

Data systems

1. Right access to right data

Gaining and managing access to data sets within MADIP in particular is a current pain point for both data users and ABS's Data Services staff. For MADIP, ABS Data Services are the intermediaries or brokers that facilitate access to the data asset; much of this facilitation involves vetting requests for access and obtaining permission to access the data from custodians. Final technical access to the data itself is given via profiles to log into ABS-created virtual machines (The DataLab). Efforts are underway to improve current processes through platforms like the current beta trial of the myDATA Portal.

In contrast, GenV is not currently at the stage where data access has scaled, and therefore GenV's data sharing challenges are less than those experienced by MADIP staff. To the extent that GenV does scale access in the future, a different set of challenges will likely arise because it aims to make data access and analysis available to a greater diversity of user groups (including the community, private sector, government, and service providers).

The process for eventually integrating GenV with MADIP is in the early planning phase, and even so, similar challenges to those already discussed in this report have readily been identified. The interviewees further elaborate on the following issues around data access:

Practical lessons for accessing the right data as concluded from the interview findings

- If the ABS Data Services did not exist, each researcher or data user would have to contact and negotiate with each data custodian separately to gain access to the different data sets that each custodian holds in the MADIP asset. The data user would have to navigate the complex legal requirements governing each data custodian, as well as the complex organisational relationships among the data custodians. This would mean an overwhelming and highly-repetitive process for all parties, with vastly increased administrative overhead. Thus, although the ABS Data Services appears to be (or is experienced as) a bottleneck to accessing MADIP, in actual fact they offer an already streamlined access process. Alongside existing onboarding and informational processes, we suggest a revision/refresh of web presence to consolidate scattered MADIP information into a single place online (an interim solution would be to create landing pages that collate links). The result of this consolidation would help new MADIP data users better understand the function of the ABS Data Service and the systemic complexities inherent in MADIP. Such instructional

[130] Interview transcript PRF3

materials will sensitise data users to the processes and timescales involved in working with MADIP data. Given that new users are required to have a myDATA Portal profile, this material could be embedded into profile creation and sign up.

- A more far-reaching solution for improving data users' experience of the MADIP system would be to automate existing processes that are almost entirely manual. With the exception of myDATA Portal, workflows depend on back-and-forth emailing to ask for and clarify information, and arranging meetings to discuss access requests and consent. Form-filling is also largely paper-based. Setting up DataLab VMs also depends on this manual process, as does the final output vetting. Transitioning MADIP infrastructure to an infrastructure like ANDII would support the different stakeholders to streamline workflows and collaborate more easily, consolidating communication and paperwork for each data user/access request into one secure shared service, available to all parties to log into and have visibility.

  If MADIP transitions to a cloud-based service such as ANDII, which is currently under construction, some challenges stand to be minimised because ANDII offers the opportunity to design and build a comprehensive cloud service (under the DAT Scheme) that would not only offer streamlined entry and exit processes into data sets, but also offer a national linkage spine, data integration environment (architecture), and analysis environment for data users, data services, and data custodians.

Interview findings on right access to right data

1. Data users express a clear understanding of the need for rigorous and comprehensive data access procedures on the grounds of data security and ethics governance. They do not necessarily have awareness of the operational, organisational, and legal challenges that ABS Data Services have to navigate in order to progress their access applications. Administrative hurdles often lead to a frustrating experience, where the data users often feel mired in a slow process where they have to rely on indirect communications through the ABS Data Services. Progress and process opacity, where there is a lack of visibility as to how a researcher's data access application is being funnelled through the pipeline, is therefore a significant hurdle for cooperation amongst researchers, ABS Data Services, and MADIP data custodians. The myDATA Portal which has been recently commissioned by ABS to facilitate data access within the DataLab environment, will eventually allow for greater visibility of the data access application process, as well as application history.

2. Facilitating MADIP access consists largely of managing relationships and data custodians and sometimes jurisdictional consent and approvals. Given the variety of possible state and federal agencies involved in 'feeding' data into MADIP, and given the variety of pathways for ingesting data, no single overriding streamlined process exists for ABS Data Services to execute. Further, when requests to access data that is not already linked within MADIP arise, additional consent must be sought from the relevant bodies.

3. ABS Data Services are also responsible for 'research translation' in the course of administering access; that is, they are tasked with explaining the research project and methods, and making the case for data access on behalf of the researchers who apply. This function often means that the ABS Data Services team are forced to contend with 'translating' complex or difficult research methods or techniques to non-expert data custodians, a process that can be time consuming.

4. Further to the initial 'entry' phase, ABS Data Services also has to administer an 'exit' process for data users (output vetting). The 'exit' involves vetting all data-derived results to ensure data privacy is preserved on the publication of the research, as well as deleting data sets from the VMs and checking that no unauthorised copies have been made to ensure data security.

5. Data service providers (ABS Data Services) openly acknowledge that their processes currently 'focus too much on governance' Data services personnel [131]. She outlines the extensive processes for protecting the privacy and rights of data subjects within MADIP, which are managed through multiple processes and layers of governance. For example, ABS Data Services have to 'manage getting all of the approvals in place from various people that are required to approve access for these types of projects', while at the same time 'mak[ing] sure that all of the researchers have signed the appropriate undertakings and completed their [5 SAFES] training'. Further, '[a]ny mistakes in these processes are logged and reviewed by teams in MADIP' as incidents to ensure 'safe project, the safe people, and [...] safe data considerations'.

6. The data and statistical response to COVID, which enabled real-time reporting of infection and vaccination and contact tracing, offers a case study of how the MADIP system can operate at exceptional efficiency. This extraordinary circumstance is an exemplar of what can be achieved when all parties have 'buy-in' into a shared goal that everyone recognises is critically urgent and important. Acute awareness of the need for timely and accurate COVID data reporting meant that parallel processes of working were created (contra the usual linear processes). For instance, data preparation was occurring as the consent process for that data was also underway, while at the same time the analysis architectures for data visualisation were also being stood up. Health agencies across the state jurisdictions were cooperating with the federal agencies to feed data into MADIP, while data researchers in universities were validating the data input to verify official numbers. In addition to the parallel processes, constant communication via teleconference among all stakeholders meant that critical informational feedback loops reduced knowledge gaps around the COVID numbers reporting logistics and rates (e.g. updating lags), improving coordination. Comparative statistics were generated by other research agencies (such as CSIRO). Identifying communities with low uptake of vaccination (often culturally and linguistically diverse groups) was a key pragmatic result of this cooperation, which allowed government to direct more resourcing to those affected in order to support Australia's wider COVID-zero policy.

2. Appropriate and fit-for-purpose data tools and platforms

Having the right data tools and platforms for data management and analysis are critical for efficient data work. We found that although MADIP – and GenV going forward – uses or seeks to use industry-standard languages and software programmes for data manipulation, the subsystems that feed data into these assets do not necessarily do so. This discrepancy acts as a drag on whole-of-system operational efficiencies in the ways detailed below.

Practical lessons for using fit-for-purpose data tools as concluded from the interview findings

- Although upgrades can require substantial commitments of time, resource, and expertise, we suggest that it is critical for data custodians to undertake this where it is required (i.e.

[131] Research Interviewee PRF7

where industry-standard tools are not already being used). As part of the DAT Act (2022), which legislates for improving data integration, data custodians are obligated to facilitate public data sharing. Aligning key software programmes used in-house with those used by MADIP will improve whole-of-systems operations. In particular, this alignment would provide the opportunity to support MADIP data ingestion by pre-processing data using the same or equivalent tools. Currently, there are no prescribed data standards for data custodians to pre-process the data before it is ingested into MADIP, and therefore all data formats / structures accepted into the asset must be first cleaned and standardised by MADIP integrators. Although final quality assurance by the MADIP team is non-negotiable, we suggest that data custodians can work to provide data in such a way that reduces MADIP workload. Using industry-standard tools means that the entire data sharing system will not be slowed down by any one custodian operating within inefficient, out-dated data environments.

- Within data custodians, training and awareness-raising for management is especially crucial given that without appropriate technical literacy in new tools, it becomes harder for them to weigh up technical options and take the right decisions to facilitate better data integration and analysis within their organisation. The level of technical literacy required would not be at the level of implementation but rather at the level of understanding the ramifications of use (or lack thereof), especially for emerging techniques like machine learning. Additional benefit of technical literacy would include upstream benefits to MADIP users, as a broader understanding of data tools and techniques would make access requests more intelligible to these custodians. If data custodians have the same understanding of data techniques as the external researchers applying for MADIP access, then the data access application process would become streamlined.

Interview findings on appropriate and fit-for-purpose data tools and platforms

1. Data custodians / workers who are dealing with data sets have noted that 'clunky tools' limit their ability to effectively handle and interpret the data. PRF 4 describes how lack of institutional access to industry-standard software and hardware means they have to resort to using packages that are unfit for big data analysis. When using unsuitable programs, analysis suffers from slow processing, minimal or no error detection, and platform instability (crashes). Using the right tools would 'exponentially' speed up data analysis, especially given the size of the data sets.

2. Lack of right tooling arises because of management's reluctance to approve changes to tools and platforms, in large part due to the technical difficulties and organisational complexities of upgrades. Specifically, retaining the current technical environment allows the current security measures and protocols to be preserved; upgrading would require an evaluation of the security repercussions and new security solutions to be developed. This arrangement, however, means that the prohibitive environment only permits 'blunt instruments' to be used in a bid to ensure that 'sharp instruments' do not expose the organisation to security breaches, rather than making their internal systems more robust overall. There is no cost-benefit analysis as to whether this kind of environment should be maintained; instead, negative perceptions of technical change seem to dominate and promote a (perhaps disproportionate) conservative culture of controversy avoidance.

3. Technical upgrades within an organisation do not necessarily have to be blanket upgrades across all machines; bespoke upgrades with selected machines can also be implemented. This is especially pertinent in cases where available licensing is limited to

a few computers or users. Despite technical upgrades themselves not needing to be organisation-wide, decisions to upgrade would nevertheless necessitate coordination and consensus across internal teams who have divergent opinions on the best tools to use and who might have access to them. Data teams within sub-systems are themselves often insular and have incompatible work processes (e.g., Agile versus Waterfall project management styles).

4. In order to facilitate technical upgrades, there is a parallel need for upskilling and capability building in subsystems' personnel, not only at the analyst but also at the managerial level. Data analysts are unlikely to find upskilling difficult given existing technical backgrounds and 'constantly learning new programs over time', and thus overheads for training would not be onerous. In certain cases, where more specialist skills are required, for instance when platform migration to a data lake or SaaS model occurs, there might be extra need for outsourcing or more advanced training, which would likely also be limited to a few key personnel.

3. Funding/resourcing structures and priorities

Due to their differences, the funding profiles for GenV and MADIP are also necessarily different. However, there are shared concerns expressed by the interviewees, in particular around sustained funding for long term data assets and research. Currently, most funding is short term in nature, and risks falling away after the initial phase of a data asset's lifecycle. More specifically, funding can be divided into these strands:

Practical lessons for resourcing data assets as concluded from the interview findings

- Funders should be aware of the inter-related and inter-dependent nature of the strands of funding that data assets require. Data use is likely to become more resource-intensive over time, as the data asset evolves. Funding has significant implications for future-proofing assets like GenV and MADIP. As technologies for working with data become more sophisticated, and as access demand grow exponentially, so too will administrative burdens and complexities increase. Efficiencies will have to be found in order to make the most effective use of funding; an example already in play is the making available of 'Tier 3' data in MADIP --- that is, custom-linked data paid for by researchers' grants and optionally made available to other researchers at no further cost to them.

Interview findings on funding/resourcing structures and priorities

1. Maintenance: MADIP has funding secured for maintenance (as an 'enduring asset').1 Maintenance includes ensuring that the asset is continuing to ingest federal/Commonwealth (and increasingly other sources) data at regular intervals. GenV has secured funding to set up the data asset for the first 10 years, which includes seed funding and development funding from state government and charitable sector. However, given the longitudinal ambition for GenV (100 years+), these early funding tranches will not be enough to sustain the next phases, and fundraising will be an ongoing burden.

2. Enhancement: GenV faces ongoing costs for improving the data. Three ways of ensuring data richness are deployed: data integration/linkage, data collection, data enrichment (data enrichment comes from additional capture of participant data via wearables and

other automated means and factoring in the ingesting of other relevant data for future use when the legislative environment allows). MADIP data enhancement involves ingesting more data from state jurisdictions (increasingly allowed for under the DAT Act (2022)). Jurisdictional data requires extra funding beyond maintenance funding.2 Much of the funding for this integration comes in a piecemeal fashion from researchers requesting custom integrations and paying for that out of research budgets.

3. Access/Use: GenV is not at a stage where its datasets are widely accessible; however, significant forward-planning for eventual use is already underway. The envisaged use is intended to be broad-based access, including access by non-experts/non-specialists and communities, which will place heavy demand on provision of appropriate tools for data analysis, interpretation, and insight within GenV's Solutions Hub. Therefore, significant funding tranches will be necessary in future; this is a major concern for GenV. MADIP use consumes significant ABS resources, since ABS acts as the intermediary/broker between researchers and data custodians. Up to 50% of ABS resourcing for MADIP is spent on governance and approval requirements (including essential 5 SAFES training), with a significant cost burden also spent on building the technical platforms and tools that allow researchers to access MADIP data in a secure environment.

APPENDIX 2:

RESEARCH OUTPUTS

Some of the notable outputs that have been generated from this 12-month project are:

Research report:

This report is a public facing research report, summarising the outcomes of our research that was done in collaboration with data decision makers.

Presentations:

- Chris Mesiku and Maia Gould presented a data and disadvantage talk to invited guests as part of the 2022 launch of the ANU School of cybernetics.

Discussions and workshops:

- The data and disadvantage team convened a data and disadvantage expert panel to discuss preliminary findings of the research in June 2023.
- Chris Mesiku represented on the project during a Humanitech Summit 2023 Panel guest appearance – Hacking the future of service and humanitarian service delivery.
- Chris Mesiku spoke at a panel discussion as part of the 2022 Spark Festival at PRF Yirranma Place –imagining better futures with cybernetics.
- Chris Mesiku and Adrian Schmidt facilitated the Designing Data Solutions for Good workshop as part of the 2022 Spark Festival at PRF Yirranma Place.

Published articles:

- Project Position paper: Do more data equal more truth? Towards a cybernetic approach to data.
- What can cybernetics tell us about the Optus and Medibank data hacks?
- Data Science Central: It takes a village to protect and steer data flow
- Enhancing Human Flourishing: The Synergy between Data and a Systems Approach

Community of Practice:

- As a result of this project, we have developed relationships with experts and stakeholders working within data and disadvantage. We are now working with this group of experts to grow a community of people interested in cybernetics, data, and disadvantage. We hope this community of practice will support possible future phases of the research through testing and iterating our proposed intervention in a pilot setting and evaluating the effectiveness of that mechanism.

REFERENCES

ANU. “Paul Ramsay Foundation Supports ANU to End Disadvantage.” The Australian National University, April 29, 2022. https://www.anu.edu.au/giving/impact-stories/paul-ramsay-foundation-supports-anu-to-end-disadvantage.

Ashby, William Ross. *An introduction to cybernetics*, 1961.

Australian Government. “Australian Data Strategy: The Australian Government’s Whole-of-Economy Vision for Data,” October 1, 2022. https://www.finance.gov.au/sites/default/files/2022-10/australian-data-strategy.pdf.

Australian Institute of Health and Welfare. “Australia’s Welfare Indicators - Australia’s Welfare Indicators,” 2022. https://www.aihw.gov.au/reports-data/indicators/australias-welfare-indicators.

Beer, Stafford. “Brain of the Firm; a Development in Management Cybernetics.” New York: Herder and Herder, 1972.

Bell, Genevieve, Maia Gould, Brenda Martin, Amy McLennan, and Ellen O’Brien. “Do More Data Equal More Truth? Toward a Cybernetic Approach to Data.” *Australian Journal of Social Issues* 56, no. 2 (2021): 213–22.

Bennett Moses, Lyria, Kylie Valentine, and Janet Chan. “Data Practices in a Web of Values: Reflections on the Gap between Ethical Principles and Data-Driven Social Policy.” SSRN Scholarly Paper. Rochester, NY, October 1, 2021. https://papers.ssrn.com/abstract=4054085.

Birhane, Abeba. “Algorithmic Injustice: A Relational Ethics Approach.” *Patterns* 2, no. 2 (February 12, 2021): 100205. https://doi.org/10.1016/j.patter.2021.100205.

Broad, Ellen. “‘We Need to Let Go of Regulating ‘Artificial Intelligence.’” *InnovationAus.Com* (blog), August 8, 2022. https://www.innovationaus.com/we-need-to-let-go-of-regulating-artificial-intelligence/.

Brotherhood of St. Laurence. “Our Social Exclusion Measure,” 2020. https://www.bsl.org.au/research/our-research-and-policy-work/social-exclusion-monitor/social-exclusion-measure/.

Bruce, Peter C. *Responsible Data Science.* Newark: John Wiley & Sons, Incorporated, 2021.

CEDA. “Disrupting Disadvantage: Setting the Scene.” Trove, 2019. https://nla.gov.au/nla.obj-2056065312.

CSIRO. “Our Future World.” CSIRO. Accessed October 4, 2023. https://www.csiro.au/en/research/technology-space/data/our-future-world.

data.org. “Workforce Wanted: Data Talent for Social Impact.” *Data.Org* (blog). Accessed November 22, 2022. https://data.org/reports/workforce-wanted/.

Davis, Kord. *Ethics of Big Data*. 1st edition. Sebastopol, California: O’Reilly, 2012.

Doan, AnHai. *Principles of Data Integration*. 1st edition. Waltham, Mass: Morgan Kaufmann, 2012.

Finlay, Linda. “Championing ‘Reflexivities.’” *Qualitative Psychology (Washington, D.C.)* 4, no. 2 (2017): 120–25. https://doi.org/10.1037/qup0000075.

Foerster, H. von. “On Self-Organizing Systems and Their Environments.” In *Understanding Understanding: Essays on Cybernetics and Cognition*, edited by Heinz von Foerster, 1–19. New York, NY: Springer, 2003. https://doi.org/10.1007/0-387-21722-3_1.

Foerster, Heinz von. “Cybernetics of Cybernetics.” In *Understanding Understanding: Essays on Cybernetics and Cognition*, edited by Heinz von Foerster, 283–86. New York, NY: Springer, 2003. https://doi.org/10.1007/0-387-21722-3_13.

Foerster, Heinz Von. “Notes On An Epistemology For Living Things.” *BCL Report #9.3*, 1972. https://semiorganized.com/resources/other/foerster/epistemology.pdf.

Forrester, Jay W. *Principles of Systems*. Productivity Press, 1990.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. “Datasheets for Datasets.” arXiv, December 1, 2021. http://arxiv.org/abs/1803.09010.

Gruen, David. “Realising the Potential of Data in Government.” Australian Bureau of Statistics, 2022. https://www.abs.gov.au/about/our-organisation/australian-statistician/speeches/realising-potential-data-government.

Hayles, N. Katherine. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago, IL: University of Chicago Press, 1999. https://press.uchicago.edu/ucp/books/book/chicago/H/bo3769963.html.

Horton, John. “Centring Reflexivity, Positionality and Autoethnographic Practices in Creative Research.” In *Creative Methods for Human Geographers*, edited by Nadia von Benzon, Mark Holton, Catherine Wilkinson, and Samantha Wilkinson. SAGE Publications Ltd, 2021.

Hu, Yanhong Jessika, Melissa Wake, Susan Clifford, and Jo Said. “1059Innovative Epidemiological Methods in a Whole-of-State Cohort of Children and Parents: Generation Victoria (GenV).” *International Journal of Epidemiology* 50, no. Supplement_1 (September 1, 2021): dyab168.290. https://doi.org/10.1093/ije/dyab168.290.

Infoxchange. “Introducing the Data Catalyst Network,” April 18, 2023. https://www.infoxchange.org/au/news/2023/04/introducing-data-catalyst-network.

“Introducing the Data and Digital Government Strategy | DDGS.” Accessed October 4, 2023. https://www.dataanddigital.gov.au/.

Jamieson, Michelle K., Gisela H. Govaart, and Madeleine Pownall. “Reflexivity in Quantitative Research: A Rationale and Beginner’s Guide.” *Social and Personality Psychology Compass* 17, no. 4 (2023): e12735. https://doi.org/10.1111/spc3.12735.

Johnson, Tess, Konrad Kollnig, and Pierre Dewitte. “Towards Responsible, Lawful and Ethical Data Processing: Patient Data in the UK.” *Internet Policy Review* 11, no. 1 (March 18, 2022). https://policyreview.info/articles/analysis/towards-responsible-lawful-and-ethical-data-processing-patient-data-uk.

Kentico. “CEDA - Addressing Entrenched Disadvantage in Australia,” 2015. https://www.ceda.com.au/ResearchAndPolicies/Research/Population/Addressing-entrenched-disadvantage-in-Australia.

Kitchin, Rob, and Martin Dodge. *Code/Space: Software and Everyday Life*. The MIT Press, 2011. https://doi.org/10.7551/mitpress/9780262042482.001.0001.

Krippendorff, Klaus. “Cybernetics’s Reflexive Turns,” January 1, 2008, 14.

Luke, Joanne, Ebony Verbunt, Angela Zhang, Muriel Bamblett, Gabrielle Johnson, Connie Salamone, David Thomas, et al. “Questioning the Ethics of Evidence-Based Practice for Indigenous Health and Social Settings in Australia.” *BMJ Global Health* 7, no. 6 (June 1, 2022): e009167. https://doi.org/10.1136/bmjgh-2022-009167.

Lynch, Melody. “Data Lives: How Data Are Made and Shape Our World.” *The AAG Review of Books* 10, no. 2 (April 3, 2022): 20–21. https://doi.org/10.1080/2325548X.2022.2036546.

MacKay, Donald MacCrimmon. *Information, Mechanism and Meaning*. Cambridge, Massachusetts: MIT Press, 1969.

McCallum, Kerry, Tess Ryan, and Jo Caffery. “Deficit Metrics in Australian Indigenous Education: Through a Media Studies Lens.” *Discourse: Studies in the Cultural Politics of Education* 43, no. 2 (March 4, 2022): 266–81. https://doi.org/10.1080/01596306.2020.1828285.

McDonald, Catherine, and Lesley Chenoweth. "(Re) Shaping Social Work: An Australian Case Study." *The British Journal of Social Work* 39, no. 1 (2009): 144–60.

McKay, S. "Poverty or Preference: What Do 'Consensual Deprivation Indicators' Really Measure?" *Fiscal Studies* 25, no. 2 (2004): 201–23. https://doi.org/10.1111/j.1475-5890.2004.tb00102.x.

OAIC. "What Is Privacy?" OAIC, March 10, 2023. https://www.oaic.gov.au/privacy/your-privacy-rights/your-personal-information/what-is-privacy.

Perry, Laura. "Educational Disadvantage Is a Huge Problem in Australia – We Can't Just Carry on the Same." The Conversation, 2017. http://theconversation.com/educational-disadvantage-is-a-huge-problem-in-australia-we-cant-just-carry-on-the-same-74530.

Phillips, Shawn, Ami Seivwright, Susan Young, and Et Al. "Insights into Hardship and Disadvantage in Perth, Western Australia: The 100 Families WA Report." 100 Families WA, 2021. https://doi.org/10.25916/B914-1J34.

Provost, Foster, and Tom Fawcett. *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. 1st, 1st ed. ed. Sebastopol: O'Reilly Media, Incorporated, 2013.

Ryan, Louise, and Anne Golden. "'Tick the Box Please': A Reflexive Approach to Doing Quantitative Social Research." *Sociology* 40, no. 6 (December 2006): 1191–1200. https://doi.org/10.1177/0038038506072287.

Scutella, Rosanna, Michael Horn, and Roger Wilkins. "Measuring Poverty and Social Exclusion in Australia: A Proposed Multidimensional Framework for Identifying Socio-Economic Disadvantage." *SSRN Electronic Journal*, 2009. https://doi.org/10.2139/ssrn.1639867.

Shannon, C. E. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27, no. 4 (1948): 623–56. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.

Staines, Zoe, Charlotte Moore, Greg Marston, and Louise Humpage. "Big Data and Poverty Governance under Australia and Aotearoa/New Zealand's 'Social Investment' Policies." *Australian Journal of Social Issues* 56, no. 2 (2021): 157–72. https://doi.org/10.1002/ajs4.129.

"Supporting Analysis of The Life Course | Australian Bureau of Statistics," August 16, 2023. https://www.abs.gov.au/about/our-organisation/australian-statistician/speeches/supporting-analysis-life-course.

Tigard, Daniel W. "Big Data and the Threat to Moral Responsibility in Healthcare." In *Datenreiche Medizin und das Problem der Einwilligung: Ethische, rechtliche und sozialwissenschaftliche Perspektiven*, edited by Gesine Richter, Wulf Loh, Alena Buyx, and Sebastian Graf von Kielmansegg, 11–25. Berlin, Heidelberg: Springer, 2022. https://doi.org/10.1007/978-3-662-62987-1_2.

"To Create Fairness and Accountability in the Use of Government Decision Making Algorithms - Churchill Trust." Accessed August 29, 2022. https://www.churchilltrust.com.au/fellow/dr-owen-churches-sa-2018.

VanderWeele, Tyler J. "On the Promotion of Human Flourishing." *Proceedings of the National Academy of Sciences* 114, no. 31 (August 2017): 8148–56. https://doi.org/10.1073/pnas.1702996114.

Veenstra, Anne Fleur van, and Bas Kotterink. "Data-Driven Policy Making: The Policy Lab Approach." In *Electronic Participation*, edited by Peter Parycek, Yannis Charalabidis, Andrei V. Chugunov, Panos Panagiotopoulos, Theresa A. Pardo, Øystein Sæbø, and Efthimios Tambouris, 100–111. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-64322-9_9.

Wacquant, Loïc. *Punishing the Poor: The Neoliberal Government of Social Insecurity*. duke university Press, 2009.

Wake, Melissa, Sharon Goldfeld, and Andrew Davidson. “Embedding Life Course Interventions in Longitudinal Cohort Studies: Australia’s GenV Opportunity.” *Pediatrics* 149, no. Suppl 5 (May 2022): e2021053509R. https://doi.org/10.1542/peds.2021-053509R.

Wake, Melissa, Yanhong Jessika Hu, Hayley Warren, Margie Danchin, Michael Fahey, Francesca Orsini, Maurizio Pacilli, Kirsten P. Perrett, Richard Saffery, and Andrew Davidson. “Integrating Trials into a Whole-Population Cohort of Children and Parents: Statement of Intent (Trials) for the Generation Victoria (GenV) Cohort.” *BMC Medical Research Methodology* 20 (September 24, 2020): 238. https://doi.org/10.1186/s12874-020-01111-x.

Whitehair, K. J. “Hasselbalch, Gry. Data Ethics of Power: A Human Approach in the Big Data and Al Era.” *CHOICE: Current Reviews for Academic Libraries* 60, no. 2 (2022): 182-.

Wiener, Norbert. *Cybernetics or Control and Communication in the Animal and the Machine*. Technology Press, 1948.

Wolff, Jonathan, and Avner De-Shalit. *Disadvantage*. Oxford: Oxford University Press, 2007.

Zook, Matthew, Solon Barocas, Danah Boyd, Kate Crawford, Emily Keller, Seeta Peña Gangadharan, Alyssa Goodman, et al. “Ten Simple Rules for Responsible Big Data Research.” *PLOS Computational Biology* 13, no. 3 (March 30, 2017): e1005399. https://doi.org/10.1371/journal.pcbi.1005399.

Zuber, Niina, Jan Gogoll, Severin Kacianka, Alexander Pretschner, and Julian Nida-Rümelin. “Empowered and Embedded: Ethics and Agile Processes.” *Humanities and Social Sciences Communications* 9, no. 1 (June 6, 2022): 1–13. https://doi.org/10.1057/s41599-022-01206-4.

BIBLIOGRAPHY

"1950 | Timeline of Computer History | Computer History Museum." Accessed September 6, 2022. https://www.computerhistory.org/timeline/1950/.

"A Global History of Cybernetics," March 25, 2016. https://doi.org/10.7551/mitpress/9800.003.0004.

"A Snapshot of Poverty in Australia 2022 – Poverty and Inequality." Accessed November 21, 2022. https://povertyandinequality.acoss.org.au/a-snapshot-of-poverty-in-australia-2022/.

Abad, Zahra Shakeri Hossein, Wendy Thompson, Joon Lee. "Physical Activity, Sedentary Behavior, and Sleep on Twitter: Multicountry and Fully Labeled Public Data Set for Digital Public Health Surveillance Research." JMIR Public Health and Surveillance 8, no. 2 (February 2022): e32355. https://doi.org/10.2196/32355.

ABC News. "The Australian behind Technology That Could 'end Privacy as We Know It' Says It Is Being Used Here." January 22, 2020. https://www.abc.net.au/news/2020-01-23/australian-founder-of-clearview-facial-recognition-interview/11887112.

ABC Radio National. "Alleviating Child Poverty," October 20, 2022. https://www.abc.net.au/radionational/programs/bigideas/alleviating-child-poverty-/14087778.

Abe, A.K. "Developing Deprivation Index for Children Taking into Account of Adaptive Preferences." Child Indicators Research 12, no. 2 (2019): 647–65. https://doi.org/10.1007/s12187-018-9553-4.

Abel-Smith, Brian, and Peter Townsend. The Poor and the Poorest: A New Analysis of the Ministry of Labour's Family Expenditure Surveys of 1953-54 and 1960. Bell, 1966.

Abowd, John, Simon Woodcock, Benoit Dostie, Sam Hawala, Janet Heslop, Paul Massell, Carol Murphree, et al. "Disclosure Limitation in Longitudinal Linked Data." Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies, January 1, 2001.

Academy of the Social Sciences in Australia. "Big Data for Australian Social Policy," December 16, 2021. https://socialsciences.org.au/publications/big-data-for-australian-social-policy/.

Adams, Emma K., Kirsten J. Hancock, Francis Mitrou, Daniel Christensen, Catherine L. Taylor, and Stephen R. Zubrick. "Transfers of Disadvantage across Three Generations Using Latent Class Associations within Families." Australian Journal of Social Issues 57, no. 3 (2022): 497–523. https://doi.org/10.1002/ajs4.206.

Alexander, J.T., Davern, M., Stevenson, B.: Inaccurate Age and Sex Data in the [United States] Census PUMS Files: Evidence and Implications. Public Opinion Quarterly 74(3), 551–569 (2010), Doi:10.1093/Poq/Nfq033.

Allam, Zaheer, Gourav Dey, and David S. Jones. "Artificial Intelligence (AI) Provided Early Detection of the Coronavirus (COVID-19) in China and Will Influence Future Urban Health Policy Internationally." AI 1, no. 2 (June 2020): 156–65. https://doi.org/10.3390/ai1020009.

Amara, Nabil, Mathieu Ouimet, and RÉjean Landry. "New Evidence on Instrumental, Conceptual, and Symbolic Utilization of University Research in Government Agencies." Science Communication 26, no. 1 (2004): 75–106. https://doi.org/10.1177/1075547004267491.

Ambrey, Christopher L, Cameron Parsell, Melanie Spallek, and Richard N S Robinson. "An Investigation Into Repeat Requests for Charity: Evidence From the St Vincent de Paul Society Queensland, Australia." Nonprofit and Voluntary Sector Quarterly 48, no. 1 (2019): 91–107. https://doi.org/10.1177/0899764018794300.

ANU. “Paul Ramsay Foundation Supports ANU to End Disadvantage.” The Australian National University, April 29, 2022. https://www.anu.edu.au/giving/impact-stories/paul-ramsay-foundation-supports-anu-to-end-disadvantage.

ARDC Leadership Series: Sharing Sensitive and Identifiable Human Data, 2022. https://www.youtube.com/watch?v=X7bN_JumOBE.

Ashby, William Ross. An introduction to cybernetics, 1961

Atkinson, A.B. “On the Measurement of Poverty.” Econometrica 55, no. 4 (1987): 749–64.

“Australia : Funding the Australian Bureau of Statistics to Better Collect Data on Disadvantage.” MENA Report. London, 2022. https://www.proquest.com/docview/2711767530/fulltext/17487D5DD6404FC8PQ/1?accountid=8330&forcedol=true.

Australian Bureau of Statistics. “Business Longitudinal Analysis Data Environment (BLADE).” Accessed July 24, 2023. https://www.abs.gov.au/about/data-services/data-integration/integrated-data/business-longitudinal-analysis-data-environment-blade.

Australian Bureau of Statistics. “Data Integration Service.” Accessed July 24, 2023. https://www.abs.gov.au/about/data-services/data-integration/data-integration-service.

Australian Bureau of Statistics. “MADIP Data and Legislation.” Accessed July 24, 2023. https://www.abs.gov.au/about/data-services/data-integration/integrated-data/multi-agency-data-integration-project-madip/madip-data-and-legislation.

Australian Bureau of Statistics. “MADIP Privacy Impact Assessment.” Accessed July 24, 2023. https://www.abs.gov.au/about/data-services/data-integration/integrated-data/multi-agency-data-integration-project-madip/madip-privacy-impact-assessment.

Australian Bureau of Statistics. “Multi-Agency Data Integration Project (MADIP).” Accessed July 24, 2023. https://www.abs.gov.au/about/data-services/data-integration/integrated-data/multi-agency-data-integration-project-madip.

Australian Bureau of Statistics. “Undertakings and Declarations,” 2021. https://www.abs.gov.au/statistics/microdata-tablebuilder/responsible-use-abs-microdata/undertakings-and-declarations.

“Australian Census Longitudinal Dataset - Voluntary Work - Dataset - Data.Gov.Au.” Accessed July 18, 2022. https://data.gov.au/data/dataset/australian-census-longitudinal-dataset-voluntary-work.

Australian Government. “Australian Charities Report - 8th Edition.” Australian Charities and Not-for-Profits Commission, 2022. https://www.acnc.gov.au/tools/reports/australian-charities-report-8th-edition.

“National Centre for Place-Based Collaboration (Nexus Centre).” Department of Social Services, Australian Government. Accessed May 31, 2023. https://www.dss.gov.au/place-based-collaboration.

“Stronger Places, Stronger People.” Department of Social Services, Australian Government. Accessed May 31, 2023. https://www.dss.gov.au/families-and-children-programs-services/stronger-places-stronger-people.

Australian Human Rights Commission. “Aboriginal and Torres Strait Islander Social Justice.” Accessed July 24, 2023. https://humanrights.gov.au/our-work/aboriginal-and-torres-strait-islander-social-justice.

Australian Institute of Health and Welfare. “Australia’s Welfare Indicators - Australia’s Welfare Indicators.” Accessed October 6, 2022. https://www.aihw.gov.au/reports-data/indicators/australias-welfare-indicators.

Australian Institute of Health and Welfare. “New Report Shows Long-Term Disadvantage for Australia’s Stolen Generations.” Accessed October 6, 2022. https://www.aihw.gov.au/news-media/media-releases/2018/august/new-report-shows-long-term-disadvantage-for-austra.

Australian Institute of Health and Welfare. “Social Isolation and Loneliness.” Accessed September 21, 2022. https://www.aihw.gov.au/reports/australias-welfare/social-isolation-and-loneliness-covid-pandemic.

Australian Institute of Health and Welfare. “The Five Safes Framework.” Accessed October 11, 2022. https://www.aihw.gov.au/about-our-data/data-governance/the-five-safes-framework.

Azpitarte, F., and G. Kalb. “Measuring Income Poverty in Australia: A Review of Methods and Recent Trends.” Revisiting Henderson: Poverty, Social Security and Basic Income, 2019, 163–87.

“Measuring Income Poverty in Australia: A Review of Methods and Recent Trends.” Revisiting Henderson: Poverty, Social Security and Basic Income, 2019, 163–87.

Azpitarte, Francisco. “Was Pro-Poor Economic Growth in Australia for the Income-Poor? And for the Multidimensionally- Poor?” Social Indicators Research 117, no. 3 (2014): 871–905. https://doi.org/10.1007/s11205-013-0378-8.

Bäckman, Guy. “Reframing Social Policy for Welfare and Equal Opportunities in a Digital Culture.” Social Development Issues 43, no. 2 (2021). https://doi.org/10.3998/sdi.1821.

Bagot, John. “Progress and Poverty.” The Westminster Review 172 (1909): 371–75.

Bailey, Kenneth D. “Sociocybernetics and Social Entropy Theory.” Kybernetes 35, no. 3/4 (2006): 375-384,279. https://doi.org/10.1108/03684920610653683.

Banks, Emily. “How Do We Decide Which Disease to Prevent next? Long-Term Studies Help.” The Conversation. Accessed July 13, 2022. http://theconversation.com/how-do-we-decide-which-disease-to-prevent-next-long-term-studies-help-73833.

Beer, Stafford. “Brain of the Firm; a Development in Management Cybernetics.” New York: Herder and Herder, 1972.

Bennett Moses, Lyria, Kylie Valentine, and Janet Chan. “Data Practices in a Web of Values: Reflections on the Gap between Ethical Principles and Data-Driven Social Policy.” SSRN Scholarly Paper. Rochester, NY, 2021. https://papers.ssrn.com/abstract=4054085.

Berthoud, R., M. Bryan, R. Berthoud, and F. Zantomio. “‘Deprivation Indicators.’” Measuring Poverty: Seven Key Issues, 2008, 14–15.

Best, Stephen. “Walter J. Ong, Orality and Literacy (1982).” Public Culture 32, no. 2 (91) (May 1, 2020): 431–39. https://doi.org/10.1215/08992363-8090173.

“Big Data for Australian Social Policy | Academy of the Social Sciences in Australia.” Accessed February 28, 2023. https://socialsciences.org.au/publications/big-data-for-australian-social-policy/.

Blueblood. “Mapping the Potential.” Catholic Social Services Australia. Accessed November 29, 2022. https://cssa.org.au/mapping-the-potential/.

Bonnery, Daniel, Yi Feng, Angela Henneberger, Tessa Johnson, Mark Lachowicz, Bess Rose, Terry Shaw, Laura Stapleton, Michael Woolley, and Yating Zheng. “The Promise and Limitations of Synthetic Data as a Strategy to Expand Access to State-Level Multi-Agency Longitudinal Data.” Journal of Research on Educational Effectiveness 12 (August 2, 2019): 1–29. https://doi.org/10.1080/19345747.2019.1631421.

Brady, Marnie, Kathleen Dunn, and Jamie McCallum. “Commentary: Culture of Poverty: Don’t Call It a Comeback!” Formations: The Graduate Center Journal of Social Research 1, no. 1 (2010): 103–6.

Braun, Virginia. “Thematic Analysis: A Practical Guide.” Los Angeles: SAGE, 2022.

Bronfenbrenner, Martin. Review of Review of Ends and Means of Reducing Income Poverty, by Robert J. Lampman. The Journal of Human Resources 9, no. 2 (1974): 290–93. https://doi.org/10.2307/144978.

Brotherhood of St. Laurence. “Our Social Exclusion Measure.” Accessed September 16, 2022. https://www.bsl.org.au/research/our-research-and-policy-work/social-exclusion-monitor/social-exclusion-measure/.

“Business Longitudinal Analysis Data Environment (BLADE) | Australian Bureau of Statistics,” July 16, 2021. https://www.abs.gov.au/about/data-services/data-integration/integrated-data/business-longitudinal-analysis-data-environment-blade.

Butler, Danielle C., Ellie Paige, Jennifer Welsh, Hsei Di Law, Lynelle Moon, Emily Banks, and Rosemary J. Korda. “Factors Related to Under-Treatment of Secondary Cardiovascular Risk, Including Primary Healthcare: Australian National Health Survey Linked Data Analysis.” Australian and New Zealand Journal of Public Health 46, no. 4 (2022): 533–39. https://doi.org/10.1111/1753-6405.13254.

Bynum, Terrell Ward. “Flourishing Ethics.” Ethics and Information Technology 8, no. 4 (November 1, 2006): 157–73. https://doi.org/10.1007/s10676-006-9107-1.

Cabinet, Prime Minister and. “Data Integration Partnership for Australia.” Text, August 29, 2017. https://www.pmc.gov.au/public-data/data-integration-partnership-australia.

Cantillon, S. “Measuring Differences in Living Standards Within Households.” Journal of Marriage and Family 75, no. 3 (2013): 598–610. https://doi.org/10.1111/jomf.12023.

CEDA. “Disrupting Disadvantage: Setting the Scene,” 2019. https://cedakenticomedia.blob.core.windows.net/cedamediacontainer/kentico/media/general/publication/pdfs/ceda_disruptingdisadvantagess1nov2019_com.pdf.

“CEDA - Disrupting Disadvantage 3: Finding What Works,” 2023. https://www.ceda.com.au/ResearchAndPolicies/Research/Government-Regulation/DD3.

Chan, Janet, and Peter Saunders. “Big Data for Australian Social Policy: Developments, Benefits and Risks.” Canberra: Academy of the Social Sciences in Australia, 2021. https://socialsciences.org.au/wp-content/uploads/2021/12/Big-Data-for-Australian-Social-Policy.pdf.

Clemetson, Lynette. “Homeland Security Given Data on Arab-Americans.” New York Times, Late Edition (East Coast), July 30, 2004, sec. A.

Cleveland, Lara, Robert McCaa, Steven Ruggles, and Matthew Sobek. “When Excessive Perturbation Goes Wrong and Why IPUMS-International Relies Instead on Sampling, Suppression, Swapping, and Other Minimally Harmful Methods to Protect Privacy of Census Microdata.” In Privacy in Statistical Databases, edited by Josep Domingo-Ferrer and Ilenia Tinnirello, 179–87. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2012. https://doi.org/10.1007/978-3-642-33627-0_14.

Clifford, Susan, and Melissa Wake. “GenV Measures Selection Principles Working Paper.” Murdoch Children’s Research Institute. Murdoch Childrens Research Institute, 2019. https://doi.org/10.25374/MCRI.9750983.v3.

Clough, Patricia Ticineto, Karen Gregory, Benjamin Haber, and R. Joshua Scannell. “The Datalogical Turn.” In Non-Representational Methodologies. Routledge, 2015.

Collins, Samuel Gerald. “Cutting Loose in the Fun Palace: Reengaging Cybernetics in a Complex World.” Social Studies of Science 43, no. 2 (2013): 314–18. https://doi.org/10.1177/0306312712461712.

Coyle, Diane, Paul Allin, and Tim Jackson. “Beyond GDP: Changing How We Measure Progress Is Key to Tackling a World in Crisis – Three Leading Experts.” The Conversation. Accessed October 11, 2022. http://theconversation.com/beyond-gdp-changing-how-we-measure-progress-is-key-to-tackling-a-world-in-crisis-three-leading-experts-186488.

Crawford, Kate. “The Hidden Biases in Big Data.” Harvard Business Review 1, no. 4 (2013).

Crettaz, E., and C. Suter. “The Impact of Adaptive Preferences on Subjective Indicators: An Analysis of Poverty Indicators.” Social Indicators Research 114, no. 1 (2013): 139–52. https://doi.org/10.1007/s11205-013-0388-6.

Dahl, Cortland J, Christine D Wilson-Mendenhall, and Richard J Davidson. “The Plasticity of Well-Being: A Training-Based Framework for the Cultivation of Human Flourishing.” Proceedings of the National Academy of Sciences 117, no. 51 (2020): 32197–206. https://doi.org/10.1073/pnas.2014859117.

Dang, Hai-Anh, Dean Jolliffe, and Calogero Carletto. “Data Gaps, Data Incomparability, and Data Imputation: A Review of Poverty Measurement Methods for Data-Scarce Environments.” Journal of Economic Surveys 33, no. 3 (2019): 757–97. https://doi.org/10.1111/joes.12307.

“Data Confidentiality Guide | Australian Bureau of Statistics.” Accessed September 16, 2022. https://www.abs.gov.au/book/export/28232/print.

“Data Driven Policy – Where to from Here?” Accessed September 16, 2022. https://www.themandarin.com.au/163988-data-driven-policy-where-to-from-here/.

data innovation alliance. “Data Ethics.” Accessed November 8, 2022. https://data-innovation.org/data-ethics/.

data.org. “Workforce Wanted: Data Talent for Social Impact.” Data.Org (blog). Accessed November 22, 2022. https://data.org/reports/workforce-wanted/.

Datt, G. “Distribution-Sensitive Multidimensional Poverty Measures.” World Bank Economic Review 33, no. 3 (2019): 551–72. https://doi.org/10.1093/wber/lhx017.

Davidson, P., P. Saunders, B. Bradbury, and M. Wong. “Poverty in Australia 2020: Part 1, Overview.” Poverty in Australia 2020: Part 1, Overview, 2020.

Davies, William. “How Statistics Lost Their Power – and Why We Should Fear What Comes Next.” The Guardian, January 19, 2017, sec. Politics. https://www.theguardian.com/politics/2017/jan/19/crisis-of-statistics-big-data-democracy.

Dean, Hartley, and Lucinda Platt, eds. Social Advantage and Disadvantage. First edition. Oxford, United Kingdom: Oxford University Press, 2016.

Doan, AnHai. Principles of Data Integration. 1st edition. Waltham, Mass: Morgan Kaufmann, 2012.

Dreier, Thomas. “Big Data Book Review.” Journal of Intellectual Property, Information Technology and Electronic Commerce Law 5, no. 1 (2014): 60–63.

Duncan, G. T., M. Elliot, and J. -. J. Salazar-González. Statistical Confidentiality: Principles and Practice. Heidelberg: Springer, 2011. https://doi.org/10.1007/978-1-4419-7802-8.

“The Role of Data in Disrupting Disadvantage - Taylor Fry,” December 9, 2021. https://taylorfry.com.au/articles/the-role-of-data-in-disrupting-disadvantage/.

Edwards, Meredith. “Social Science Research and Public Policy: Narrowing the Divide.” Australian Journal of Public Administration 64, no. 1 (2005): 68–74. https://doi.org/10.1111/j.1467-8500.2005.00417.x.

Ellyatt, Wendy. “Education for Human Flourishing — A New Conceptual Framework for Promoting Ecosystemic Wellbeing in Schools.” Challenges 13, no. 2 (2022): 58. https://doi.org/10.3390/challe13020058.

Espinosa Salazar, Angela Ma, and Allenna Leonard. Cybernetics and Public Administration. Kybernetes the International Journal of Systems & Cybernetics ; v. 35, No. 1/2. Bradford, England: Emerald Group Publishing, 2006.

Eubanks, Virginia. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin’s Press, 2018.

Evaluation, Centre for Education Statistics and. “Pathways for the Future Program.” NSW Department of Education, August 5, 2022. https://education.nsw.gov.au/about-us/educational-data/cese/publications/pathways-for-the-future-program.html.

Ferragina, Emanuele, Mark Tomlinson, and Robert Walker. “Poverty, Participation and Choice: The Legacy of Peter Townsend,” 2013. https://www.jrf.org.uk/sites/default/files/jrf/migrated/files/society-poverty-participation-full.pdf.

Fincher, Ruth, and Peter Saunders. Creating Unequal Futures?: Rethinking Poverty, Inequality and Disadvantage. Sydney, AUSTRALIA: Taylor & Francis Group, 2001. http://ebookcentral.proquest.com/lib/anu/detail.action?docID=314104.

Fiske, Amelia, Barbara Prainsack, and Alena Buyx. “Data Work: Meaning-Making in the Era of Data-Rich Medicine.” Journal of Medical Internet Research 21, no. 7 (July 9, 2019): e11672. https://doi.org/10.2196/11672.

Flood, Robert L, and Michael C Jackson. Creative Problem Solving: Total Systems Intervention. Chichester; New York: Wiley, 1991.

Foerster, H. von. “On Self-Organizing Systems and Their Environments.” In Understanding Understanding: Essays on Cybernetics and Cognition, edited by Heinz von Foerster, 1–19. New York, NY: Springer, 2003. https://doi.org/10.1007/0-387-21722-3_1.

Foerster, Heinz von. “Cybernetics of Cybernetics.” In Understanding Understanding: Essays on Cybernetics and Cognition, edited by Heinz von Foerster, 283–86. New York, NY: Springer, 2003. https://doi.org/10.1007/0-387-21722-3_13.

Geary, Daniel. “The Moynihan Report.” The Atlantic, September 15, 2015. https://www.theatlantic.com/politics/archive/2015/09/the-moynihan-report-an-annotated-edition/404632/.

Generation Victoria. “Supporters and Partners.” Accessed July 24, 2023. https://genv.org.au/about-genv/supporters-and-partners/.

George, Henry. Progress and Poverty: An Inquiry Into the Cause of Industrial Depressions, and of Increase of Want with Increase of Wealth. The Remendy. D. Appleton, 1879.

Geyer, Felix. “The March of Self-Reference.” Kybernetes 31, no. 7/8 (2002): 1021–42. https://doi.org/10.1108/03684920210436318.

Giest, Sarah, and Reuben Ng. “Big Data Applications in Governance and Policy.” Politics and Governance 6, no. 4 (November 21, 2018): 1–4.

Giest, Sarah, and Annemarie Samuels. “‘For Good Measure’: Data Gaps in a Big Data World.” Policy Sciences 53, no. 3 (2020): 559–69. https://doi.org/10.1007/s11077-020-09384-1.

Gillingham, Philip, and Timothy Graham. “Big Data in Social Welfare: The Development of a Critical Perspective on Social Work’s Latest ‘Electronic Turn.’” Australian Social Work 70, no. 2 (2017): 135–47.

Gilliom, John. Overseers of the Poor: Surveillance, Resistance, and the Limits of Privacy. University of Chicago Press, 2001.

Global Indigenous Data Alliance. “CARE Principles of Indigenous Data Governance.” Accessed September 17, 2022. https://www.gida-global.org/care.

Goldenberg, Marti, Meghan McCormick, Kristine Ruppert, Enrico M Novelli, and Ram Kalpatthi. “Vaso-Occlusive Events in Pediatric Sickle Cell Disease: Quantifying Social Disadvantage and Its Impact on Hospitalizations.” Blood 138, no. Supplement 1 (November 5, 2021): 1949. https://doi.org/10.1182/blood-2021-153986.

Gordon, D. “Measuring Poverty in the UK.” In Poverty and Social Exclusion in the UK, 1:17–39, 2017.

Governance Institute of Australia. “Australia’s Charities in Numbers.” Governance Directions 70, no. 1 (2018): 4–5.

Graff, Kirby, Gabriela Quiñonez-Riegos, Jane Scott, and Heidi Senior. “Ethical Issues in Academic Libraries: An Annotated Bibliography.” Journal of Information Ethics 29, no. 1 (Spring 2020): 65–88.

Gray, Professor Matthew. “Better Data for the Delivery of Grant Funded Social Services,” Accessed September 10, 2022, https://csrm.cass.anu.edu.au/research/publications/better-data-delivery-grant-funded-social-services.

Gruen, David. “Realising the Potential of Data in Government.” Australian Bureau of Statistics, 2022. https://www.abs.gov.au/about/our-organisation/australian-statistician/speeches/realising-potential-data-government.

Guenther, John, Melodie Bat, and Sam Osborne. “Red Dirt Thinking on Remote Educational Advantage.” Australian and International Journal of Rural Education 24, no. 1 (2014): 51–67.

Gupta, Shiv Prakash. “Conceptualizing Poverty.” Sociological Bulletin 33, no. 1/2 (1984): 63–74.

Halleröd, B., J. Bradshaw, and H. Holmes. “Adapting the Consensual Definition of Poverty.” In Breadline Britain in the 1990s, 213–34, 2018. https://doi.org/10.4324/9780429460173-11.

Harrington, Michael. “The Other America; Poverty in the United States.” New York: Macmillan, 1962.

Hayes, Alan, and Andrew Hacker. “Persistent Disadvantage in Australia: Extent, Complexity and Some Key Implications.” Australia’s Welfare Series. Canberra: Australian Institute of Health and Welfare, 2017. https://www.aihw.gov.au/getmedia/9592571c-801c-46be-9c9d-75d0faffbb5b/aihw-australias-welfare-2017-chapter1-6.pdf.aspx.

Head, Brian. “Evidence-Based Policy: Principles and Requirements.” In Strengthening Evidence-Based Policy in the Australian Federation, edited by Productivity Commission, 13–26. Melbourne, Australia: Productivity Commission, 2010.

Henderson, Ronald Frank. “People in Poverty: A Melbourne Survey.” University of Melbourne. Institute of Applied Economic and Social Research. Monographs, 4. Melbourne: Cheshire for the Institute of Applied Economic and Social Research, University of Melbourne, 1970.

Hick, R. “Material Poverty and Multiple Deprivation in Britain: The Distinctiveness of Multidimensional Assessment.” Journal of Public Policy 36, no. 2 (2016): 277–308. https://doi.org/10.1017/S0143814X14000348.

Höltge, J, R G Cowden, Matthew T Lee, A O Bechara, S Joynt, S Kamble, V V Khalanskyi, et al. “A Systems Perspective on Human Flourishing: Exploring Cross-Country Similarities and Differences of a Multisystemic Flourishing Network.” The Journal of Positive Psychology 18, no. 5 (2022): 695–710. https://doi.org/10.1080/17439760.2022.2093784.

“Clearview AI Breached Australians’ Privacy.” Accessed September 15, 2022. https://www.oaic.gov.au/updates/news-and-media/clearview-ai-breached-australians-privacy.

Homel, Jacqueline, Ross Homel, Tara Renae McGee, Pauline Zardo, Sara Branch, Kate Freiberg, Matthew Manning, and Gabriel Wong. “Evaluation of a Place-Based Collective Impact Initiative through Cross-Sectoral Data Linkage.” Australian Journal of Social Issues 56, no. 2 (2021): 301–18. https://doi.org/10.1002/ajs4.147.

“How Better Use of Data Can Help Address Key Challenges Facing the NHS - The Health Foundation.” Accessed August 9, 2022. https://www.health.org.uk/publications/long-reads/how-better-use-of-data-can-help-address-key-challenges-facing-the-nhs.

“How Big Data Creates Human Flourishing - McCrindle,” February 19, 2019. https://mccrindle.com.au/article/blog/how-big-data-creates-human-flourishing/.

Hu, Yanhong Jessika, Anna Fedyukova, Jing Wang, Joanne M. Said, Niranjan Thomas, Elizabeth Noble, Jeanie L. Y. Cheong, Bill Karanatsios, Sharon Goldfeld, and Melissa Wake. “Improving Cohort-Hospital Matching Accuracy through Standardization and Validation of Participant Identifiable Information.” Children 9, no. 12 (December 7, 2022): 1916. https://doi.org/10.3390/children9121916.

Hu, Yanhong Jessika, Melissa Wake, Susan Clifford, and Jo Said. “Innovative Epidemiological Methods in a Whole-of-State Cohort of Children and Parents: Generation Victoria (GenV).” International Journal of Epidemiology 50, no. Supplement_1 (September 1, 2021): dyab168.290. https://doi.org/10.1093/ije/dyab168.290.

Hubal, Elaine A. Cohen, Nicole M. DeLuca, Ashley Mullikin, Rachel Slover, John C. Little, and David M. Reif. “Demonstrating a Systems Approach for Integrating Disparate Data Streams to Inform Decisions on Children’s Environmental Health.” BMC Public Health 22, no. 1 (February 15, 2022): 313. https://doi.org/10.1186/s12889-022-12682-3.

Hunter, Wade Robert. “Global Growth, Inequality, and Poverty: The Globalization Argument and the ‘Political’ Science of Economics.” In Global Political Economy, edited by John Ravenhill, 319–55. Oxford University Press, 2017.

Infoxchange. “Introducing the Data Catalyst Network,” April 18, 2023. https://www.infoxchange.org/au/news/2023/04/introducing-data-catalyst-network.

IPAA 2021 Garran Oration with Distinguished Professor Genevieve Bell AO, 2022. https://vimeo.com/692091571.

Johns, Gary. “Charities Reform in Australia.” Agenda: A Journal of Policy Analysis and Reform 11, no. 4 (2004): 293–306.

Johnson, David. “Poverty Lines and the Measurement of Poverty.” The Australian Economic Review 29, no. 1 (January 1996): 110–26. https://doi.org/10.1111/j.1467-8462.1996.tb00919.x.

Johnson, Tess, Konrad Kollnig, and Pierre Dewitte. “Towards Responsible, Lawful and Ethical Data Processing: Patient Data in the UK.” Internet Policy Review 11, no. 1 (March 18, 2022). https://policyreview.info/articles/analysis/towards-responsible-lawful-and-ethical-data-processing-patient-data-uk.

Karunananthan, S., and C. Wolfson. “MISSING DATA IN LONGITUDINAL STUDIES OF AGING: THE GOOD, THE BAD, AND THE UGLY.” Innovation in Aging 1, no. Suppl 1 (June 30, 2017): 965. https://doi.org/10.1093/geroni/igx004.3481.

Kenny, Christopher T., Shiro Kuriwaki, Cory McCartan, Evan T. R. Rosenman, Tyler Simko, and Kosuke Imai. “The Use of Differential Privacy for Census Data and Its Impact on Redistricting: The Case of the 2020 U.S. Census.” Science Advances 7, no. 41 (n.d.): eabk3283. https://doi.org/10.1126/sciadv.abk3283.

Kentico. “CEDA - Addressing Entrenched Disadvantage in Australia,” 2015. https://www.ceda.com.au/ResearchAndPolicies/Research/Population/Addressing-entrenched-disadvantage-in-Australia.

Kentico. “Disrupting Disadvantage Part 2,” 2021. https://www.ceda.com.au/ResearchAndPolicies/Research/Economy/Disrupting-Disadvantage-Part-2.

Klein, E. “The Curious Case of Using the Capability Approach in Australian Indigenous Policy.” Journal of Human Development and Capabilities 17, no. 2 (2016): 245–59. https://doi.org/10.1080/19452829.2016.1145199.

Kockelman, Paul. “Information Is the Enclosure of Meaning: Cybernetics, Semiotics, and Alternative Theories of Information.” Language & Communication 33, no. 2 (April 1, 2013): 115–27. https://doi.org/10.1016/j.langcom.2013.01.002.

Koziol, Michael. “No Child Will Live in Poverty? 30 Years on, Bob Hawke’s Promise Remains an Elusive Goal.” The Sydney Morning Herald, June 22, 2017. https://www.smh.com.au/politics/federal/no-child-will-live-in-poverty-30-years-on-bob-hawkes-promise-remains-an-elusive-goal-20170621-gwvdya.html.

Kripke, Gawain. “‘Poor’ versus ‘Low-Income’: What Term Should We Use?” Oxfam, 2015. https://politicsofpoverty.oxfamamerica.org/poor-versus-low-income-what-term-should-we-use/.

Krippendorff, Klaus. “Cybernetics’s Reflexive Turns,” 2008, 14.

Kurian, Anji. “Interview with David Crosbie: Invest in Changing the Public Discourse about Charities in Australia.” Governance Directions 71, no. 10 (2019): 531–34.

Lam, Maria Lai-Ling, and Kei-Wing Wong. “Human Flourishing in Smart Cities.” In Digital Transformation in Business and Society: Theory and Cases, edited by Babu George and Justin Paul, 13–37. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-08277-2_2.

Lessof, Carli. “Ethical Issues in Longitudinal Surveys.” In Methodology of Longitudinal Surveys, edited by Peter Lynn, 35–54. Chichester, UK: John Wiley & Sons, Ltd, 2009. https://doi.org/10.1002/9780470743874.ch3.

“Longitudinal Data - an Overview | ScienceDirect Topics.” Accessed July 13, 2022. https://www.sciencedirect.com/topics/mathematics/longitudinal-data.

Luke, Joanne, Ebony Verbunt, Angela Zhang, Muriel Bamblett, Gabrielle Johnson, Connie Salamone, David Thomas, et al. “Questioning the Ethics of Evidence-Based Practice for Indigenous Health and Social Settings in Australia.” BMJ Global Health 7, no. 6 (2022): e009167. https://doi.org/10.1136/bmjgh-2022-009167.

Machado, Inês Araújo, Carlos Costa, and Maribel Yasmina Santos. “Data Mesh: Concepts and Principles of a Paradigm Shift in Data Architectures.” Procedia Computer Science 196 (2022): 263–71. https://doi.org/10.1016/j.procs.2021.12.013.

MacIntyre, C. Raina. “Social Media for Tracking Disease Outbreaks – Fad or Way of the Future?” The Conversation. Accessed July 13, 2022. http://theconversation.com/social-media-for-tracking-disease-outbreaks-fad-or-way-of-the-future-66401.

MacKay, Donald M. Information, Mechanism and Meaning. The MIT Press. Cambridge: MIT Press, 1969. https://doi.org/10.7551/mitpress/3769.001.0001.

Markus, Andrew. Mapping Social Cohesion the Scanlon Foundation Surveys 2019. Caulfield East, Vic.: Monash University, 2019.

Marston, Greg, and Rob Watts. "'Just the Facts Ma'am': A Critical Appraisal of Evidence-Based Policy." Just Policy: A Journal of Australian Social Policy, no. 30 (2003): 32–46.

Martinez, Arturo, and Francisco Perales. "The Dynamics of Multidimensional Poverty in Contemporary Australia." Social Indicators Research 130, no. 2 (January 1, 2017): 479–96. https://doi.org/10.1007/s11205-015-1185-1.

Maslow, Abraham H., and John J. Honigmann. "Synergy: Some Notes of Ruth Benedict." American Anthropologist 72, no. 2 (1970): 320–33.

"Mastering Systems Thinking in Practice." Accessed October 6, 2022. https://www.open.edu/openlearn/mod/oucontent/science-maths-technology/mastering-systems-thinking-practice.

McCaa, R., S. Ruggles, M. Davern, T. Swenson, and K. M. Palipudi. "IPUMS-International High Precision Population Census Microdata Samples: Balancing the Privacy-Quality Tradeoff by Means of Restricted Access Extracts." In Privacy in Statistical Databases, edited by J. Domingo-Ferrer and L. Franconi. Heidelberg: Springer, 2006. https://doi.org/10.1007/11930242_31.

McCaa, R., S. Ruggles, and M. Sobek. "IPUMS-International Statistical Disclosure Controls: 159 Census Microdata Samples in Dissemination, 100+ in Preparation." In Privacy in Statistical Databases, edited by J. Domingo-Ferrer and E. Magkos. Heidelberg: Springer, 2010. https://doi.org/10.1007/978-3-642-15838-4_7.

McCallum, Kerry, Tess Ryan, and Jo Caffery. "Deficit Metrics in Australian Indigenous Education: Through a Media Studies Lens." Discourse: Studies in the Cultural Politics of Education 43, no. 2 (2022): 266–81. https://doi.org/10.1080/01596306.2020.1828285.

McCosker, Anthony, Frances Shaw, Xiaofang Yao, and Kath Albury. "A Data Capability Framework for the Not-for-Profit Sector." Report. Swinburne University of Technology, November 16, 2022. Australia. https://apo.org.au/node/320678.

McDonald, Catherine, and Lesley Chenoweth. "(Re) Shaping Social Work: An Australian Case Study." The British Journal of Social Work 39, no. 1 (2009): 144–60.

McGraw, Deven, and Carolyn Petersen. "From Commercialization to Accountability: Responsible Health Data Collection, Use, and Disclosure for the 21st Century." Applied Clinical Informatics 11, no. 2 (March 2020): 366–73. https://doi.org/10.1055/s-0040-1710392.

McKay, S. "Poverty or Preference: What Do 'Consensual Deprivation Indicators' Really Measure?" Fiscal Studies 25, no. 2 (2004): 201–23. https://doi.org/10.1111/j.1475-5890.2004.tb00102.x.

McLachlan, Rosie, Geoff Gilfillan, and Jenny Gordon. "Deep and Persistent Disadvantage in Australia." Productivity Commission Staff Working Paper. Productivity Commission, 2013. https://www.pc.gov.au/research/supporting/deep-persistent-disadvantage/deep-persistent-disadvantage.pdf.

Mendes, Philip. "Why Isn't Poverty a Priority for Australia's Major Political Parties?" Monash Lens, 2022. https://lens.monash.edu/@politics-society/2022/05/04/1384645?slug=why-is-poverty-not-a-priority-for-australias-major-political-parties.

Mendes, Philip, and Steven Roche. "How Do Australian Policymakers Frame the Causes of and Policy Solutions to Poverty? A Critical Examination of Anti-Poverty Week Parliamentary

Debates from 2012 to 2021." Australian Journal of Social Issues 00, no. n/a (2023): 1–15. https://doi.org/10.1002/ajs4.254.

Middleton, Karen. "'Pretty Creepy': Agencies Illegally Obtained Emails, Voicemails and Texts." Text. The Saturday Paper, September 17, 2022. https://www.thesaturdaypaper.com.au/news/politics/2022/09/17/pretty-creepy-agencies-illegally-obtained-emails-voicemails-and-texts.

Mitchell, W. Bede. "Elena S. Danielson. The Ethical Archivist. Chicago: Society of American Archivists, 2010. 440p. Alk. Paper, $49.00 (ISBN 978-1-931666-34-2). LC2010-026680. | Mitchell | College & Research Libraries," April 25, 2017. https://doi.org/10.5860/0720302.

Monat, Jamie P., and Thomas F. Gannon. "What Is Systems Thinking? A Review of Selected Literature Plus Recommendations." American Journal of Systems Science 4, no. 1 (2015): 11–26.

Moynihan, Daniel Patrick. "(1965) The Moynihan Report: The Negro Family, the Case for National Action," 2007. https://www.blackpast.org/african-american-history/moynihan-report-1965/.

Musgrove, Erica, Loretta Gasparini, Katie McBain, Susan A. Clifford, Simon A. Carter, Helena Teede, and Melissa Wake. "Synthesizing Core Outcome Sets for Outcomes Research in Cohort Studies: A Systematic Review." Pediatric Research 92, no. 4 (October 2022): 936–45. https://doi.org/10.1038/s41390-021-01801-2.

New York Times. "CENSUS OF INCOME CALLED HARMFUL: Dr. McFall, Who Headed First Check of Business, Fears Distortion of Markets." 1950, sec. Business & Finance.

Nguyen, C. Thi. "Transparency Is Surveillance." Philosophy and Phenomenological Research, August 2, 2021, phpr.12823. https://doi.org/10.1111/phpr.12823.

Nkyekyer, Joanna, Susan A Clifford, Fiona K Mensah, Yichao Wang, Lauren Chiu, and Melissa Wake. "Maximizing Participant Engagement, Participation, and Retention in Cohort Studies Using Digital Methods: Rapid Review to Inform the Next Generation of Very Large Birth Cohorts." Journal of Medical Internet Research 23, no. 5 (May 14, 2021): e23499. https://doi.org/10.2196/23499.

Obermeyer, Ziad, and Sendhil Mullainathan. "Dissecting Racial Bias in an Algorithm That Guides Health Decisions for 70 Million People." In Proceedings of the Conference on Fairness, Accountability, and Transparency, 89. FAT* '19. New York, NY, USA: Association for Computing Machinery, 2019. https://doi.org/10.1145/3287560.3287593.

OECD. Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-Use across Societies. OECD, 2019. https://doi.org/10.1787/276aaca8-en.

Olic, Cath. "Mapping the Potential." Catholic Social Services Australia. Accessed November 29, 2022. https://cssa.org.au/research/mapping-the-potential/.

Oliver, Kathryn A, and Frank de Vocht. "Defining 'evidence' in Public Health: A Survey of Policymakers' Uses and Preferences." European Journal of Public Health 27, no. suppl_2 (2017): 112–17. https://doi.org/10.1093/eurpub/ckv082.

Ong, Walter J., John Hartley, and John Hartley. Orality and Literacy: The Technologizing of the Word. 3rd ed. London: Routledge, 2012. https://doi.org/10.4324/9780203103258.

Owens, John, Vikki A. Entwistle, Luke K. Craven, and Ina Conradie. "Understanding and Investigating Relationality in the Capability Approach." Journal for the Theory of Social Behaviour 52, no. 1 (2022): 86–104. https://doi.org/10.1111/jtsb.12310.

Palla, Gergely, Imre Derenyi, Illes Farkas, and Tamas Vicsek. "Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society." Nature 435, no. 7043 (June 2005): 814–18. https://doi.org/10.1038/nature03607.

Panek, Tomasz, and Jan Zwierzchowski. “Examining the Degree of Social Exclusion Risk of the Population Aged 50 + in the EU Countries Under the Capability Approach.” Social Indicators Research 163, no. 3 (October 1, 2022): 973–1002. https://doi.org/10.1007/s11205-022-02930-9.

Parliament of Australia. “Chapter 2 - Defining and Measuring Poverty.” Australia. Accessed March 6, 2023. https://www.aph.gov.au/parliamentary_business/committees/senate/community_affairs/completed_inquiries/2002-04/poverty/report/c02.

PASK, Gordon. An Approach to Cybernetics. First Edition. Harper and Brothers, 1961.

Pawar, Manohar, Dominic O’Sullivan, Belinda Cash, Richard Culas, Kiprono Langat, Andrew Manning, Ndungi Mungai, John Rafferty, Satyan Rajamani, and Wesley S. Ward. “The Sustainable Development Goals: An Australian Response.” The International Journal of Community and Social Development 2, no. 4 (2020): 374–93. https://doi.org/10.1177/2516602620983716.

Perry, B. “The Mismatch between Income Measures and Direct Outcome Measures of Poverty.” Social Policy Journal of New Zealand 19, no. 19 (2002): 101–27.

Perry, Laura. “Educational Disadvantage Is a Huge Problem in Australia – We Can’t Just Carry on the Same.” The Conversation. Accessed September 21, 2022. http://theconversation.com/educational-disadvantage-is-a-huge-problem-in-australia-we-cant-just-carry-on-the-same-74530.

Peters, Benjamin. How Not to Network a Nation: The Uneasy History of the Soviet Internet. Information Policy. Cambridge: MIT Press, 2016. https://doi.org/10.7551/mitpress/9800.001.0001.

Phillips, Shawn, Ami Seivwright, Susan Young, and Et Al. “Insights into Hardship and Disadvantage in Perth, Western Australia: The 100 Families WA Report.” 100 Families WA, 2021. https://doi.org/10.25916/B914-1J34.

Pickering, Andrew. The Cybernetic Brain: Sketches of Another Future. Chicago, USA: University of Chicago Press, 2011. http://ebookcentral.proquest.com/lib/anu/detail.action?docID=530446.

Piketty, Thomas. “7. Inequality and Concentration: Preliminary Bearings.” In 7. Inequality and Concentration: Preliminary Bearings, 297–338. Harvard University Press, 2017. https://doi.org/10.4159/9780674982918-010.

Pilkington, Rhiannon, Alicia Montgomerie, Janet Grant, Angela Gialamas, Catia Malvaso, Lisa Smithers, Catherine Chittleborough, and John Lynch. “An Innovative Linked Data Platform to Improve the Wellbeing of Children — the South Australian Early Childhood Data Project,” 2019. https://www.aihw.gov.au/getmedia/75c9f692-c72c-438d-b68b-91318f780154/Australias-Welfare-Chapter-8-summary-18Sept2019.pdf.aspx.

Poel, Ibo van de. “The Relation Between Forward-Looking and Backward-Looking Responsibility.” In Moral Responsibility: Beyond Free Will and Determinism, edited by Nicole A. Vincent, Ibo van de Poel, and Jeroen van den Hoven, 37–52. Library of Ethics and Applied Philosophy. Dordrecht: Springer Netherlands, 2011. https://doi.org/10.1007/978-94-007-1878-4_3.

“Policy and Practice Forum on the Multi-Agency Data Integration Project (MADIP) | National Centre for Epidemiology and Population Health.” The Australian National University. Accessed July 28, 2022. https://nceph.anu.edu.au/research/tools-resources/policy-and-practice-forum-multi-agency-data-integration-project-madip.

Prinsloo, Shaun. “Henderson Poverty Line.” Melbourne Institute: Applied Economic & Social Research, 2019. https://melbourneinstitute.unimelb.edu.au/research/labour/henderson-poverty-line.

"Privacy Impact Assessments | Australian Bureau of Statistics," February 6, 2022. https://www.abs.gov.au/about/legislation-and-policy/privacy/privacy-impact-assessments.

Purdam, K., and M. Elliot. "A Case Study of the Impact of Statistical Disclosure Control on Data Quality in the Individual UK Samples of Anonymised Records." Environment and Planning A 39 (2007). https://doi.org/10.1068/a38335.

Ramage, Magnus. Systems Thinkers. 2nd ed. 2020. London: Springer London, 2020. https://doi.org/10.1007/978-1-4471-7475-2.

Rauch, Christopher B., Mat Kelly, John A. Kunze, and Jane Greenberg. "FAIR Metadata: A Community-Driven Vocabulary Application." In Metadata and Semantic Research, edited by Emmanouel Garoufallou, María-Antonia Ovalle-Perandones, and Andreas Vlachidis, 1537:187–98. Communications in Computer and Information Science. Cham: Springer International Publishing, 2022. https://doi.org/10.1007/978-3-030-98876-0_16.

"Return of Cybernetics." Nature Machine Intelligence 1, no. 9 (September 2019): 385–385. https://doi.org/10.1038/s42256-019-0100-x.

Rid, Thomas. Rise of the Machines – A Cybernetic History. Illustrated edition. New York: W. W. Norton & Company, 2017.

Robeyns, Ingrid, and Morten Fibieger Byskov. "The Capability Approach." In The Stanford Encyclopedia of Philosophy, edited by Edward N. Zalta and Uri Nodelman, Summer 2023. Metaphysics Research Lab, Stanford University, 2023. https://plato.stanford.edu/archives/sum2023/entries/capability-approach/.

Rodgers, Joan R, and John L Rodgers. "Contributions of Longitudinal Data to Poverty Measurement in Australia." Economic Record 85, no. s1 (2009): S35–47. https://doi.org/10.1111/j.1475-4932.2009.00587.x.

Roser, Max. "Global Poverty in an Unequal World: Who Is Considered Poor in a Rich Country? And What Does This Mean for Our Understanding of Global Poverty?" Our World in Data, 2021. https://ourworldindata.org/higher-poverty-global-line.

Sarwar, T., S. Seifollahi, J. Chan, X. Zhang, V. Aksakalli, I. Hudson, K. Verspoor, and L. Cavedon. "The Secondary Use of Electronic Health Records for Data Mining: Data Characteristics and Challenges." ACM Computing Surveys 55, no. 2 (2023). https://doi.org/10.1145/3490234.

Saunders, P., and Y. Naidoo. "Mapping the Australian Poverty Profile: A Multidimensional Deprivation Approach." Australian Economic Review 51, no. 3 (2018): 336–50. https://doi.org/10.1111/1467-8462.12266.

Saunders, P., Y. Naidoo, and M. Griffiths. "Towards New Indicators of Disadvantage: Deprivation and Social Exclusion in Australia." Australian Journal of Social Issues 43, no. 2 (2008): 175–94. https://doi.org/10.1002/j.1839-4655.2008.tb00097.x.

Saunders, P., Y. Naidoo, and M. Wong. "Comparing the Monetary and Living Standards Approaches to Poverty Using the Australian Experience." Social Indicators Research 162, no. 3 (2022): 1365–85. https://doi.org/10.1007/s11205-022-02888-8.

Saunders, P., K. Sutherland, P. Davidson, A. Hampshire, S. King, and J. Taylor. Experiencing Poverty: The Voices of Low Income Australians, 2006.

Saunders, Peter. "Closing the Gap: The Growing Divide between Poverty Research and Policy in Australia." Australian Journal of Social Issues 50, no. 1 (2015): 13–35. https://doi.org/10.1002/j.1839-4655.2015.tb00332.x.

Saunders, Peter, Yuvisthi Naidoo, and Megan Griffiths. “Towards New Indicators of Disadvantage: Deprivation and Social Exclusion in Australia.” Australian Journal of Social Issues 43, no. 2 (2008): 175–94. https://doi.org/10.1002/j.1839-4655.2008.tb00097.x.

Saunders, Peter, Yuvisthi Naidoo, and Melissa Wong. “The Crumbling Pillar: Assessing the Impact of Housing Costs on Recent Trends in Poverty and Deprivation in Australia.” International Journal of Social Welfare 31, no. 4 (2022): 421–32. https://doi.org/10.1111/ijsw.12549.

Savage, Gill. “What Does Social Cohesion Mean in Australia?” The Strategist, September 6, 2021. https://www.aspistrategist.org.au/what-does-social-cohesion-mean-in-australia/.

Scanlon institute. “What Is Social Cohesion?” Accessed September 21, 2022. https://scanloninstitute.org.au/research/mapping-social-cohesion/what-social-cohesion.

Schooling, Jennifer, Mark Enzer, and Didem Gurdur Broo. “Flourishing Systems: Re-Envisioning Infrastructure as a Platform for Human Flourishing.” Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction 173, no. 1 (2020): 166–74. https://doi.org/10.1680/jsmic.20.00023.

Scott, Christy K., and William L. White. “Ethical Issues in the Conduct of Longitudinal Studies of Addiction Treatment.” Journal of Substance Abuse Treatment 28 Suppl 1 (2005): S91–101. https://doi.org/10.1016/j.jsat.2004.10.012.

Scutella, Rosanna, Michael Horn, and Roger Wilkins. “Measuring Poverty and Social Exclusion in Australia: A Proposed Multidimensional Framework for Identifying Socio-Economic Disadvantage.” Melbourne Institute Working Paper Series, 2009. https://melbourneinstitute.unimelb.edu.au/downloads/working-paper-series/wp2009n04.pdf.

Sen, Amartya. Resources, Values and Development. Oxford: Blackwell, 1984.

Shadbolt, Nigel, Alys Brett, Min Chen, Glenn Marion, Iain J. McKendrick, Jasmina Panovska-Griffiths, Lorenzo Pellis, Richard Reeve, and Ben Swallow. “The Challenges of Data in Future Pandemics.” Epidemics 40 (September 1, 2022): 100612. https://doi.org/10.1016/j.epidem.2022.100612.

Shannon, C. E. “A Mathematical Theory of Communication.” Bell System Technical Journal 27, no. 4 (1948): 623–56. https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.

Sittig, Dean F., Elisabeth Belmont, and Hardeep Singh. “Improving the Safety of Health Information Technology Requires Shared Responsibility: It Is Time We All Step Up.” Healthcare 6, no. 1 (March 1, 2018): 7–12. https://doi.org/10.1016/j.hjdsi.2017.06.004.

Sloot, Bart van der. “Privacy as Human Flourishing: Could a Shift towards Virtue Ethics Strengthen Privacy Protection in the Age of Big Data.” Journal of Intellectual Property, Information Technology and Electronic Commerce Law 5, no. 3 (2014): 230–44.

Social Ventures Australia. “Rethinking Social Networks for Equity, Excellence and Flourishing.” Accessed September 27, 2022. https://www.socialventures.com.au/sva-quarterly/rethinking-social-networks-for-equity-excellence-and-flourishing/.

Sontag, Lester W. “The History of Longitudinal Research: Implications for the Future.” Child Development 42, no. 4 (1971): 987–1002. https://doi.org/10.2307/1127786.

Spradley, James P. The Ethnographic Interview. New York: Holt, Rinehart And Winston, 1979.

Stahl, Bernd Carsten, Aequitas Andreou, Philip Brey, Tally Hatzakis, Alexey Kirichenko, Kevin Macnish, Stephanie Laulhé Shaelou, Andrew Patel, Mark Ryan, and David Wright. “Artificial Intelligence for Human Flourishing – Beyond Principles for Machine Learning.” Journal of Business Research 124 (2021): 374–88. https://doi.org/10.1016/j.jbusres.2020.11.030.

Staines, Zoe, Charlotte Moore, Greg Marston, and Louise Humpage. “Big Data and Poverty Governance under Australia and Aotearoa/New Zealand’s ‘Social Investment’ Policies.” Australian Journal of Social Issues 56, no. 2 (2021): 157–72. https://doi.org/10.1002/ajs4.129.

Stanford, Jessica. “Why Some Americans Don’t Trust the Census.” The Conversation. Accessed July 18, 2022. http://theconversation.com/why-some-americans-dont-trust-the-census-130109.

Stromberg, Brian. “The Limitations of Data-Driven Policymaking.” National Housing Conference, 2016. https://nhc.org/the-limitations-of-data-driven-policymaking/.

Sung, Valerie, Katrina Williams, Ella Perlow, Yanhong J. Hu, Susannah Ahern, Joanne M. Said, Bill Karanatsios, et al. “Enhancing Value and Uptake for Whole-Population Cohorts of Children and Parents: Methods to Integrate Registries into the Generation Victoria Cohort.” Children 8, no. 4 (April 7, 2021): 285. https://doi.org/10.3390/children8040285.

Taylor, David. “Wellbeing and Welfare: A Psychosocial Analysis of Being Well and Doing Well Enough.” Journal of Social Policy 40, no. 4 (2011): 777–94. https://doi.org/10.1017/S0047279411000249.

“The Effects of the 1930s HOLC ‘Redlining’ Maps (REVISED August 2020) - Federal Reserve Bank of Chicago.” Accessed August 26, 2022. https://www.chicagofed.org/publications/working-papers/2017/wp2017-12.

“The Evolving Privacy Landscape: 30 Years After the OECD Privacy Guidelines.” OECD Digital Economy Papers. Vol. 176. OECD Digital Economy Papers, April 6, 2011. https://doi.org/10.1787/5kgf09z90c31-en.

“The Story of Informatics in Human Flourishing.” Accessed February 21, 2023. https://flourish.umbc.edu/home/the-story-of-informatics-in-human-flourishing/.

“The Uses of Census Data: An Analytical Review.” Accessed July 18, 2022. https://clintonwhitehouse4.archives.gov/WH/EOP/CEA/html/censusreview.html.

Thienen, Lars Von. Decoding the Hidden Market Rhythm: Dynamic Cycles. CreateSpace Independent Publishing Platform, 2017.

Thurber, Katherine A., Joanne Thandrayen, Emily Banks, Kate Doery, Mikala Sedgwick, and Raymond Lovett. “Strengths-Based Approaches for Quantitative Data Analysis: A Case Study Using the Australian Longitudinal Study of Indigenous Children.” SSM - Population Health 12 (December 1, 2020): 100637. https://doi.org/10.1016/j.ssmph.2020.100637.

Thurman, Sabrina L. “Ethical Considerations in Longitudinal Studies of Human Infants.” Infant Behavior and Development 38 (February 1, 2015): 116–25. https://doi.org/10.1016/j.infbeh.2014.12.012.

Tigard, Daniel W. “Big Data and the Threat to Moral Responsibility in Healthcare.” In Datenreiche Medizin und das Problem der Einwilligung: Ethische, rechtliche und sozialwissenschaftliche Perspektiven, edited by Gesine Richter, Wulf Loh, Alena Buyx, and Sebastian Graf von Kielmansegg, 11–25. Berlin, Heidelberg: Springer, 2022. https://doi.org/10.1007/978-3-662-62987-1_2.

“To Create Fairness and Accountability in the Use of Government Decision Making Algorithms - Churchill Trust.” Accessed August 29, 2022. https://www.churchilltrust.com.au/fellow/dr-owen-churches-sa-2018.

Toros, Halil, and Daniel Flaming. “Silicon Valley Triage Tool.” Economic Roundtable Research Report. Rochester, NY, 2016. https://papers.ssrn.com/abstract=2772768.

Townsend, Peter. Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living. University of California Press, 1979.

"Transfers of Disadvantage across Three Generations Using Latent Class Associations within Families - Australian National University." Accessed January 19, 2023. https://anu.primo.exlibrisgroup.com.

"United Nations Economic Commission for Europe. Conference of European Statisticians. Managing Statistical Confidentiality & Microdata Access: Principles and Guidelines of Good Practice. United Nations, Geneva (2007), See Online Edition Annex 1.23, Http://Www.Unece.Org/Fileadmin/DAM/Stats/Publications/Managing.Statistical.Confidentiality.and.Microdata.Access.Pdf," n.d. http://www.unece.org/fileadmin/DAM/stats/publications/Managing.statistical.confidentiality.and.microdata.access.pdf.

Vallentyne, Peter. "Jonathan Wolff and Avner De-Shalit, Disadvantage (Oxford: Oxford University Press, 2007), Pp. Ix + 231." Utilitas 21, no. 4 (2009): 532–35. https://doi.org/10.1017/S0953820809990306.

Van den Bosch, K. Identifying the Poor: Using Subjective and Consensual Measures, 2001.

VanderWeele, Tyler J. "On the Promotion of Human Flourishing." Proceedings of the National Academy of Sciences 114, no. 31 (2017): 8148–56. https://doi.org/10.1073/pnas.1702996114.

Veenstra, Anne Fleur van, and Bas Kotterink. "Data-Driven Policy Making: The Policy Lab Approach." In Electronic Participation, edited by Peter Parycek, Yannis Charalabidis, Andrei V. Chugunov, Panos Panagiotopoulos, Theresa A. Pardo, Øystein Sæbø, and Efthimios Tambouris, 100–111. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-64322-9_9.

Vincent, Nicole A., Ibo van de Poel, and Jeroen van den Hoven. Moral Responsibility: Beyond Free Will and Determinism. Dordrecht, NETHERLANDS, THE: Springer Netherlands, 2011. http://ebookcentral.proquest.com/lib/anu/detail.action?docID=798715.

Von Foerster, Heinz. Understanding. New York, NY: Springer, 2003. https://doi.org/10.1007/b97451.

Wacquant, Loïc. Punishing the Poor: The Neoliberal Government of Social Insecurity. duke university Press, 2009.

Wake, Melissa, Sharon Goldfeld, and Andrew Davidson. "Embedding Life Course Interventions in Longitudinal Cohort Studies: Australia's GenV Opportunity." Pediatrics 149, no. Suppl 5 (2022): e2021053509R. https://doi.org/10.1542/peds.2021-053509R.

Wake, Melissa, Yanhong Jessika Hu, Hayley Warren, Margie Danchin, Michael Fahey, Francesca Orsini, Maurizio Pacilli, Kirsten P. Perrett, Richard Saffery, and Andrew Davidson. "Integrating Trials into a Whole-Population Cohort of Children and Parents: Statement of Intent (Trials) for the Generation Victoria (GenV) Cohort." BMC Medical Research Methodology 20 (September 24, 2020): 238. https://doi.org/10.1186/s12874-020-01111-x.

Walter, Maggie. "The Voice of Indigenous Data: Beyond the Markers of Disadvantage." Griffith REVIEW, no. 60 (2018): 256–63.

Wang, Jing, Yanhong Hu, Lana Collins, Anna Fedyukova, Varnika Aggarwal, Fiona Mensah, Jeanie Cheong, and Melissa Wake. "Study Protocol: Generation Victoria (GenV) Special Care Nursery Registry." International Journal of Population Data Science 8, no. 1 (2023). https://doi.org/10.23889/ijpds.v8i1.2139.

Wang, Jing, Yanhong Jessika Hu, Susan Clifford, Sharon Goldfeld, and Melissa Wake. "Selecting Life Course Frameworks to Guide and Communicate Large New Cohort Studies: Generation Victoria (GenV) Case Study." Journal of Developmental Origins of Health and Disease 12, no. 6 (December 2021): 829–48. https://doi.org/10.1017/S2040174420001245.

Watson, Robert, John Thwaites, Dave Griggs, Tahl Kestin, and Kim McGrath. “Sustainable Development Goals and Targets for Australia,” 2014. https://ap-unsdsn.org/wp-content/uploads/2013/10/SDGs-for-Australia_Interim-Report.pdf.

Weir, Michael. Goal-Directed Behaviour. Studies in Cybernetics, v. 6. New York: Gordon and Breach Science Publishers, 1984.

“What Is Community Disadvantage? Understanding the Issues, Overcoming the Problem.” Accessed February 6, 2023. https://aifs.gov.au/resources/policy-and-practice-papers/what-community-disadvantage-understanding-issues-overcoming.

Wiley.com. “Systems Thinking, Systems Practice: Includes a 30-Year Retrospective | Wiley.” Accessed December 5, 2022.

Wilson, William J. “The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.” Chicago; University of Chicago Press, 2012.

Wolfert, Sjaak, Lan Ge, Cor Verdouw, and Marc-Jeroen Bogaardt. “Big Data in Smart Farming – A Review.” Agricultural Systems 153 (May 2017): 69–80. https://doi.org/10.1016/j.agsy.2017.01.023.

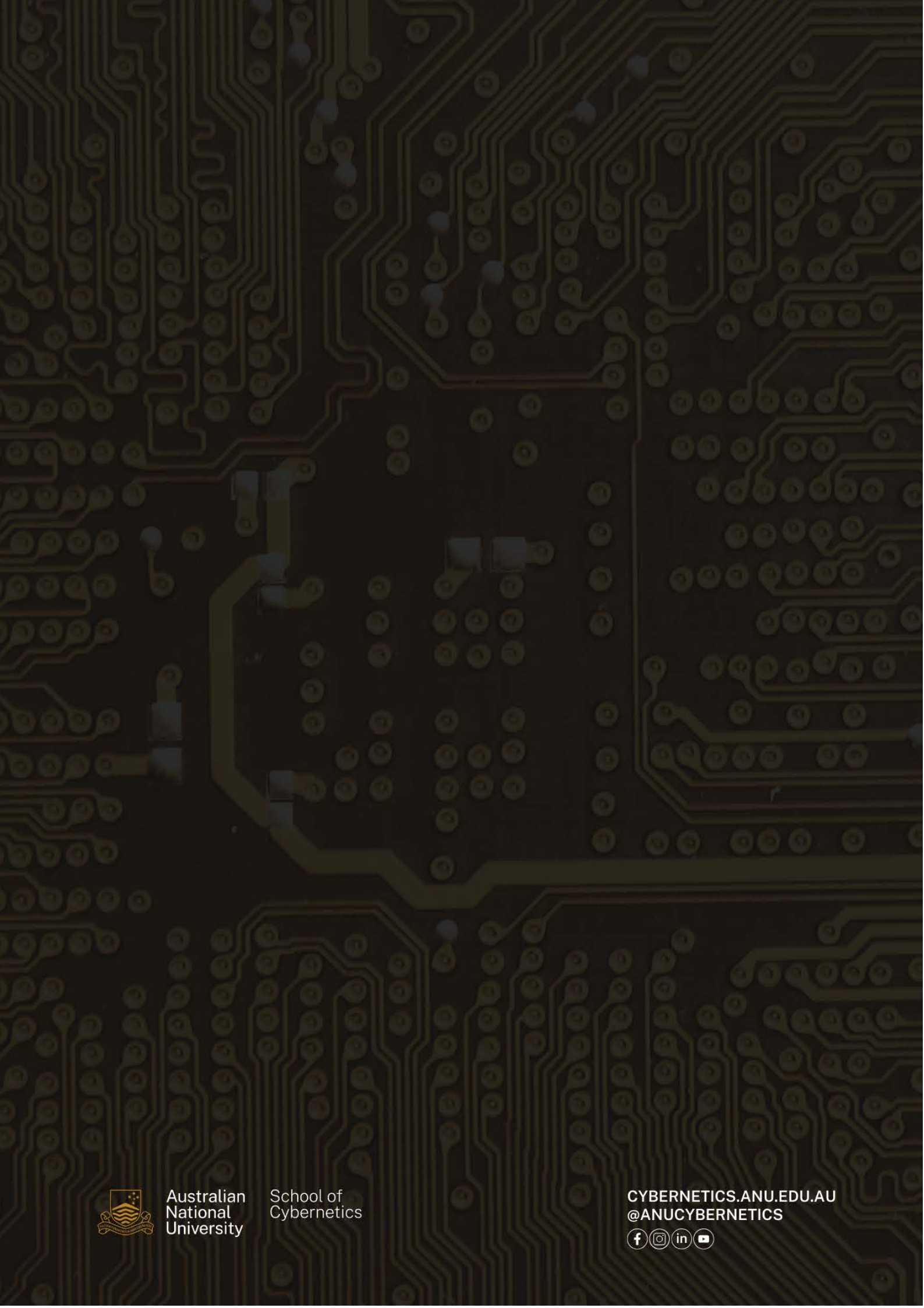
Wolff, Jonathan, and Avner De-Shalit. Disadvantage. Oxford: Oxford University Press, 2007.

World Bank. “Fact Sheet: An Adjustment to Global Poverty Lines,” 2022. https://www.worldbank.org/en/news/factsheet/2022/05/02/fact-sheet-an-adjustment-to-global-poverty-lines.

Yolles, Maurice, and Paul Iles. “Exploring Public–Private Partnerships through Knowledge Cybernetics.” Systems Research and Behavioral Science 23, no. 5 (2006): 625–46. https://doi.org/10.1002/sres.803.

Zhao, Juan, QiPing Feng, Patrick Wu, Roxana A. Lupu, Russell A. Wilke, Quinn S. Wells, Joshua C. Denny, and Wei-Qi Wei. “Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction.” Scientific Reports 9 (January 24, 2019): 717. https://doi.org/10.1038/s41598-018-36745-x.

Zook, Matthew, Solon Barocas, Danah Boyd, Kate Crawford, Emily Keller, Seeta Peña Gangadharan, Alyssa Goodman, et al. “Ten Simple Rules for Responsible Big Data Research.” PLOS Computational Biology 13, no. 3 (March 30, 2017): https://doi.org/10.1371/journal.pcbi.1005399.



Australian National University

School of Cybernetics

CYBERNETICS.ANU.EDU.AU
@ANUCYBERNETICS