



Joint Parliamentary Committee of Public Accounts and Audit Inquiry into Public Sector AI Use

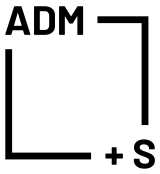


Authors

Kimberlee Weatherall, José-Miguel Bello y Villarino, Gerard Goggin, Jake Goldenfein, Paul Henman, Rita Matulionyte, Christine Parker, Lyndal Sleep, Georgia van Toorn.

ARC Centre of Excellence for Automated
Decision-Making and Society

25 October 2024



Acknowledgement of Country

In the spirit of reconciliation, we acknowledge the Traditional Custodians of Country throughout Australia and their connections to land, sea and community. We pay our respect to their Elders past and present and extend that respect to all Aboriginal and Torres Strait Islander peoples today.

Suggested Citation

Weatherall, K; Bello y Villarino, J-M, Goggin, G., Goldenfein, J., Henman, H., Matulionyte, R., Parker, C., Sleep, L., van Toorn, G. (2024). Joint Parliamentary Committee of Public Accounts and Audit Inquiry into Public Sector AI Use. ARC Centre of Excellence for Automated Decision-Making and Society. DOI: 10.60836/mn49-wy88

Copyright 2024

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited.

ARC Acknowledgement

This research was funded by the Australian Government through the Australian Research Council's Centre of Excellence for Automated Decision Making and Society (ADM+S) [CE200100005].



Australian Government
Australian Research Council

About ADM+S

The ADM+S is pleased to have this opportunity to engage with the Joint Committee on the important issue of public sector AI use. The [ARC Centre of Excellence for Automated Decision-Making and Society \(ADM+S\)](#) is a cross-disciplinary, national research centre established and supported by the Australian Research Council to create the knowledge and strategies necessary for responsible, ethical, and inclusive automated decision-making (ADM). More information about ADM+S and its research may be found on our website, www.admscentre.org.au.

This submission

This submission is the product of a collaborative process involving direct contributions from the above researchers from ADM+S, as led and consolidated by Kimberlee Weatherall. ADM+S researchers come from many different institutions, disciplines and perspectives. It should not be assumed that every contributing author, or every member of the Centre subscribes to every comment or recommendation in this submission. The submission represents our best effort to consolidate research and thinking in a way that can be useful to the government's deliberations on safe and responsible AI and mandatory guardrails. We are happy to put members of the government's team in touch with experts in the Centre for further discussion.

Executive Summary

The public sector should be an exemplar of safe and responsible artificial intelligence (AI) use. Not only that, the public sector should demonstrate the positive potentials of technology, and should, in its use of AI, demonstrate the **positive** impacts that technology can have in achieving important public goals, such as promoting access, inclusion, and better public services.

This submission, drawing on research from across the ARC Centre of Excellence on Automated Decision-Making and Society (ADM+S) seeks to provide material for consideration of the Committee.

Addressing **TOR1**, we summarise evidence from ADM+S world-leading studies of the use of automated decision-making (ADM) and artificial intelligence in the public sector. That evidence shows widespread, and increasing use of automated decision-making, including AI - across portfolios; across domains; from the mundane to the challenging. Both the research and methods used are public. We would be happy to elaborate on any of this evidence.

Addressing **TOR2**, we consider some existing, and emerging legal, regulatory and policy frameworks governing AI use in the public sector. We show, first, that existing legal frameworks - including privacy law, administrative law, and freedom of information all require consideration and updating in light of public sector AI use and its impacts. In short, we agree with Commissioner Holmes' recommendation, in the context of the Royal Commission into the Robodebt Scheme, that the Commonwealth should clarify the law relating to the use of automation in administrative decision making.

More pertinently, we address the current proliferation of frameworks around AI use. These frameworks get to the heart of whether government will achieve its stated goal of being an **exemplar** in AI use. We conclude that there is a high risk of failure in this goal at present.

To date, there has been a proliferation of frameworks, guidance, and activity by governments 'about AI'. Unfortunately, the standards the Commonwealth government has set for itself do not establish government as an exemplar. They are weaker than (admittedly voluntary) frameworks drafted for the private sector, and weaker than mandatory rules which apply in equivalent jurisdictions overseas.

In addition, the proliferating guidelines and frameworks **fail** to provide consistent, coherent, clear rules for AI development, deployment or use. Too many rules; too many of them optional or weak; too many differences: the goal of consistency, coherence, and building public trust or trustworthiness cannot be met as things stand.

We have previously made submissions in relation to the government's proposed **Mandatory Guardrails**. In summary, and as relevant to this Committee, we think the Mandatory Guardrails provide an opportunity to take steps to achieve the goal of establishing the Australian government as an exemplar of good AI use:

- We support the enactment of horizontal, generally applicable, mandatory guardrails, along with mechanisms for updating the guardrails over time.
- We think this is especially important across the public sector (since government is bound by higher standards and ought to be an exemplar of safe and responsible use).
- The guardrails proposed — requirements such as risk management, testing, transparency, accountability, and data governance — represent elements of good governance and process that are appropriate across a wide range of different policy domains and different kinds of organisations. They are suitable for application across the Commonwealth government.
- To avoid continuing proliferation and weak standards, a coordinating body with power to impose and monitor requirements across the Commonwealth government is required.

On **TOR5 (line of sight)** we outline research that suggests there is no adequate line of sight over AI development, deployment or use in NSW governments, and express our doubt that recently announced policy in the Commonwealth will create adequate accountability.

On **TOR6 (internal capacity)** we address one element of public sector capacity around AI adoption: the important capacity to engage with a range of stakeholders to ensure democratic control over, and benefits from AI use. We argue that considerable work is required to improve broad engagement — but the foundations are available.

On **TOR8 (other issues)** we highlight three key areas of our research that raise important considerations for the committee:

1. **Disability and accessibility:** ADM+S research raises doubts whether accessibility is sufficiently considered in AI development and adoption. This is critical: people with disabilities depend on many public and social services, and so are highly likely to be impacted by public sector AI use.
2. **Environmental impacts:** the environmental impacts of AI use, including by the public sector, must be taken into account, and methods developed for assessing such impacts.
3. **Trauma-informed approaches:** considerable proportions of the population are affected by trauma of a range of kinds, and many will interact with public sector AI and ADM uses. ADM+S researchers have developed a trauma-informed toolkit that can help ensure that uses of AI and ADM do not increase trauma suffered by members of the Australian population.

Contents

About ADM+S.....	1
This submission.....	1
Executive Summary	2
The Scope of the Inquiry.....	6
TOR 1: Purposes for current, planned and likely future uses of AI in the public sector ..	7
General observations.....	7
The evidence about AI/ADM use in NSW	9
Use of ADM.....	10
Purposes of ADM	10
Evidence about ADM/AI use in social services in Australia.....	12
AI/ADM in public sector beyond Australia.....	16
The United States of America	16
Europe	17
United Kingdom.....	17
TOR 2: The effectiveness of existing legislative, regulatory and policy frameworks relating to AI	19
Pre-existing legislative, regulatory and policy frameworks	19
Transparency and freedom of information	20
Administrative law.....	21
Privacy and security	22
Legislative, regulatory and policy frameworks explicitly covering AI	22
National Framework for the Assurance of AI in Government	23
Policy for the responsible use of AI in government	24
Technology Protection Framework.....	25
Other frameworks and guidelines	25
The move to mandated guardrails	26
TOR4: Internal framework, policies or additional controls used for assessing risks associated with the use and possible misuse of AI	28
TOR5: The capacity for line of sight to the output of AI, and assisted decisions, to ensure ethical and responsible use	28
TOR6: Internal capabilities of the public sector to effectively adopt and utilise AI	30
Internal capabilities and public participation	30
Enhancing democratic legitimacy and trust in public sector AI/ADM	32

TOR8: Other issues relating to the use of AI in the public sector	33
Disability and Accessibility.....	33
Accessibility implications of AI uses: are these being considered?.....	35
Bias and Discrimination: Are These Being Considered?	36
Uses of Public Sector AI that promote digital inclusion	37
Environmental Impact	38
Trauma-Informed AI in the Public Sector	40

The Scope of the Inquiry

As a preliminary matter, we would argue that **public sector use of artificial intelligence (AI) should be investigated and understood as part of a broader move to automate government activities**, for two reasons:

1. The distinction between automation and AI is not **clear**;¹
2. The distinction is not **important**: whether it involves AI or not, public sector automation can significantly affect citizen's rights and good public sector administration — and in similar ways.

The Robodebt scheme did not involve AI, but highlighted the **impact** of automation on the public, and the **challenges** for public administrators when designing, managing and applying automated systems, and the difficulties for **governance** of systems that are not understood by affected individuals or the wider public.

The key thing to be governed is **the government practice**, not the particular technology through which it occurs. An example is **risk profiling**. The Australian government has long deployed risk profiling of jobseekers: allocating levels of funding according to Job Network providers according to jobseekers' assessed risk of long-term unemployment. The original Job Seekers Classification Index was developed following standard statistical (multiple regression) analysis and coded into the tool.² We understand the more recent replacement, the Job Seeker Assessment Framework uses machine learning techniques for the same purposes. Either way, to the wider public and jobseekers, and likely many within the system, the nature of these profiling calculations is a 'black box'.³

¹ Definitions of AI vary, but the NSW Government's definition captures commonly-used approaches: "AI is the ability of a computer system to perform tasks that would normally require human intelligence, such as learning, reasoning, and making decisions."

<https://www.digital.nsw.gov.au/policy/artificial-intelligence/a-common-understanding-simplified-ai-definitions-from-leading#anchor-what-is-ai>

² McDonald, C., Marston, G., & Buckley, A. (2003). 'Risk technology in Australia: The role of the job seeker classification instrument in employment services.' *Critical Social Policy*, 23(4), 498-525; <https://www.dss.gov.au/sites/default/files/documents/08_2022/job-seeker-classification-instrument-guidelines-v-14.pdf>. For an international perspective see: Desiere, S., Langenbacher, K., & Struyven, L. (2019). *Statistical profiling in public employment services: An international comparison*, OECD Papers, Paris.

³ Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.

TOR 1: Purposes for current, planned and likely future uses of AI in the public sector

ADM+S researchers have conducted the most significant, systematic, and independent research into the use of automated decision-making systems (ADM), including artificial intelligence, in Australia, and indeed globally. We describe **some** of the results of that research below. Much more data, including examples of systems, case studies and more are available in our larger research report. We refer the Committee in particular to the Executive Report of our project mapping ADM use in NSW for an accessible (shorter) version of our findings.⁴

We are also aware that the Australian Digital Transformation Agency conducted a survey of ADM in Federal agencies in 2023. Their findings have not been made public.

General observations

We make some general observations from our research and that of others before presenting more evidence below.

First, some AI use in the public sector is **commonplace, even mundane** — and has been for some time. Some AI use can be so mundane and everyday that people overlook it. For example, some voice to text natural language processing systems (NLP) — similar to those used in Amazon’s Alexa — have for many years been in use without being considered as AI. Services Australia has used such systems for automated call routing and identification for over 5 years. Even newer technologies, such as generative AI, are becoming everyday. Generative AI tools such as Microsoft Copilot are increasingly being used in mundane, low risk ways, such as producing transcripts of meetings and editing internal and external communications from drafts, as is evidenced in the recent Federal government’s Microsoft 365 Copilot trial.⁵ Chatbots for customer-public agency interaction have rapidly grown, and have

⁴ Weatherall, K., Henman, P., Bello y Villarino, J-M., Matulionyte, R., Sleep, L., Trezise, M., Van Der Arend, J., Wilcock, S., Executive Report - Automated decision-making in NSW: Mapping and analysis of the use of ADM systems by state and local governments (Research Report) (ADM+S, 2024), <<https://www.ombo.nsw.gov.au/Find-a-publication/publications/reports-to-parliament/other-special-reports/a-map-of-automated-decision-making-in-the-nsw-public-sector-a-special-report-to-parliament>> ('ADM in NSW').

⁵ Australian Government, 'Australian Government trial of Microsoft 365 Copilot' <<https://www.digital.gov.au/initiatives/copilot-trial>>

accelerated with Generative AI, and also in assisting public sector employees within agencies such as in navigating complex legislation or identifying overlaps in patents.⁶

Even mundane or everyday use of ADM and AI requires care, attention, and good governance. Chatbots may appear to provide a low risk example of use of AI, but considerable care is needed to ensure accurate information is provided, and in a way that is appropriate and trauma-informed.⁷

Second, our research and that of others demonstrates expansion of automation and AI use **upstream of final decision-making**. Triage for services or compliance reviews are a common example. In these cases, AI/ML highlights cases for consideration, or provides assessment of available knowledge that can then be used by human officers to determine next steps or make official administrative decisions. The strength of AI/ML is in identifying patterns and operating on probabilities, making it less suitable for administrative decision-making which involves the application of well coded laws to cases (the 'if X, then Y' model). Rather, they operate more like "judgement machines"⁸ — making assessments within ambiguity and complexity. When used in relation to administrative decisions that have legal effects, they are typically used to inform human judgement and, as far as we can tell, at least for the present, rarely operate solely entirely autonomously.

Upstream AI use also presents challenges and requires good governance:

- Use of 'triage' or collecting and analysing data can render assumptions and limitations of AI/ML tools invisible to both operators and people affected, which is especially problematic if human officers defer to the information provided by the digital tools;⁹
- Computerisation can progressively reduce administrative discretion.¹⁰

⁶ Chen, T., Gascó-Hernandez, M., & Esteve, M. (2024) 'The Adoption and Implementation of Artificial Intelligence Chatbots in Public Organizations: Evidence from US State Governments', *The American Review of Public Administration* 54(3), 255-270; Cortés-Cediel, M.E., Segura-Tinoco, A., Cantador, I., & Bolívar, M.P.R. (2023). 'Trends and challenges of e-government chatbots: Advances in exploring open government data and citizen participation content', *Government Information Quarterly*, 40(4), 101877.

⁷ See for example the case of Tessa which when updated to a more Generative AI model gave inappropriate and dangerous advice to people with eating disorders: Hoover, A., 'An Eating Disorder Chatbot Is Suspended for Giving Harmful Advice', *Wired*, (1 June 2023), <<https://www.wired.com/story/tessa-chatbot-suspended>>.

⁸ Henman, P. (2020). 'Governing by algorithms and algorithmic governmentality: Towards machinic judgement' in *The Algorithmic Society* (pp. 19-34) Routledge.

⁹ Gillingham, P. (2011). 'Decision-making tools and the development of expertise in child protection practitioners: are we 'just breeding workers who are good at ticking boxes'?' *Child & Family Social Work* 16(4), 412-421.

¹⁰ Busch, P. A., & Henriksen, H.Z. (2018). 'Digital discretion: A systematic literature review of ICT and street-level discretion', *Information Polity* 23(1), 3-28; Bovens, M., and Zouridis, S. (2002)

Third, **more autonomy is possible, and may be difficult to identify when it happens.** While humans are mostly making final decisions for now, in our research we found many technical systems that are **technically capable of autonomous operation.** This means a decision could be made to automate a process or part of a process, without there being a system change or procurement. For example (and hypothetically), while AI used for assessing traffic infringements for non compliance with seat belts or mobile phone use have typically involved human operators, as accuracy increases humans could be removed. The process for doing so could be technically simple, and not observable to politicians, the general public or even higher level public servants. This underlines a need for very clear policies and governance systems, including line of sight over decisions to automate within existing systems.

Finally, we note that the **values** embedded in automated systems are also important. Researchers have noted the ways in which AI (and ADM) are deployed reflect both administrative agendas (e.g. efficiency, managing complexity, managing workloads) and wider socio-political agendas. These values get embedded and accentuated in the tools that are adopted. Researchers have argued that alternative approaches to designing and deploying AI/ML are needed, such as for building understanding,¹¹ providing and advancing wellbeing.

The evidence about AI/ADM use in NSW

In 2023, ADM+S mapped the use of **ADM systems** by state and local governments in NSW, as part of a project supported by the NSW Ombudsman's Office.¹² ADM+S researchers used:

- direct surveys to public servants asking them to report, categorise and briefly describe the systems used by their department, agency, or local council;
- a systematised search and human review of publicly available material published by every agency, government department and local council;
- and a small set of case studies based on interviews with public servants.

'From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control', Public administration review 62(2), 174-184.

¹¹ Hofman, J. M., et al. (2021). 'Integrating explanation and prediction in computational social science.' Nature 595(7866), 181-188.

¹² Kimberlee Weatherall, Paul Henman, Jose-Miguel Bello y Villarino, Rita Matulionyte, Lyndal Sleep, Melanie Trezise, Jenny Van Der Arend, Scarlet Wilcock, Executive Report - Automated decision-making in NSW: Mapping and analysis of the use of ADM systems by state and local governments (Research Report) (ADM+S, 2024), <<https://www.ombo.nsw.gov.au/Find-a-publication/publications/reports-to-parliament/other-special-reports/a-map-of-automated-decision-making-in-the-nsw-public-sector-a-special-report-to-parliament>>.

Use of ADM

The project found that **use of ADM systems was widespread and increasing** across the state government and local councils, with further **widespread interest in utilising various forms of AI**, including predictive analytics, natural language processing, and generative AI.

Of the 77 responding NSW state government departments and agencies, a total of 136 ADM systems were reported in use, in development, piloting or planned. A majority (60%) of NSW state government departments and agencies used or planned to use ADM systems, reporting a potential increase of 50% in the number of ADM systems planned to be used in the next three years from the number currently reported as 'in use'.

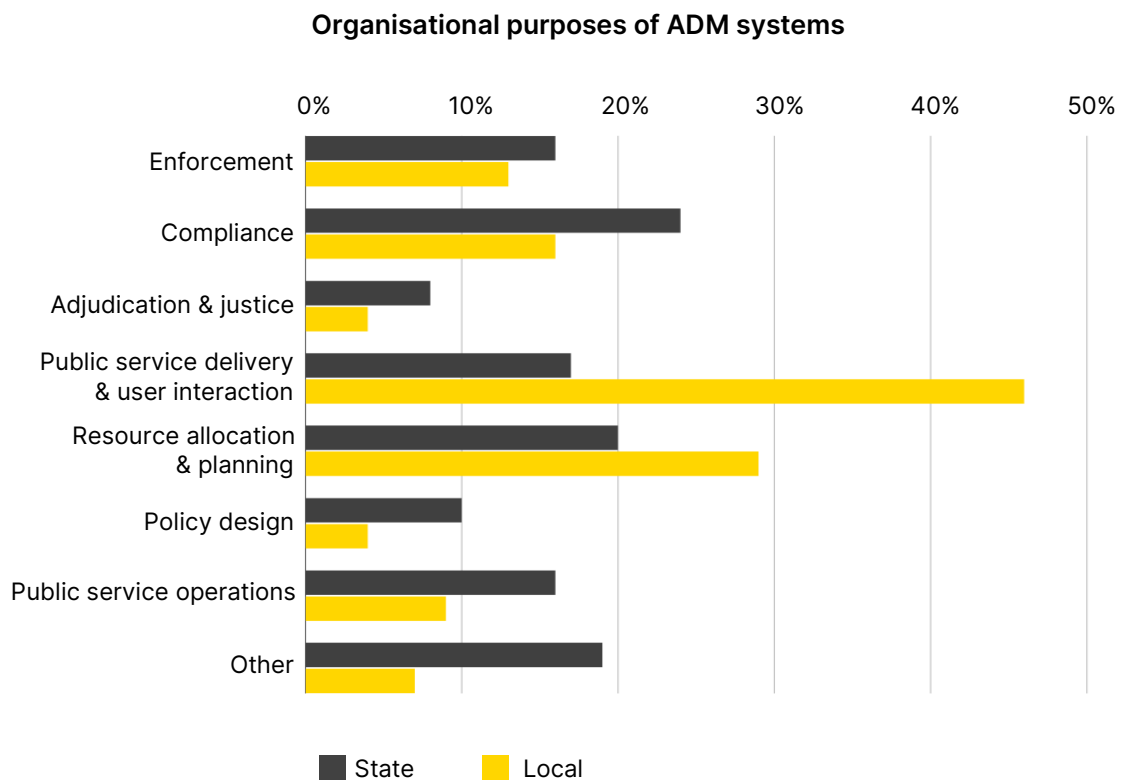
Of the 35 NSW councils that responded to the survey, just under a half (40%) reported use (or planned use) and ADM systems, with a total of 77 such systems reported. Where local councils reported having ADM systems in use, piloting, development or planned future use, these tended to be in metropolitan and city areas with no concrete systems reported in rural councils. Several local councils reported use of computer vision and automated sensors for a variety of purposes, ranging from licence plate recognition through to more advanced uses.

Purposes of ADM

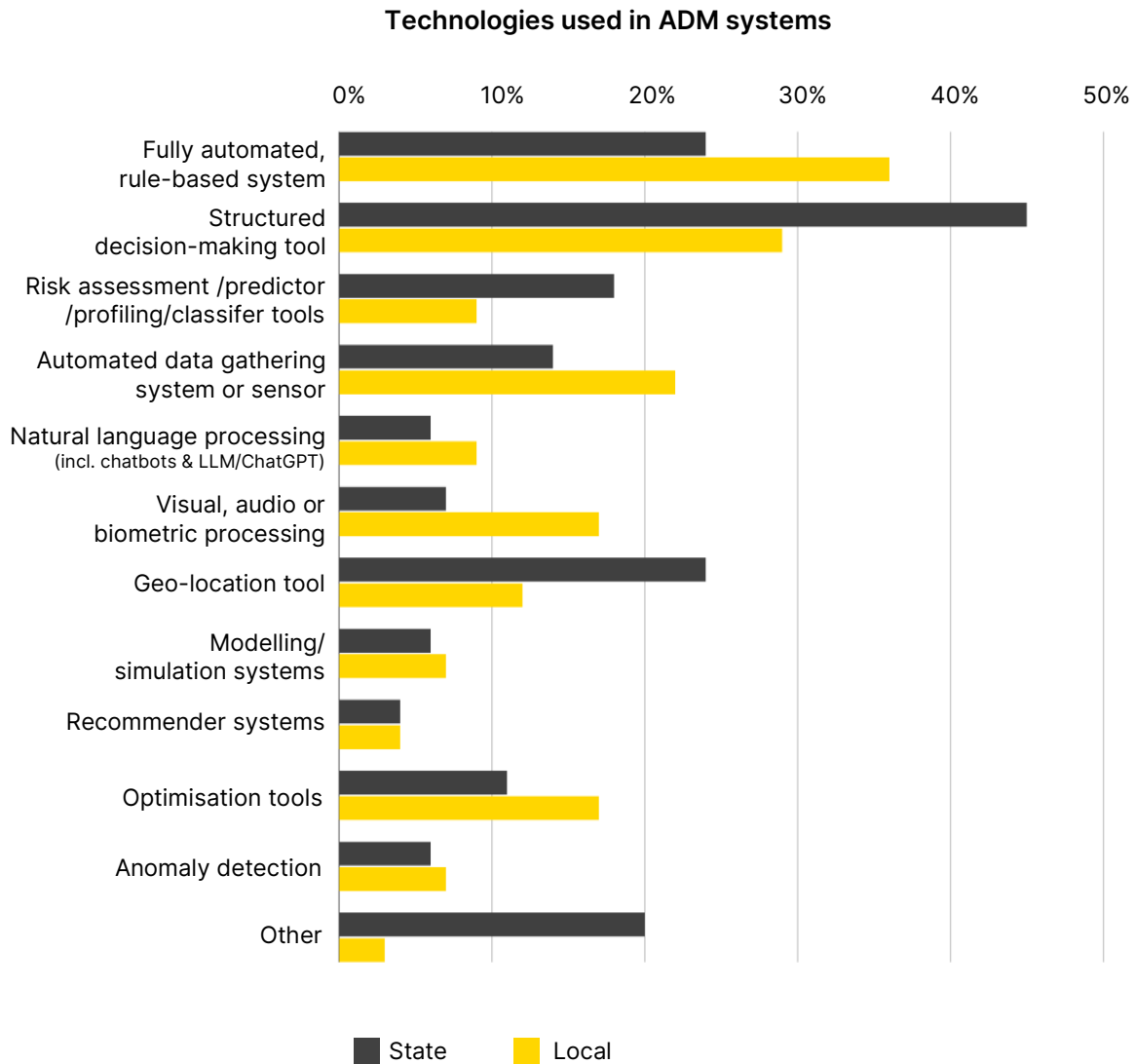
ADM systems are (or planned to be) used for a wide range of **purposes**. Our survey identified the following purposes:

- **Enforcement:** including identifying infringers and sending notices; licence or permit termination; preliminary assessment of possible infringements; application and collection of fines
- **Compliance:** including systems that enable compliance (e.g. systems for applying for/renewing licences and permissions; systems that enable regulated actors to submit information)
- **Adjudication and justice:** tasks that support formal or informal agency adjudication or rights or entitlements
- **Public service delivery and user interaction:** direct provision of services to the public (e.g. delivery of education services to students); chatbots and other automated systems for engagement with, or service delivery to, the public
- **Resource allocation and planning:** using data-driven insights to make operational and resource allocation decisions (e.g. identifying communities to prioritise for street maintenance, policing or public health interventions)
- **Policy design:** monitoring or analysing effectiveness of government actions or policies; profiling or cohort analysis for policy purposes
- **Public service operations:** e.g. procurement; monitoring service delivery and performance; internal fraud detection, and
- **Other:** with respondents invited to 'please describe'.

As the Figure highlights, there are different patterns in the priority purposes for ADM systems between state and local government agencies. We expect that the Federal government agencies would have a different profile. Regardless, dominant purposes across both levels of government are for public service delivery and user interaction, compliance, and resource allocation and planning.



A wide range of technologies are in use, including both AI and non-AI systems. The most common **type** of ADM system reported was structured decision-making, suggesting that ADM systems are being used in collecting, filtering and presenting information, and guiding decision-making, rather than replacing human decision-makers (although as we noted above, with many systems technically capable of further automation, this could change). There was also a sizeable proportion of rule-based automation.



Evidence about ADM/AI use in social services in Australia

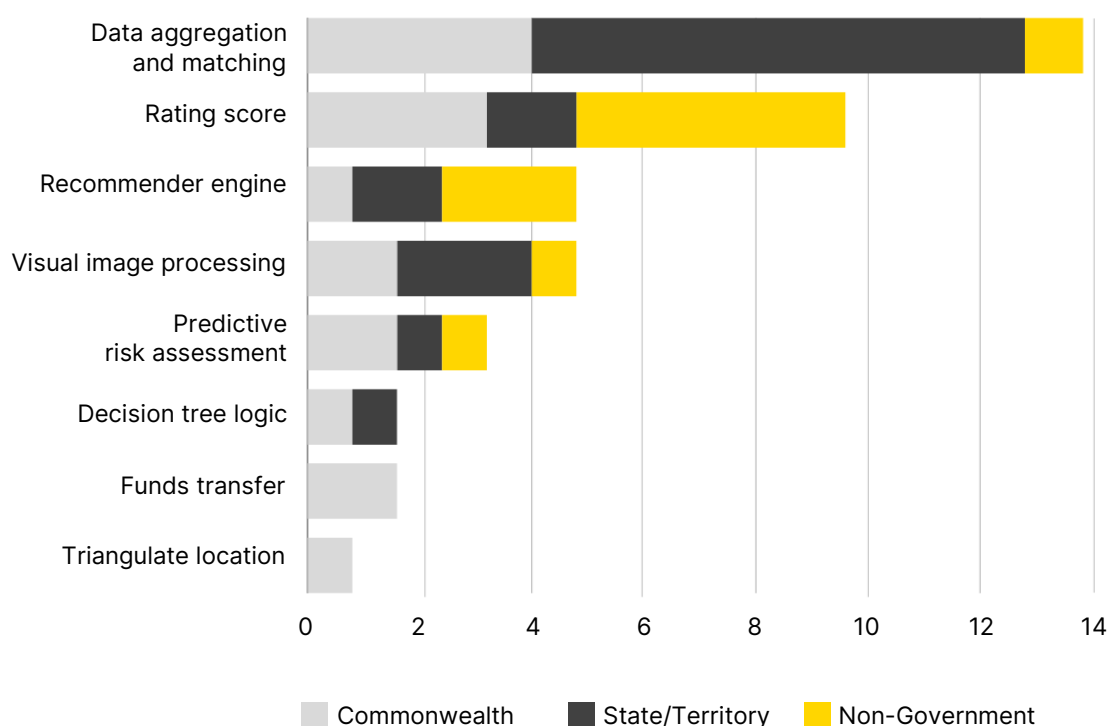
ADM+S also conducted a review in 2021-22 of ADM systems in Australian social services.¹³ A different methodology was used: instead of a systematic survey approach, the research team sought to identify how ADM systems were being used through **publicly available information** in social security and employment services, child protection and family services, disability services, and housing services (a bottom-up case study approach).

¹³ Sleep, L., Coco, B. A., & Henman, P. (2022). Mapping ADM in Australian social services. ADM+S Working Paper 006, DOI: 10.25916/xxj4-n968, <<https://apo.org.au/node/321337>>.

29 ADM systems were identified. Most were deployed by Australian governments, with several in use by the non-government (commercial and non-commercial) sector. Most were based on machine learning technologies (i.e. AI), though for a significant minority, it was unclear if ML was in use or not.

Nine general types of systems were identified in the ADM systems considered in the study: rating/scoring; natural language processing; triangulate location; data-matching; decision tree logic; funds transfer; and information access.¹⁴ A single ADM system may have had more than one function involved in its service delivery. Data aggregation and matching was the most common type of system among those included in the study, followed by rating scores.

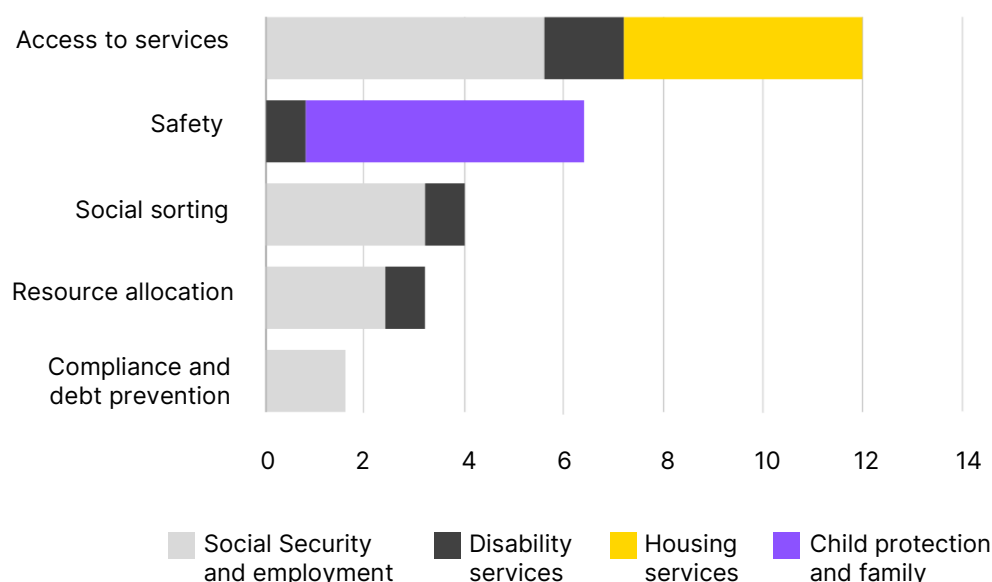
Frequency of function of ADM system in Australia by level of government



¹⁴ According to this study: Natural language processing typically involves ML technology to ‘read’ and/or ‘write’ written or spoken language because written or spoken language is complex and not easily interpreted for meaning by a computer. Visual image processing uses ML to recognise visual images. It has been used in ID verification in the rollout of Digital Drivers Licences in SA and NSW and was used by Services Australia during the 2020 bushfires. Geo-location uses satellite or other signals to calculate the location of an object. In the case of domestic violence monitoring systems, the object is a device strapped to the ankle of a perpetrator. This device can be paired with a hand-held device that is in the possession of the victim-survivor.

Of five different groups of purposes, providing access to services was the most common use of ADM systems, followed by safety, social sorting,¹⁵ resource allocation, and finally compliance and debt prevention.

Frequency of ADM system purpose according to social services sector



The figure below shows the ADM systems' purpose according to whether they are utilised by a Commonwealth, state/territory or non-government agency. Using ADM systems for accessing services was common across jurisdictions.

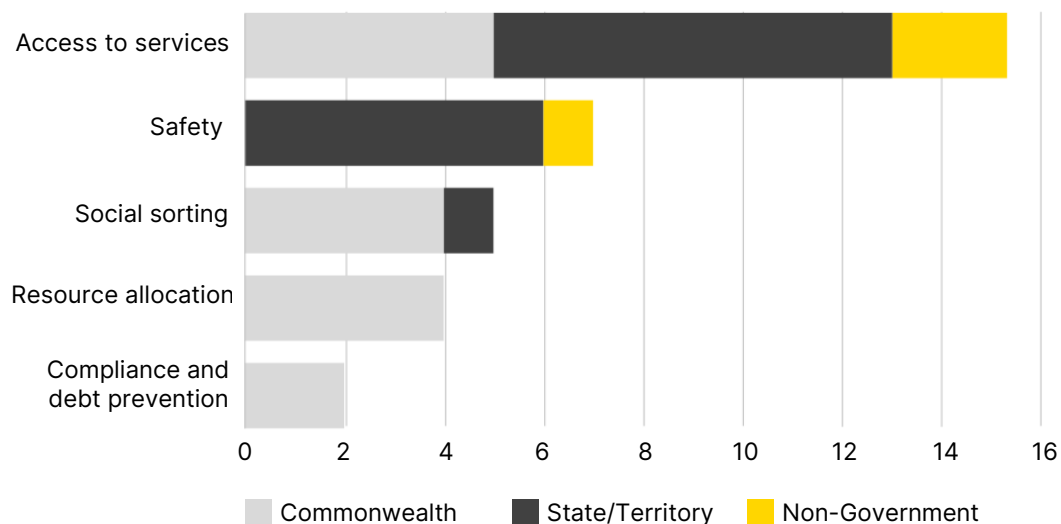
Only the Commonwealth government used ADM systems for compliance and debt reduction and resource allocation. This may reflect the fact that the Commonwealth government does not typically provide services, but funds services delivered by state and territory governments or outsourced. ADM systems were used for social sorting in the Commonwealth government and private sector: the private sector use was part of Commonwealth-funded employment services.

The use of ADM systems for safety only occurred in state/territory governments and the non-government sector, perhaps reflecting their jurisdictional responsibility for safety in policing and child and family services, although it is perhaps worth reflecting

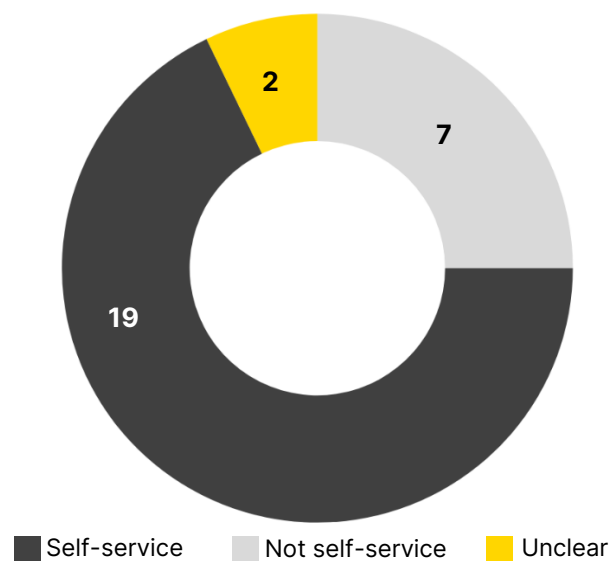
¹⁵ Social sorting refers to the process of automated classification of people into categories or groups, typically by risk assessment or scoring systems: Lyon, D., (ed.) *Surveillance as social sorting: Privacy, risk, and digital discrimination*. Routledge, 2003.

on why the Commonwealth government does not appear to see a role for ADM systems in ensuring or enhancing the safety of its service users.

Frequency of ADM systems purpose according to whether Commonwealth, state/territory or non-government agency



Another key finding was that of the 29 ADM systems included in the study, 19 of them involved **self service** (see below). This pattern indicates a use of digital technology to shift social services in Australia towards user self-services, which is important for a number of reasons including digital access and accessibility for vulnerable service users, since with self-service systems service users become responsible for monitoring systems and inputting data.¹⁶



¹⁶ Casey, S. J. (2022). 'Towards digital dole parole: A review of digital self-service initiatives in Australian employment services', *The Australian Journal of Social Issues*, 57(1), 111-124 <<https://doi.org/10.1002/ajs4.156>>; Bovens, M. & Zouridis, S. (2002). 'From street-level to

A further key finding was that **many of the systems have been developed by companies outside of Australia**, adding challenges for governance and accountability, as well as the appropriateness to Australian conditions.

AI/ADM in public sector beyond Australia

Here we describe evidence from research on other major OECD countries. While these do not directly relate to Australia, they may indicate the likely current and future purposes and uses of AI/ADM in the Australian Federal Government.

The United States of America

In 2020, a systematic review of AI across 142 USA Federal government departments, agencies and sub-agencies was conducted using publicly available data, finding 157 different use cases.¹⁷ At that time, the researchers found that almost half (45%) had experimented with AI and related machine learning (ML) tools.

The researchers classified their governance task into five general categories:

- Regulatory research analysis and monitoring;
- Enforcement;
- Public services and engagement;
- Internal management; and
- Adjudication.

Each AI case might contribute to more than one governance task. Significantly, regulatory research analysis and monitoring was evident in about half of all reviewed AI cases. About a third had enforcement as a governance task, while adjudication was the least of these uses in around 10 percent of cases.

This pattern of governance use type is arguably reflected in the policy areas in which AI is deployed across US Federal agencies. Law enforcement had the most (33), well ahead of health (18), then financial regulation (15), social welfare (13) and commerce (13). Education had the least with only one AI use case detected.¹⁸ In contrast with NSW's ADM systems, the USA's then development of AI was more often in-house.

system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control', *Public Administration Review*, 62(2), 174–184 <<https://doi.org/10.1111/0033-3352.00168>>.

¹⁷ Engstrom, D.F. et al. (2020), 'Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies', NYU School of Law, Public Law Research Paper No. 20-54.

¹⁸ Ibid, Figure 1 (p.17).

Since 2023, the USA Federal government has maintained an AI Use Case Inventory (last updated 1 September 2023), which is mandatory.¹⁹ The current Inventory includes 710 use cases, with the Department of Energy and the Department of Homeland Security having the most use cases (178 and 157 respectively), which are both three times more than that of the next most populated agency (Department of Commerce).

In August 2024 the Office and Management and Budget issued a detailed instruction about reporting for the current year.²⁰ One of its most interesting aspects is the **mandate to agencies to be expansive, noting that, in case of doubt, systems should be included. The 2024 inventory, when published, should be a complete overview of relevant AI uses in the US Federal Government apart from defence and intelligence.**

Europe

An early 2020 survey of AI use in public services in the EU²¹ identified 85 applications. They were found to be most used in the General Public Services policy sector, followed by both Economic Affairs and Public Order and Safety. Education, Defence, and Housing and Community Services were the areas with least use. Of three classified drivers for AI initiatives, performance was overwhelmingly the greatest driver, with inclusion and openness having a limited role. Unsurprisingly, the effects of AI deployment were mostly in improved administration, then improved public services, and a quite minor outcome was improved social value. In 2024, the European Commission published a study on the uses of AI in the public sector across Europe.²² The first part of that study is a scoping exercise to identify the state of play on the uptake and public procurement of AI technologies by the public sector in the EU. The report based on an analysis of current practice, policies and literature identifies four key sectors for AI procurement: eGovernment, mobility, healthcare and education.

United Kingdom

Although the UK has an algorithmic transparency recording system, its voluntary character gives it very limited value.²³ More reliable data on the use of AI was offered

¹⁹ AI.gov, 'The Government is Using AI to Better Serve the Public' <<https://ai.gov/ai-use-cases/>>.

²⁰ The White House, (2024) Guidance for 2024 Agency Artificial Intelligence Reporting Per EO 14110 <<https://www.whitehouse.gov/wp-content/uploads/2024/08/Instructions-for-2024-Agency-AI-Reporting-per-EO-14110.pdf>>.

²¹ Misuraca, G., van Noordt, C., & Boukli, A. (2020) 'The use of AI in public services: Results from a preliminary mapping across the EU.' Proceedings of the 13th international conference on theory and practice of electronic governance.

²² European Commission, (2024) Adopt AI Study: Final study report <<https://ec.europa.eu/newsroom/dae/redirection/document/108555>>

²³ Central Digital and Data Office and Department for Science, Innovation and Technology, (2023) Algorithmic Transparency Recording Standard Hub

by the UK National Audit Office survey of AI use across government with data from September-November 2023 and published in 2024 about 87 government bodies (98% response rate).²⁴ The results were evenly distributed among respondent bodies between those actively using AI (37%) and those taking initial steps to use AI (37%, of which 25% were piloting and 11% planning). **The data is relatively similar to our findings for NSW agencies and departments.**

In terms of uses, the leading current use of AI systems is to support operational decision-making, closely followed by systems supporting research and monitoring and to improve internal processes. The latter is the bigger target for new systems with 87 systems piloted or planned with this function. At the other end, only 19 of the current systems directly provide a public service or engage with the public, with 30 more planned or piloted in the pipeline.

In terms of objectives of the current systems, the usual reasons apply. The report identified as the leading reasons: improved service quality, a concern for roughly 90% of the bodies; and cost reduction and improved service speed (roughly two thirds).

A UK official entity, the Central Digital and Data Office (CDDO), also generated around the same period a very rough study on possible productivity gains from large-scale adoption of AI across the civil service and wider public sector.²⁵ This study estimated the productivity gains to be in the billions and considered that close to a third of tasks in the civil service could be automated. The study did not include a proper consideration of direct and opportunity costs.

<<https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub>>

²⁴ National Audit Office, (2024) Use of artificial intelligence in government (Report) <<https://www.nao.org.uk/wp-content/uploads/2024/03/use-of-artificial-intelligence-in-government.pdf>>.

²⁵ Cited in National Audit Office, (2024) Use of artificial intelligence in government (Report) at 2.8 <<https://www.nao.org.uk/wp-content/uploads/2024/03/use-of-artificial-intelligence-in-government.pdf>>.

TOR 2: The effectiveness of existing legislative, regulatory and policy frameworks relating to AI

Use of AI is governed by both **existing general** legal, regulatory and policy frameworks, and **emerging specific** frameworks. A combination of law and regulation, and policies and frameworks are in operation, so we address TOR and a later TOR concerning internal framework/policies or additional controls together.

Pre-existing legislative, regulatory and policy frameworks

A range of Commonwealth Acts have recognised computer decision-making since the 1980s. The Commonwealth Ombudsman published an Automated Decision-making: Better Practice Guide in 2007 (updated 2020).²⁶

Nevertheless, existing frameworks **have not been effective** to ensure that automated decision-making (let alone the introduction of AI) occurs legally, or in accordance with good administrative practice and public governance values. Despite numerous efforts to use legal and review mechanisms and multiple red flags,²⁷ existing mechanisms failed to prevent Robodebt, ‘a shameful chapter in the administration of the Commonwealth social security system and a massive failure of public administration’.²⁸ The issuance by governments of debt notices that do not explain how debts were calculated, as occurred in Robodebt, is a continuation of a long-standing approach in Services Australia, and importantly, even after Robodebt, **continues to this day**. The recent issuance of old tax debts also occurred without any explanation of how they were calculated.²⁹

We therefore agree with Commissioner Holmes’ recommendation, in the context of the Royal Commission into the Robodebt Scheme, that the Commonwealth should

²⁶ Commonwealth Ombudsman, [Automated Decision-making: Better Practice Guide](#).

²⁷ Carney, T. (2019). Robo-debt illegality: The seven veils of failed guarantees of the rule of law? *Alternative Law Journal*, 44(1), 4-10. <<https://doi.org/10.1177/1037969X18815913>>.

²⁸ *Prygodicz v Commonwealth of Australia (No 2)* [2021] FCA 634; 173 ALD 278 (Robodebt Representative Action approval of settlement; judgment at Annexure F), at [5].

²⁹ Barrett, J. ‘Robotax: why the ATO’s controversial tax debt clawback scheme deserves media scrutiny’, *The Guardian*, 5 April 2024 <<https://www.theguardian.com/australia-news/2024/apr/05/robotax-why-the-atos-controversial-tax-debt-clawback-scheme-deserves-media-scrutiny>>.

clarify the law relating to the use of automation in administrative decision making,³⁰ and are glad to see the strengthening of oversight mechanisms that focus on systems and administrative practices.³¹

A range of laws require clarification and updating. We address only a subset below, but refer the Committee to earlier submissions by the Centre, including in particular our submission on Safe and Responsible AI from 2023.³²

Transparency and freedom of information

Individuals have a right to request information from government agencies. However, ADM systems are normally developed by third parties, and government agencies often do not hold important information about ADM system development and functioning, or this information is protected as commercial trade secrets. For these reasons individuals have been denied access to information about ADM systems that directly affected their interests.

We note that the NSW government has sought to address some of these challenges by requiring that all contracts with external parties must ensure that the government agency is able to access necessary data and information about data processing in order to fulfil the agency's legal accountability responsibilities. However, the situation may be more complex than can be solved by such a directive.

An illustrative case is *O'Brien v Secretary, Department of Communities and Justice* [2022] NSWCATAD 100 where the applicant sought information from the NSW Department of Communities and Justice about how rental subsidies were calculated. Alongside other documentation, the freedom of information request was determined to include the source code of the software that performed the subsidy calculation. The software system that performed the calculation was procured from Northgate Public Services (NPS) rather than developed in-house. This complicated the transparency issues because accessing source code owned by a commercial operator creates risk of commercial disadvantage, highlighting the need for a more tailored transparency mechanism within public law such as a right to explanation. In addition, the software was provided by NPS as a 'service', which further frustrated transparency obligations.

Northgate's algorithmic calculator operated on computer servers that it controlled, but ran as a nested widget within the Department's software system. This meant that the subsidy decision was still formally made by the Department (using its own software), but the material calculations behind the subsidy decisions were processed externally on servers controlled by third parties. The software-as-a-service arrangement meant

³⁰ Holmes, C., Report of the Royal Commission into the Robodebt Scheme, (July 2023), Recommendation 17.1.

³¹ Oversight Legislation Amendment (Robodebt Royal Commission Response and Other Measures) Bill 2024 (Cth).

³² Weatherall et al, ADM+S Submission to the Safe and Responsible AI in Australia Discussion Paper, August 2023, DOI: 10.25916/catx-q405 <<https://apo.org.au/node/323896>>

that despite the algorithmic calculation occurring beyond the boundaries of the Department, the decision was not deemed to have been contracted out. Were the service provision explicitly outsourced, then the Department would have been obliged to include contractual provisions for an immediate right of access to relevant information.

The software-as-a-service arrangement thus enabled an operational equivalent to contracting out but free from the transparency obligations associated with service delivery. This highlights the need for **more explicit contemplation of how transparency obligations and access to records and documents should be contractually managed**. These issues become even more imperative as software supply chains complexify with the use of newer AI systems, and additional regulatory obligations like periodic audit come online.

We note the Privacy and Other Legislation Amendment Bill 2024 (Cth), which will require that Privacy Policies include information on the kinds of personal information used in computer (and computer-aided) decision-making, and the kinds of decisions where computer and computer-aided decision-making occurs. We note however that such information, provided at a high level in a privacy policy, is unlikely to provide much assistance in understanding the operation of automated or AI computer systems, or their outputs in particular cases.

Administrative law

There are uncertainties around when and how 'decisions' are made when computers are involved, and when review jurisdiction under the Administrative Decisions (Judicial Review) Act 1977 (Cth) is enlivened (enabling an affected person to seek review).³³ There are challenges applying existing mechanisms of administrative law, which tend, albeit not exclusively, to focus on review of specific, individual and final decisions, rather than underlying technical systems.³⁴ This is important, because AI and ADM is increasingly used "upstream" of administrative decision making. For example, classifying citizens/cases into different categories for different services, investigation and so on. This is the case in Australia's use of jobseeker classification, in risk assessment tools in child protection,³⁵ and in employment screening.

³³ Anna Huggins, 'Addressing disconnection: Automated decision-making, administrative law and regulatory reform' (2021) 44(3) UNSWLJ 1048-1077.

³⁴ More systemic mechanisms of oversight exist, including, for example, Ombudsman powers to consider administrative conduct, so it would be incorrect to focus only on ADJR Act-style review. Nevertheless, systems are challenging as the objectives of public accountability.

³⁵ Gillingham, P, & Humphreys, C., (2010) 'Child protection practitioners and decision-making tools: Observations and reflections from the front line.' *British journal of social work* 40(8), 2598-2616; Gillingham, P. (2020) 'The development of algorithmically based decision-making systems in children's protective services: Is administrative data good enough?' *The British Journal of Social Work* 50.2, 565-580.

Privacy and security

ADM+S researchers have expressed concern about the adequacy of current laws and internal policies in ensuring the legal and ethical use of AI by law enforcement authorities. Predictive analytics, facial recognition technologies and other AI-based tools pose significant risks, and the current governance framework does not seem to be sufficient to ensure their legal use. NSW police has explicitly stated they are using facial recognition technology, despite specific regulation around its use.

The problem is demonstrated by the Australian Federal Police's trial of Clearview AI, a facial recognition system trained on billions of photos scraped from online social media. The Information Commissioner found that the use of this facial recognition system by the AFP took place in violation of privacy law.³⁶ Clearview AI was also found to have violated (and continues to violate) privacy laws in Australia, UK and many other jurisdictions.³⁷

AI technology, including facial and biometric recognition have a legitimate role in policing and security. But their use needs to be carefully governed to protect individuals' human rights. Australia lags behind other equivalent jurisdictions in the governance of these technologies in policing and security (as well as other) contexts.

Legislative, regulatory and policy frameworks explicitly covering AI

Australian governments (at the Commonwealth, state and territory levels) have committed to being exemplars in safe and responsible use of AI by adopting a lawful and ethical approach to prioritise the rights, wellbeing and interests of people.³⁸

To date, there has been a proliferation of frameworks, guidance, and activity by governments 'about AI'. But it is not clear to us either

(a) that the standards the Commonwealth government has set for itself do or will in fact establish government as an exemplar, or

³⁶ Australian Government Office of the Australian Information Commissioner, 'AFP ordered to strengthen privacy governance', 16 December 2021 <<https://www.oaic.gov.au/news/media-centre/afp-ordered-to-strengthen-privacy-governance>>.

³⁷ Matulionyte, R. 'Australia's privacy regulator just dropped its case against 'troubling' facial recognition company Clearview AI. Now what?', The Conversation, 22 August 2024 <<https://theconversation.com/australias-privacy-regulator-just-dropped-its-case-against-troubling-facial-recognition-company-clearview-ai-now-what-237231>>.

³⁸ Australian Government, National Framework for the assurance of artificial intelligence in government, 21 June 2024. This is a joint approach agreed and released by the Data and Digital Ministers Meeting of that date.

(b) that proliferating guidelines and frameworks provide consistent, coherent, clear rules for AI development, deployment or use in the public sector.

Government *should* be an exemplar of responsible, safe AI use. Governments hold powers to affect the lives of citizens, residents and visitors, including unique powers of punishment. It is not possible for people to ‘opt out’ of a relationship with government. It is for these reasons that governments are subject to more stringent obligations to protect and promote human rights, including specifically in relation to AI under the only existing treaty on AI use to date, the Council of Europe’s Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.³⁹ That treaty differentiates between government and non-government use of AI, imposing higher standards on the former. Australia participated in the drafting of the Framework Convention, but is not (yet) a signatory.⁴⁰

Australian governments are not yet exemplars in the standards they impose on themselves. Frameworks and policies issued by Commonwealth agencies and other governments in recent times have significant weaknesses or problems. As will be evident from the below, the proliferation of guidelines and frameworks, issued by different agencies within government, based on different groundings and emphasising different features is likely to create confusion, inconsistency, and incoherence.

National Framework for the Assurance of AI in Government

The National Framework for the Assurance of AI in Government,⁴¹ has the significant **strength** of being agreed to by the Data and Digital Ministers from the Commonwealth government, and all the States and Territories. **However**, it has no binding force — it is at most a statement of intent. It is also weakened by its grounding in Australia’s dated, pre-generative AI AI Ethics Principles, rather than the newer risk-based regulatory approaches that represent emerging global best practice, including as endorsed by the Australian government itself when it signed the Bletchley Declaration in 2023.⁴² Risk-based regulation is mentioned in the Framework, but the framework itself does not set out a basis for a risk-based approach.

³⁹ [Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law](#) (CETS No. 225). The treaty was signed on 5 September 2024 at Vilnius, by the EU, US, Israel, UK and six other non-EU European countries.

⁴⁰ Bello y Villarino, J-M., Vijayarasa, R., & Weatherall, K., ‘A first step on the long road to global AI regulation’, *The Interpreter* (10 September 2024) <<https://www.lowyinstitute.org/the-interpreter/first-step-long-road-global-ai-regulation>>.

⁴¹ National Framework for the assurance of artificial intelligence in government, 21 June 2024, <<https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government>>.

⁴² The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023, <<https://www.industry.gov.au/publications/bletchley-declaration-countries-attending-ai-safety-summit-1-2-november-2023>>

Policy for the responsible use of AI in government

The Digital Transformation Agency's (Cth) Policy for the responsible use of AI in government appears to be a step forward because it is **binding**,⁴³ and requires the appointment of accountable officials (by 30 November) and a public statement on the entity's 'approach to AI adoption and use' by February 2025.⁴⁴

Unfortunately, the Policy introduces a new three-part language framework that is not aligned with any of the Australia's AI Ethics Principles, the National Framework or the proposed Mandatory Guardrails.

And the policy is **extraordinarily limited in what it requires**. The only mandatory elements relate to accountable officials and transparency statements. Many aspects of governance that look like commitments in the National Assurance Framework are largely voluntary in the Policy. Training staff about AI is "strongly recommended". Agencies are told they **should** "consider ... understanding where and how AI is being used within agencies and developing an internal register with this information". In other words, determining what AI is being used within an agency is not a requirement of the policy.

The policy is markedly weaker than counterparts in, for example, Canada where structured risk assessment processes are mandatory.⁴⁵ There are stronger requirements in Australian state policy, including NSW Government's AI Policy and AI Assurance Framework which uses mandatory considerations to enshrine Ethical Principles.⁴⁶

Under the Commonwealth policy, accountable officials (and there could be many, reducing line of sight and effective accountability) are only tasked with implementing the mostly voluntary policy, notifying the DTA of new (but not existing) high risk use cases, and being a point of contact for whole of government AI questions.

The DTA's standard for transparency is also extremely basic: it applies only to AI (not ADM), and does not provide specific disclosure requirements, only a general

⁴³ With the exception of entities in the defence portfolio and the national intelligence community (the Office of National Intelligence, the Australian Signals Directorate, the Australian Security Intelligence Organisation, the Australian Secret Intelligence Service, the Australian Geospatial Intelligence Organisation, the Defence Intelligence Organisation, the Australian Criminal Intelligence Commission, and the intelligence role and functions of AUSTRAC, the Australian Federal Police, the Department of Home Affairs and the Department of Defence.

⁴⁴ Elaborated in the Digital Transformation Agency's Standard for AI transparency statements.

⁴⁵ Government of Canada, [Directive on Automated Decision-Making- Canada.ca](#) (2023); The White House, [Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence | The White House](#) (October 30 2023).

⁴⁶ Digital NSW, [Artificial Intelligence Ethics Policy | Digital NSW](#); Digital NSW, [NSW Artificial Intelligence Assessment Framework](#) (2024).

statement. Compare this to the US, where since 2023, the USA Federal government has maintained an AI Use Case Inventory (last updated 1 September 2023), which is mandatory. The current Inventory includes 710 use cases. Guidelines for the 2024 update suggest that it will be even more comprehensive.

Technology Protection Framework

We note also that as part of passing the Workforce Australia reforms within the Morrison government, there was a commitment to introduce a Technology Protection Framework (TPF) as part of regulations to ensure that decisions made under the Act would be consistent with a range of principles, including Human Rights and anti-discrimination. It is not clear whether this has been established. The 2022 Parliamentary Select Committee on Workforce Australia Employment Services Rebuilding Employment Services report also discussed, and recommended, a Digital Protection Framework.⁴⁷ The intention would be that this would cover the use of AI, but if it only relates to decisions under the Act, then the use of AI and ADM “upstream” of decisions would continue to be unregulated.

Other frameworks and guidelines

In addition to the National Framework and Policy discussed above, different parts of government have produced a wide range of guidelines, frameworks and other documents:

- Cybersecurity: the Australian Signals Directorate has guidance on Engaging with Artificial Intelligence (2024) and Deploying AI Systems Securely (2024)
- Privacy: the Office of the Australian Information Commissioner (OAIC) has just released Guidance on Privacy and Developing and Training Generative AI Models⁴⁸ and Guidance on privacy and the use of commercially available AI products;⁴⁹

⁴⁷ Parliament of Australia, ‘Chapter 10 - Service provision and program design: Digital or hybrid services’,

<https://www.aph.gov.au/Parliamentary_Business/Committees/House/Former_Committees/Workforce_Australia_Employment_Services/WorkforceAustralia/Report/Chapter_10_-_Service_provision_and_program_design_Digital_or_hybrid_services#_ftn6>

⁴⁸ OAIC Guidance on Privacy and Developing and Training Generative AI Models

<<https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/guidance-on-privacy-and-developing-and-training-generative-ai-models>>

Department of Industry, Science and Resources, Proposals Paper for introducing mandatory guardrails for AI in high-risk settings (5 September 2024). <<https://consult.industry.gov.au/ai-mandatory-guardrails>>

⁴⁹ OAIC Guidance on privacy and the use of commercially available AI products

<<https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/guidance-on-privacy-and-the-use-of-commercially-available-ai-products>>. The

Office of the Victorian Information Commissioner has also issued resources on Artificial

- The Office of the Victorian Information Commissioner has issued resources on Artificial Intelligence – Understanding Privacy Obligations and Artificial Intelligence and Privacy – Issues and Challenges
- The Department of Education have released Australian Framework for Generative Artificial Intelligence (AI) in Schools⁵⁰

The number of often slightly different guidelines, recommendations, frameworks, and statements is overwhelming. A common baseline — and one markedly stronger than the current Commonwealth policy — is needed.

The move to mandated guardrails

In September 2024, the Department of Industry, Science and Resources published a Proposals Paper setting out proposed Mandatory Guardrails for AI in High Risk Settings.⁵¹ While AI is used in a wide range of ways in the public sector (as set out above), a number of settings recognised internationally and in the Proposals Paper as potentially high risk are public sector uses: such as uses affecting people’s access to and progress in education; health; justice and law enforcement; allocation and management of social services; disability and related services. ADM+S has written an extended submission to the Proposal which is available online.⁵²

In summary, and as relevant to this Committee:

- **We support the enactment of horizontal, generally applicable, mandatory guardrails, along with mechanisms for updating the guardrails over time.**
- We think this is **especially important** across the public sector (since government is bound by higher standards and ought to be an exemplar of safe and responsible use).
- The guardrails proposed — requirements such as risk management, testing, transparency, accountability, and data governance — represent **elements of good governance and process that are appropriate across a wide range of different policy domains and different kinds of organisations**. They are suitable for application across the Commonwealth government.
- A coordinating body with power to impose and monitor requirements across the Commonwealth government is required.

Intelligence – Understanding Privacy Obligations and Artificial Intelligence and Privacy – Issues and Challenges

⁵⁰ The Department of Education, Australian Framework for Generative Artificial Intelligence (AI) in Schools <<https://www.education.gov.au/schooling/resources/australian-framework-generative-artificial-intelligence-ai-schools>>

⁵¹ Department of Industry, Science and Resources, Proposals Paper for introducing mandatory guardrails for AI in high-risk settings (5 September 2024). <<https://consult.industry.gov.au/ai-mandatory-guardrails>>

⁵² K Weatherall, H Fraser and A Snoswell, Safe and responsible AI in Australia: proposals paper for introducing mandatory guardrails for AI in high-risk settings, 4 October 2024, DOI: 10.60836/mrgn-9e28, <<https://apo.org.au/node/328568>>.

As we noted in our earlier submission:⁵³

“Addressing issues around safe and responsible AI is a whole-of-government endeavour: it would be a problem to isolate **all** expertise and responsibility in a single entity, and there is much good expertise in various parts of the Commonwealth government. The Commonwealth government has already moved to appoint accountable officials, who can take on a larger role than currently envisaged.

But equally, government action, especially at the Commonwealth level at present **lacks coherence**: as Weatherall has explained in some detail elsewhere.⁵⁴ One reason for this lack of coherence could be the absence of any entity within the Commonwealth Government with the resources to develop coherent, technically- and socio-technically well-informed and hence authoritative positions, and the power, on the basis of those positions to impose stronger obligations, and/or the willingness to use that power.⁵⁵

The imperatives identified above — the need to act; the need to develop processes and powers to update and iterate — should drive institutional and regulatory design. This could require an AI Body (Commission, Regulator). Such an entity would coordinate with agencies and regulators across government, but to be most effective, would need to be the ‘final word’ on safe and responsible AI, with authority across government uses of AI. ...”

As we note there, a **possible explanation for the current weakness of Commonwealth policy on public sector AI use is the absence of any entity within the Commonwealth Government with the power to impose stronger requirements around the use of ADM/AI, and/or the absence of willingness to use that power.** The Digital Transformation Agency is an advisory body: it doesn’t have the power to direct people how to use AI. Oversight mechanisms such as the Commonwealth Ombudsman have power to investigate and review, but not the power to direct agencies in their approach to the use of AI. We are aware that an AI in Government Taskforce is composed of 18 secondees from 11 APS agencies who are working on whole-of-government AI application, policies, standards and guidance.

By contrast, stronger systems addressing public sector use of AI internationally are based not on we understand that the Canadian Directive is based on specific powers held by Treasury; in the US the President’s Executive Order is based on the President’s power over the US federal executive.

⁵³ Ibid at 47-48.

⁵⁴ Kimberlee Weatherall, ‘The Mandatory Guardrails Proposal Paper: An Opinionated Explainer’, 25 September 2024.

⁵⁵ Compare Canada, where the Automated Decision-Making Directive is based on specific powers held by Treasury; in the US the President’s Executive Order is based on the President’s Power over the US federal executive. At the Commonwealth level, the Digital Transformation Agency is an advisory body: it doesn’t have the power to direct people how to use AI.

If this is correct, that there is no entity in the Commonwealth Government with the legal power, the expertise *and* the willingness to mandate consistent governance principles and systems for the development, deployment and use of AI, then legislative action may be required to address the current incoherence and inconsistency. Legislation for the Mandatory Guardrails represents an opportunity to address this gap.

TOR4: Internal framework, policies or additional controls used for assessing risks associated with the use and possible misuse of AI

We have addressed the adequacy and effectiveness of legal and regulatory frameworks applying to, and specifically related to AI above in our discussion of TOR 2. Public sector use of AI is governed by a combination of legislative and regulatory frameworks and associated practices, policies and processes. We have therefore addressed these issues together above.

TOR5: The capacity for line of sight to the output of AI, and assisted decisions, to ensure ethical and responsible use

The limited visibility of ADM and AI systems used in the public sector hinders the public's ability to understand and hold governments accountable for their use of automation and AI use, the government's ability to develop best practice through knowledge-sharing, and the capacity for independent integrity agencies such as Ombudsmans' Offices and Audit bodies to maintain independent oversight.

ADM+S researchers attempting to map ADM use across NSW, found that, despite high level government support for the project, identifying the right people to inform the research was extremely difficult. It was unclear whether there were designated people or teams within these public sector entities tasks identifying new or sufficiently amended systems that may require reporting. Where these positions or teams may have existed, they were not easily identifiable from the outside.

The NSW Ombudsman has specifically stated that this lack of visibility is a key concern, particularly for members of the public whose rights and interests were impacted by decisions made with the use of ADM and for people to be able to be aware of potential rights for review.⁵⁶ They echoed our view that voluntary public reporting will not be sufficient to ensure ongoing comprehensive visibility.⁵⁷

One element of the ADM+S project mapping use of ADM systems in NSW involved reviewing every government department, agency and local council website for references to AI and ADM systems, to see what information is publicly available.⁵⁸ In their special report on the ADM+S project, the NSW Ombudsman's Office noted that the research showed that while information about the use of ADM and AI tools is increasingly available on government agency websites, there were **inconsistencies about what was publicly reported by public bodies, and where and how it is shared**.⁵⁹ They also commented on the insufficiency of voluntary reporting and concerns about insufficient legal input during the design of ADM systems, further noting that many of the systems identified by the project would not be subject to any specific ADM or AI regulatory framework, as they fell outside of the definitions and criteria of the NSW Government AI Assurance Framework.

At a **Commonwealth** level, while the Digital Transformation Agency (Cth) published the Policy for the responsible use of AI in government, which requires the designation of one or more accountable officials, these officials are tasked with implementing mostly voluntary policy; notifying the Digital Transformation Agency of new, but not already existing, high risk use cases; and serving as a point of contact for whole of government AI discussions. We note that under the policy, agencies are told they **should "consider ... understanding where and how AI is being used within agencies and developing an internal register with this information"**. In other words, determining what AI is being used within an agency is not even a mandatory requirement of the policy.

⁵⁶ NSW Ombudsman, A map of automated decision-making in the public sector: A special report under section 31 of the Ombudsman Act 1974, NSW Parliament, 8 March 2024 <https://www.ombo.nsw.gov.au/__data/assets/pdf_file/0004/145093/Introduction-and-Brief-Observations.pdf> p. 2-3.

⁵⁷ Ibid p. 7.

⁵⁸ Results of that review are published in the Executive and Research Reports from the project, which are available online at <<https://apo.org.au/node/325901>>

⁵⁹ NSW Ombudsman, A map of automated decision-making in the public sector: A special report under section 31 of the Ombudsman Act 1974, 8 March 2024 <https://www.ombo.nsw.gov.au/__data/assets/pdf_file/0004/145093/Introduction-and-Brief-Observations.pdf> p. 7.

As we have noted above under TOR 1, since 2023, the USA Federal government has maintained an AI Use Case Inventory (last updated 1 September 2023), which is **mandatory**.⁶⁰

TOR6: Internal capabilities of the public sector to effectively adopt and utilise AI

Internal capabilities and public participation

ADM+S is not in a position to comment generally about internal public sector capacities. However, we can make some comments about the important questions which arise when public sector capacity declines and there is reliance on private sector actors in the development and deployment of public sector systems. Earlier in this submission (TOR2) we talked about freedom of information issues. But there are **broader questions about who is making decisions about public sector activities, and how decision-making accords with democratic principles and public sector values.**

Internationally, there is growing focus on the implications of public sector AI/ADM for democratic participation and decision-making. Research is exploring ways to enhance trust in public sector AI/ADM systems by involving citizens in their design and implementation.⁶¹ This reflects a broader recognition that meaningful public participation is essential, not only for improving the effectiveness of AI/ADM systems but also for ensuring they reflect the needs and priorities of diverse communities.

AI and ADM systems are recognised as posing significant challenges due to their reliance on external IT expertise, namely “govtech” companies and consultancies, for their design and implementation.⁶² Studies have documented the decline in public

⁶⁰ AI.gov, ‘The Government is Using AI to Better Serve the Public’ <<https://ai.gov/ai-use-cases/>>.

⁶¹ Hintz, A., Dencik, L., Redden, J., Treré, E., Brand, J. & Warne, H. (2022) ‘Civic Participation in the Datafied Society: Towards Democratic Auditing?’ Data Justice Lab, Cardiff University, <https://datajusticelab.org/wp-content/uploads/2022/08/CivicParticipation_DataJusticeLab_Report2022.pdf>.

⁶² van Zoonen, L. (2020). ‘Data governance and citizen participation in the digital welfare state’, *Data & Policy*, 2, e10. doi:10.1017/dap.2020.10; Margetts, H., & Dunleavy, P. (2024). ‘The political economy of digital government: How Silicon Valley firms drove conversion to data science and artificial intelligence in public management’, *Public Money & Management*, 1–11. <<https://doi.org/10.1080/09540962.2024.2389915>>; Dickinson, H., & Yates, S. (2023). ‘From

sector technological expertise, which in Australia and similar countries has led to an increasing dependence on industry partners for the delivery of e-government systems, services and platforms.⁶³ When system development is shaped by procurement processes, design decisions are largely in the hands of private third-party developers.⁶⁴ This can leave communities without a meaningful voice in shaping AI/ADM systems that impact them.

This is an international issue, including for Australia. In 2021 a Federal parliamentary committee inquiring into the capability of the Australian Public Service (APS) reported that an “overreliance on external personnel has hollowed out the ICT skills of the APS workforce and represents an inefficient use of taxpayer money”.⁶⁵ The Committee was “extremely concerned” that private contractors outnumber the APS workforce and that these “ICT contractors are being used for core, ongoing work” of the public sector.⁶⁶ As the APS cedes critical functions to private companies, it risks losing institutional knowledge and the ability to respond effectively to the needs of citizens. At stake is not only the technological capabilities of the public sector but also the APS's role as a custodian of public values and democratic governance.⁶⁷

Research on public participation in policy, including in the emerging field of participatory AI/ADM, indicates that involving the public can unlock the positive potential of AI in the public sector while mitigating associated risks. Engaging affected communities in the design and deliberation processes not only enhances legitimacy and trust in government systems. It can also lead to “concrete improvements to a

external provision to technological outsourcing: lessons for public sector automation from the outsourcing literature’, *Public Management Review*, 25(2), 243-261.

⁶³ Collington, R., (2021) ‘Disrupting the Welfare State? Digitalisation and the Retrenchment of Public Sector Capacity’, *New Political Economy*, doi:10.1080/13563467.2021.1952559; van Toorn, G. (2024) ‘Automating the welfare state: The case of disability benefits and services’ in D. Primrose, R. Loeppky, & R. Chang (eds.), *Routledge handbook of the political economy of health and healthcare*, pp. 259-270; Hinkley, S., (2023) ‘Technology in the public sector and the future of government work’, UC Berkeley Labor Centre, <<https://laborcenter.berkeley.edu/wp-content/uploads/2023/01/Technology-in-the-public-sector-and-the-future-of-government-work.pdf>>.

⁶⁴ Mulligan, D. K., & Bamberger, K. A. (2019). ‘Procurement as policy: Administrative process for machine learning’, *Berkeley Technology Law Journal*, 34, 773.

⁶⁵ Australian Government (2021) APS Inc: undermining public sector capability and performance: the current capability of the Australian Public Service, Finance and Public Administration References Committee, November 2021, p. 76 <https://parlinfo.aph.gov.au/parlInfo/download/committees/reportsen/024628/toc_pdf/APSInc_underminingpublicsectorcapabilityandperformance.pdf;fileType=application%2Fpdf>.

⁶⁶ Ibid.

⁶⁷ Alston, P (2019). ‘Digital welfare states and human rights - Report of the Special Rapporteur on extreme poverty and human rights’, A/74/493, 11 October, <<https://documents.un.org/doc/undoc/gen/n19/312/13/pdf/n1931213.pdf?token=QPaHP05HlyLUfG7WNQ&fe=true>>.

program, or in the rejection of proposals that community members did not support”,⁶⁸ helping to avoid program failures which can lead to the abandonment of expensive projects.⁶⁹ By prioritising public involvement, agencies can address concerns early, ensuring that AI/ADM systems are not only effective but also aligned with community needs and values.

Enhancing democratic legitimacy and trust in public sector AI/ADM

There is a high road and a low road for government support of public participation in public sector AI/ADM. A low road approach sees public participation as a mere formality—a box to be checked in the design process. It involves limited, tokenistic involvement, where community input is solicited but rarely acted upon, and participation mechanisms are poorly resourced, inaccessible, and exclusionary.

In contrast, the high road approach recognises public participation as essential to the legitimacy and success of public sector AI/ADM systems. It involves meaningful, continuous engagement with communities at each stage—from problem identification and design to implementation and oversight. The high road approach prioritises accessibility, ensuring that diverse voices are included, and requires logistical support such as accessible venues, language services, technology access, and financial assistance for participants, to eliminate barriers to participation. It seeks to learn from impacted communities while building capacity, providing the education, resources, and technical support needed for informed and equal participation. Impacted communities are regarded as partners, rather than consultants, with the power to shape both design and, importantly, deployment.⁷⁰

The high road approach to participatory public sector AI/ADM requires institutional and regulatory support from governments:

“It is not enough for public participation to be advisory or recommended—or delivered by the benevolence of corporate entities. Rather, hard law requirements mandating public participation and creating enforceable rights to participation are foundational to ensuring

⁶⁸ Gilman, M. E. (2023). ‘Democratizing AI: Principles for meaningful public participation’, Data & Society <https://datasociety.net/wp-content/uploads/2023/09/DS_Democratizing-AI-Public-Participation-Brief_9.2023.pdf>.

⁶⁹ Redden, J., Brand, J., Sander, I., Warne, H., Grant, A., & White, D. (2022). ‘Automating public services: learning from cancelled systems.’ Cardiff, UK: Data Justice Lab, Cardiff University.

⁷⁰ van Toorn, G. (2024). ‘Inclusion interrupted: Lessons from the making of a digital assistant by and for people with disability’, *Government Information Quarterly*, 41(1), 101900.

that decision-makers take participatory requirements seriously and that the public has real influence.”⁷¹

We note that consultation with stakeholders, while an element of the AI Voluntary Safety Standard was not included in the proposed Mandatory Guardrails suggested by the Department of Industry. It may be that there is concern about how to operationalise stakeholder engagement in this area. Considerable work however has been done on this question. Michele Gilman at the School of Law at the University of Baltimore proposes an ideal statute for public participation in AI that incorporates the following key elements:

- mandating public participation;
- defining meaningful participation and the goals of participation;
- describing the public’s role at various stages of the AI life cycle;
- requiring two-way forms of deliberation at specific stages;
- directing culturally appropriate outreach to affected communities;
- identifying obligations to increase access for marginalized communities;
- committing funds for public participation and specialized support on technological issues; and
- setting forth enforcement mechanisms to ensure compliance with public participation requirements.⁷²

At this crucial moment in the evolution of public sector AI/ADM, Australia has the opportunity to lead by adopting a high-road approach centred on trust, participation and democratic legitimacy.

TOR8: Other issues relating to the use of AI in the public sector

Disability and Accessibility

Internationally, there has been considerable interest in the use of AI by public sector entities to improve service delivery and accessibility of government services to people with disability.

⁷¹ Gilman, M. E. (2023). ‘Democratizing AI: Principles for meaningful public participation’, Data & Society. <https://datasociety.net/wp-content/uploads/2023/09/DS_Democratizing-AI-Public-Participation-Brief_9.2023.pdf>.

⁷² Ibid.

In Australia, the efforts of the National Disability Insurance Agency (NDIA) to streamline assessment, provide better information, and personalised service, including the much discussed NADIA chatbot which involved significant consultation and co-design with people with disability.⁷³

A flagship initiative is the National Disability Data Asset (NDAA), aiming to bring together various sources of public sector data for integrated use of data for better service delivery. The NDAA has involved considerable discussion and has generated a notable governance model and principles for how to ensure participation in data developments and their applications (which would include AI use)–for instance, in their draft Charter.⁷⁴

Australia also has the advantage of an early benchmark consideration by the Human Rights Commission with its 2021 final report that laid out a framework for disability rights in the context of AI. These recommendations have not yet been adequately implemented. On top of these have come the evidence and additional recommendations relevant to ADM, AI, and digital inclusion from the ‘Robodebt’ Royal Commission and the Disability Royal Commission.

In the interactions of the public sector with citizens with disability, AI provides an important opportunity to revisit and provide a step-change in the longstanding and languishing area of accessibility and digital inclusion for people with disability. This is an area where state governments especially are undertaking consultation and policy development. For instance, this heightened interest in AI, disability, and digital inclusion is evident in the NSW Government’s work in developing its next 4-year NSW Disability Inclusion Plan (2025-2029).

The implications of disability for AI, especially in relation to ADM, have been a signature, cross-cutting area of ADM+S work.⁷⁵ Disability features in ADM+S’s 2 major

⁷³ Park, S., & Humphry, J. (2019). ‘Exclusion by design: intersections of social, digital and data exclusion’, *Information, Communication & Society*, 22(7), 934–953 <<https://doi.org/10.1080/1369118X.2019.1606266>>; Van Toorn, G. (2024). Inclusion interrupted: Lessons from the making of a digital assistant by and for people with disability. *Government Information Quarterly*, 41(1), 101900. <<https://doi.org/10.1016/j.giq.2023.101900>>; Henman, P.W.F., Dai, D., Borg, S.J., Hummell, E., Foster, M., & Fisher, K.R. (2024). ‘Digitally Networked Social Services: Mapping the National Disability Insurance Scheme (NDIS) online network in Queensland, Australia’, *Journal of Social Policy*, 53(3), 829–853 <<https://doi.org/10.1017/S0047279422000691>>.

⁷⁴ National Disability Data Asset (2024, April). The National Disability Data Asset Draft Charter. Department of Social Services <<https://www.ndda.gov.au/about/the-national-disability-data-asset-charter/>>.

⁷⁵ For example: Henman, P. (2019). ‘Of algorithms, Apps and advice: digital social policy and service delivery’, *Journal of Asian Public Policy*, 12(1), 71–89 <<https://doi.org/10.1080/17516234.2018.1495885>>; Van Toorn, G. (2024). ‘Automating the welfare state: The case of disability benefits and services’ in D. Primrose, R. Loeppky, & R.

reports on ADM in government services.⁷⁶ The issues are also the subject of a report on Disability and Digital Citizenship in Australia, undertaken by researchers from Western Sydney University and University of Sydney, in conjunction with Australian Communications Consumer Action Network (ACCAN).⁷⁷ From this body of work, especially the 2 reports, we would note three key issues to be addressed: accessibility implications; bias and discrimination; potential to improve digital inclusion and accessibility.

Accessibility implications of AI uses: are these being considered?

The requirement that digital technologies be accessible is set out in the 2006 United Nation Convention on the Rights of Persons with Disabilities to which Australia is a signatory. It is set out in various Australian legislation and regulation. In the ADM+S 2024 study of ADM by NSW state and local governments, disability accessibility is the second least used requirement in the assessment of systems. As the report notes

This finding is potentially concerning given the importance of universal design and requirements under disability anti-discrimination to ensure public services are accessible. This low level of testing may be because several ADM systems are internal facing and not directly interacting with the public ... Nevertheless, government employees may also have disabilities, and their ability to use internal-facing systems would be enhanced if accessible design is taken into account.⁷⁸

Chang (eds.), Routledge handbook of the political economy of health and healthcare, (Routledge), pp. 259-270; Van Toorn, G., & Cox, L. (2024). 'Digital citizenship and disability in the Covid era', *New Media & Society*, 26(3), 1249-1267 <<https://doi.org/10.1177/14614448231201650>>; Zhuang, K. V., & Goggin, G. (2024). 'New possibilities or problems for disability and inclusion? The case of AI and ADMs across work', *Telematics and Informatics*, 92, 102156 <<https://doi.org/10.1016/j.tele.2024.102156>>; van Toorn, G., Henman, P., & Soldatić, K. (2024). 'Introduction to the digital welfare state: Contestations, considerations and entanglements', *Journal of Sociology*, 60(3), 507-522 <<https://doi.org/10.1177/14407833241260890>>.

⁷⁶ Ann Coco, B., Henman, P. & Sleep, L. (2022) Mapping ADM in Australian Social Services, ADM+S Working Paper Series (6), ARC Centre of Excellence for Automated Decision-Making and Society, DOI: 10.25916/xxj4-n968; <<https://apo.org.au/node/321337>>; Weatherall et al., ADM in NSW (n 4).

⁷⁷ Goggin, G., Hawkins, W., and Schokman, A., Disability and Digital Citizenship in Australia: Australian Consumers & Citizens with Disability Navigating Digital Society, Institute for Culture and Society, Western Sydney University, Department of Media and Communications, University of Sydney, and Australian Communications Consumer Action Network (ACCAN), October 2024.

⁷⁸ Weatherall et al., ADM in NSW (n 4) p. 69.

Bias and Discrimination: Are These Being Considered?

There are mounting concerns about the implications AI uses raises for disability discrimination, bias, and injustice. Disability has come late to this area, only recently being properly recognised – and the implications of AI for public sector are likely to be significant, especially at all levels (including local and state), and in the face of widespread adoption of commercial systems for areas such as parking, counting citizens, general customer service and citizen interaction platforms, as well as obvious areas of social services, health (not least digital health), and justice.

Here many of the implications of AI use, again particularly by public sector agencies, involves data and its uses in which disability is not evident or 'visible' (e.g. often in health data or financial or tax data), or in the mix with different kinds of intersectionalities,⁷⁹ or what kinds of disability concerns are emerging also with the AI systems and uses developments.⁸⁰ So it is not clear what kind of risk AI use pose for people with disability, and what level of risk is entailed (high or low).

Australia has yet to adopt the kind of frameworks advanced by UNESCO in its AI Ethics Directive or by the EU via its AI Act.⁸¹ These frameworks are generating significant discussion among disability rights groups and disability communities, citizens and civil society, as well as public sector agencies and system providers.⁸²

⁷⁹ Simonsen, L., Steinstø, T., Verne, G., & Bratteteig, T. (2020). "I'm disabled and married to a foreign single mother". Public service chatbot's advice on citizens' complex lives' in Electronic Participation: 12th IFIP WG 8.5 International Conference, ePart 2020, Linköping, Sweden, August 31–September 2, 2020, Proceedings 12, pp. 133-146. Springer; Goggin, G., & Soldatić, K. (2022). 'Automated decision-making, digital inclusion and intersectional disabilities', *New Media & Society*, 24(2), 384–400 <<https://doi.org/10.1177/14614448211063173>>.

⁸⁰ See for example: Packin, N. G. (2021). Disability discrimination using artificial intelligence systems and social scoring: can we disable digital bias? *Journal of International and Comparative Law*, 8(2), 487–512; Special Rapporteur on the rights of persons with disabilities. (2022). A/HRC/49/52: Artificial intelligence and the rights of persons with disabilities - Report of the Special Rapporteur on the rights of persons with disabilities. <<https://www.ohchr.org/en/documents/thematic-reports/ahrc4952-artificial-intelligence-and-rights-persons-disabilities-report>>; Tilmes, N. (2022). 'Disability, fairness, and algorithmic bias in AI recruitment.' *Ethics and Information Technology*, 24(2), 1-13; Mack, K. A., Qadri, R., Denton, R., Kane, S. K., & Bennett, C. L. (2024). "They only care to show us the wheelchair": disability representation in text-to-image AI models.' Proceedings of the CHI Conference on Human Factors in Computing Systems.

⁸¹ Zhuang, K. V. & Goggin, G. (2024). 'Governing AI for disability and inclusion.' Unpublished paper.

⁸² Noori, N., A disability-inclusive Artificial Intelligence Act: A guide to monitor implementation in your country, European Disability Forum, October 2024, <<https://www.edf-feph.org/publications/a-disability-inclusive-artificial-intelligence-act-a-guide-to-monitor-implementation-in-your-country/>>.

At the very least, Australia needs to put a focus on this; and to ensure that implications for disability bias, discrimination, and injustice are routinely considered in assessment of AI uses. This is something that the National Disability Data Asset has looked at, so there is at least one model that can be built upon—to consider a general approach.

Uses of Public Sector AI that promote digital inclusion

Discourses and rationales for public sector use of AI often reference its uses for improving digital in a general sense. However, there is little information available on how AI is being used by public sector entities to improve digital inclusion—and what kinds of policies, processes, initiatives, and resources are involved. Digital inclusion is a flagship area of research at ADM+S.⁸³

There is some activity underway across various agencies (especially in data-intensive areas such as health, welfare, and social services), but it is often supported by consultants without much visibility, planning, or integration. The activity builds in a fragmented way on areas such as improvement of website accessibility and inclusive digital communications—vital areas where the COVID-19 pandemic showed where opportunity and shortfall both lie.⁸⁴

So this is an area where the Australian public sector has a major opportunity—and also imperative— to take a much more ambitious, systematic, and human rights centred policy & design approach.⁸⁵

⁸³ See our work on the Australian Digital Inclusion Index, available at <<https://www.digitalinclusionindex.org.au/>>.

⁸⁴ Ellis, K., Kent, M., Locke, K., McRae, L., Dau, D., & Peaty, G. (2021). 'Smartphones, disability and the Australian experience of the COVID-19 pandemic for people who are Blind and with low vision', *Disability Studies Quarterly*, 41(3) <<https://doi.org/10.18061/dsq.v41i3.8318>>; Goggin, G., & Ellis, K. (2020). 'Disability, communication, and life itself in the COVID-19 pandemic', *Health Sociology Review*, 29(2), 168–176 <<https://doi.org/10.1080/14461242.2020.1784020>>; Goggin, G., & Ellis, K. (2022). 'Disability and media education in the COVID-19 pandemic' in Y. Friesem, U. Raman, I. Kanižaj, & G. Choi (eds.), *Routledge handbook of media education futures post-pandemic* (Routledge) pp. 19-27.

⁸⁵ Suzor, N., Dragiewicz, M., Harris, B., Gillett, R., Burgess, J. & Van Geelen, T. (2019), 'Human Rights by Design: The Responsibilities of Social Media Platforms to Address Gender-Based Violence Online', *Policy & Internet*, 11 84-103 <<https://doi.org/10.1002/poi3.185>>; Yeung, K., Howes, A., and Pogrebna, G. 'AI Governance by Human Rights–Centered Design, Deliberation, and Oversight: An End to Ethics Washing', in M.D. Dubber, F. Pasquale, and S. Das (eds.), *The Oxford Handbook of Ethics of AI* (2020), <<https://doi.org/10.1093/oxfordhb/9780190067397.013.5>>; Henman, P. (2020). 'Improving public services using artificial intelligence: possibilities, pitfalls, governance', *Asia Pacific Journal of Public Administration*, 42(4), 209–221 <<https://doi.org/10.1080/23276665.2020.1816188>>; Australian Human Rights Commission

Environmental Impact

The environmental impact of Public Sector use of AI systems and services is an emerging area of public concern that should be considered. **Every stage of the AI lifecycle creates environmental impacts.**⁸⁶ The main stages include the extraction of resources to build new infrastructure such as data centres, servers, and undersea cables; the energy used to power data centres and water used to cool them; the deployment and application of AI in various facilities and on individual devices for more or less environmentally beneficial purposes; and the disposal and waste of infrastructure and devices.⁸⁷ While computation is generally becoming more efficient, the rapid overall expansion and growth in the training and use of AI means that AI's overall environmental impact is growing, with consequent impacts on emissions. These impacts are created and influenced by downstream users including public sector agencies. For example, Sasha Luccioni of Hugging Face and colleagues have suggested that 'the amount of energy needed to power AI now outpaces what renewable energy sources can provide' in the United States.⁸⁸ In addition, larger models generally use more power and water to train and use,⁸⁹ and the impact depends on where computation is sourced.⁹⁰

A key challenge is the ability to track environmental impacts through the whole AI supply chain so that the user can make an assessment about which systems to purchase and how to use them to minimise their environmental impact.

There are steps being taken in the EU towards requiring larger higher impact model developers and deployers to build in ways to track and make transparent carbon emissions and other environmental impacts of their AI systems. Specifically, the EU AI Act requires the European Commission to request the development of standards for all requirements for high-risk and general purpose AI models and specifically 'on reporting and documentation processes to improve AI systems' resource performance, such as reducing the high-risk AI system's consumption of energy and of other

(2021), Final Report: Human Rights and Technology, <<https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology>>.

⁸⁶ Kaack, L.H. et al, (2022) 'Aligning artificial intelligence with climate change mitigation', *Nature Climate Change*, 12 518–527; Ligozat, A-L. et al, (2022) 'Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions', *Sustainability*, 14(9) 5172.

⁸⁷ Luccioni, S. et al, 'The Environmental Impacts of AI--Primer' (2024) Hugging Face; OECD, *Measuring the Environmental Impacts of Artificial Intelligence Compute and Applications: The AI Footprint* (November 2022).

⁸⁸ Luccioni et al (2022) (n 76).

⁸⁹ Varoquaux, G. et al, *Hype, Sustainability, and the Price of the Bigger-is-Better Paradigm in AI* (2024).

⁹⁰ Pasek, A. et al, *The World Wide Web of Carbon: Toward a Relational Footprinting of Information and Communications Technology's Climate Impacts* (2023) 10 *Big Data & Society*: 20539517231158994.

resources during its lifecycle, and on the energy-efficient development of general-purpose AI models'.⁹¹ In addition the EU AI Act encourages the development and adoption of voluntary codes of conduct for other (non high-risk) AI systems for various purposes including 'assessing and minimising the impact of AI systems on environmental sustainability, including as regards energy-efficient programming and techniques for the efficient design, training and use of AI'.⁹² These reporting standards can provide a basis for Governments and public sector agencies to implement AI procurement policies aimed at minimising the environmental impact of the AI systems they use. For example, the Government of France has adopted as one of its principles for AI innovation the concept of 'frugal' or 'sufficient' AI by which they refer to a range of ways to promote 'moderation in the production and consumption of energy and material resources'.⁹³

The Australian Public Service Net Zero Emissions Reporting Framework requires all Commonwealth entities to consistently report their greenhouse gas emissions.⁹⁴ However it is currently difficult to track the significant emissions from energy (and water) use by the data centres required to train and deploy AI systems and AI enabled services. These would generally count as 'scope 3' emissions, that is emissions from supply chain sources rather than those directly controlled by the Commonwealth entity. The Reporting Framework does not currently require reporting of such Scope 3 emissions, but scope 3 emissions from data centres are currently under active consideration for possible future expansions of emissions reporting requirements.⁹⁵ The Australian Government is already requiring data centres hosting government operations to meet the highest energy rating possible as of 2025.⁹⁶

Similarly the proposed implementation plan for the Australian Government's Environmentally Sustainable Procurement Policy⁹⁷ would require information and

⁹¹ EU AI Act Article 40(2); see also article 112(6)

⁹² See EU AI Act Articles 95(2)(b); see also article 112(7). See also Hacker, P. (2024) 'Sustainable AI Regulation' (2024) *Common Market Law Review* 2(61) pp. 345-386.

⁹³ AFNOR SPEC 2314, 'General Framework for frugal AI: An AFNOR SPEC to measure and reduce the environmental impact of AI', (June 2024), *Ministere de la Transition Ecologique et de law cohesion des territoires*.

⁹⁴ Australian Government Department of Finance, Australian Public Service Net Zero Emissions Reporting Framework, 5 July 2024 <<https://www.finance.gov.au/government/climate-action-government-operations/commonwealth-emission-reporting/australian-public-service-net-zero-emissions-reporting-framework>>.

⁹⁵ Ibid.

⁹⁶ See Australian Government Department of Finance, APS Net Zero Emissions by 2030 Roadmap (2024) <<https://www.finance.gov.au/government/climate-action-government-operations/aps-net-zero-emissions-2030>> and Australian Government Department of Climate Change, Energy, the Environment and Water, Data Centres, <<https://www.energy.gov.au/business/equipment-guides/data-centres>>.

⁹⁷ Australian Government Department of Climate Change, Energy, the Environment and Water, Environmentally Sustainable Procurement Policy (July 2024)

communication technology goods to be procured as sustainably as possible from 1 July 2025. It is not yet contemplated that software and AI systems and services be required to be procured sustainably, but it is likely that the environmental credentials of these systems will come under scrutiny in the future.

It is therefore suggested that Australian Public Sector entities prepare for emissions reporting requirements in the near future and likely sustainable procurement policies in the mid future by using their buyer power to ask AI system suppliers to make available information and estimates of the carbon intensity and other environmental impacts of any AI systems and services they are supplying, and seek to purchase the lowest emissions and environmental impact services available.

Trauma-Informed AI in the Public Sector

Internationally and nationally, there have been efforts to deliver more trauma-informed services in the service delivery sector.⁹⁸ This is a response to the overwhelming experience of trauma across populations, the cognitive and behavioural impacts of trauma on individuals and families, and how this colours their interaction with service providers, including public service organisation and bureaucracies. Trauma responses like fight or flight, shutdown, and hypervigilance can occur many years or even decades after the trauma and lead to negative outcomes for individuals when dealing with bureaucratic institutions. A fight response may mean an individual is excluded from a service. A shutdown may mean that essential documents are not provided in time. In addition, digital service delivery has been shown to have the potential to cause significant harm to people already in situations of international disadvantage

<<https://www.dcceew.gov.au/environment/protection/waste/sustainable-procurement/environmentally-sustainable-procurement-policy#dcceew-main>>.

⁹⁸ Substance Abuse and Mental Health Services Administration, Practical Guide for Implementing a Trauma-Informed Approach (PEP23-06-05005) (2023) <<https://store.samhsa.gov/sites/default/files/pep23-06-05-005.pdf>>; National Trauma Transformation Programme, Implementation: A Roadmap for Creating Trauma-Informed and Responsive Change: Guidance for Organisations, Systems and Workforces in Scotland (2023) <<https://www.traumatransformation.scot/implementation/>>; Mental Health Coordinating Council, Trauma Informed Care and Practice: Towards a cultural shift in policy reform across mental health and human services in Australia. A National Direction, Position Paper and Recommendations of the National Trauma-Informed Care and Practice Advisory Working Group (2013) <https://mhcc.org.au/wpcontent/uploads/2018/05/ticp_awg_position_paper_v_44_final_07_11_13-1.pdf>; Huang, L. N., Flatow, R., Biggs, T., Afayee, S., Smith, K., Clark, T., & Blake, M. (2014). SAMHSA's Concept of Trauma and Guidance for a Trauma-Informed Approach <<https://archive.hshsl.umaryland.edu/handle/10713/18559>>; Henderson, C., Everett, M., & Isobel, S. (2018). Trauma-Informed Care and Practice Organisational Toolkit (TICPOT): An Organisational Change Process Resource, Stage 1—Planning and Audit. Mental Health Coordinating Council (MHCC).

and re-traumatise those with a history of trauma.⁹⁹ **It is essential for the use of AI by public service, especially those that deliver social services, to be trauma-informed.**

It is a long-term issue for service provision professionals that the requirements of bureaucratic systems often conflict with their professional code of ethics and the needs of their clients.¹⁰⁰ Similarly, there are tensions between the assessment and evaluation of AI and social service practice principles, including those that use trauma-informed approaches. Concepts predominating the AI assessment literature – like bias, fairness, explainability, trustworthiness, and accuracy, are often in tension with the concepts of trauma-informed service delivery – like empowerment and choice, collaboration, trust and transparency, safety, and intersectionality.¹⁰¹ Removing bias in the data does not necessarily mean the system is safer – particularly when the system is meant to provide an equal playing field. When service users are required to engage with an AI-supported system to access essential services or otherwise become homeless, empowerment and choice are not outcomes.

Future discussions about the use of AI in the public sector include consideration of the impact on people who have experienced trauma at its core. A trauma-informed approach is increasingly becoming the international standard in public service delivery,¹⁰² and there have been moves to develop a trauma-informed assessment tool for AI-assisted services that can provide some guidance on the types of questions that need to be asked and reflections that need to be made towards this end.¹⁰³

⁹⁹ Artificial Intelligence Database (2024) <<https://incidentdatabase.ai/>>.

¹⁰⁰ AASW, [Code of Ethics](#) (2020); AASW, [Practice Standards](#) (2023).

¹⁰¹ Sleep, L., Introducing a Trauma-Informed toolkit for assessment of AI assisted services, (2024) International Evaluation Conference, Melbourne, 20 September.

¹⁰² NHS Education for Scotland, National Trauma Transformation Program (2024) <<https://www.nes.scot.nhs.uk/our-work/trauma-national-trauma-transformation-programme/>>; Homes, A., and G. Grandison, Trauma-informed practice: toolkit, Scottish Government <<https://www.gov.scot/publications/trauma-informed-practice-toolkit-scotland/>>; NHS Education for Scotland, The National Trauma Training Programme (NTTP) (2021) <<https://www.nes.scot.nhs.uk/news/the-national-trauma-training-programme-nttp/>>; Scottish Government, National Trauma Transformation Programme: Trauma-Informed Substance Use Pathfinders – Learning Report (30 November 2023) <<https://www.gov.scot/publications/national-trauma-transformation-programme-trauma-informed-substance-use-pathfinders-learning-report/pages/3/>>.

¹⁰³ Maitra, S., L. Sleep, S. Fey, and P. Henman, [Building a Trauma-Informed Algorithmic Assessment Toolkit](#) (2024) ADM+S Working Paper Series.