

ARC Centre of
Excellence for
**Automated
Decision-Making
and Society**



Submission to the Attorney-General Department Consultation paper on ADM Reform



Authors

Dr José-Miguel Bello y Villarino, University of Sydney
Professor Emeritus Terry Carney, University of Sydney
Prof Kimberlee Weatherall, University of Sydney
Dr Rita Matulionyte, Macquarie University
Prof Julian Thomas, RMIT University
Prof Paul Henman, University of Queensland
Veronica Lenard, University of Sydney

ARC Centre of Excellence for Automated
Decision-Making and Society

15 January 2025

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

Bello y Villarino, J.M., Carney, T., Weatherall, K., Matulionyte, R., Thomas, J., Henman, P., Lenard, V. (2025). Submission to the Attorney-General Department consultation paper on ADM reform. ARC Centre of Excellence for Automated Decision-Making and Society and University of Sydney.

DOI: 10.60836/v2r1-1g50

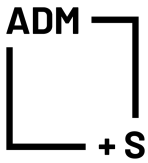
Copyright 2025

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





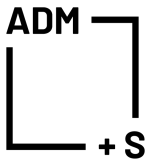
About ADM+S

The ADM+S Centre is pleased to have this opportunity to engage with the Consultation Paper on ADM Reform. 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

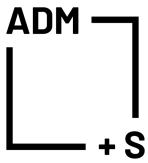
ADM+S commends this initiative of the Attorney General's Department and the reasoning behind the paper. Several members of the Centre have already contributed to the round table discussions organised by the Department and individual members may submit more detailed comments to specific parts of the paper.

The following submission should be read as an invitation to rethink some of the issues discussed in the consultation paper, with a view to achieving a more effective response to the Royal Commission recommendations and, more broadly, a useful regulatory framework for the use of ADM systems in the public sector that balances the protection of rights, reliability, adaptability and a stable regulatory framework without unnecessarily hindering the opportunities that automation offers for advancing social wellbeing and improving the allocation of public resources.



Contents

EXECUTIVE SUMMARY	5
COMMENTS ON THE CONSULTATION PAPER	6
Public interest values	6
Critique: how to achieve these aspirations	6
A focus on automation	7
Critique: ADM in the Australian public sector	8
Applying principles of administrative law and practice to automation	10
Critique: The impact of automation and the need for systemic, preventative measures	10
Critique: Consistency with related reforms, including new privacy rights and the proposed mandatory guardrails for AI in high-risk settings	11
Issues beyond privacy and data collection	11
Critique: We should not ignore the role of Government as an exemplar	12
The regulation of automation cannot be restricted to 'decisions' under the Administrative Decisions (Judicial Review) Act 1977 (ADJR Act)...	13
Critique: The decision to automate	14
Not all ADM systems are the same	16
Critique: Impacts of large-scale automation in administrative functions	16
Balancing transparency requirements with the need to protect sensitive information	17
Critique: Commercial confidentiality concerns may distract us from more pressing issues	17
Ensuring transparency around ADMs	18
Critique: We need to protect sensitive information	20
The right to review	20
Critique: Ex-ante and active monitoring procedures will be vital	20
The limited capacity of the public sector to manage complex automation	22
Critique: Critical engagement with stakeholders	22
Critique: Mechanisms for allocating responsibility for automated systems	22
An Australian response to public sector automation is urgently needed	23
Critique: We can learn from the work of other jurisdictions	23

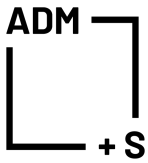


Executive Summary

We are concerned that the Consultation Paper focuses primarily on the reform of legislation and implementation of regulation (Recommendation 17.1), without giving the same level of consideration to enforcement and accountability (covered by Recommendation 17.2 on the establishment of a body to monitor and audit automated decision-making). Improving legislation and regulation is important. However, as both litigation and the Royal Commission revealed, Robodebt was illegal even under existing law: the chief failures in that case were failures of culture, accountability and enforcement.[GU1]

We query whether it is possible to craft an ADM framework that is sufficiently universal in its application without becoming so general and anodyne as to be ineffective in some or many portfolio domains. If a general approach is preferred, as it seems to be the case in the paper, it should be accompanied by provision for any necessary differentiation in the way common policy goals are achieved across different areas of practice. For example, mechanisms for enforcement may require adaptation for particular administrative sectors. Policing monetary fines, determining risk assessment in the health sector, and administering social security benefits may call for different approaches in order to realise the goals of a common set of rules. The framework also needs to recognise that both a commitment to developing a jurisprudence that gives content to the general principles and resources for its implementation will be necessary to ensure they are respected.

Legal reform is still required. We cannot assume that just because existing general administrative law principles are applicable also to automation, the current administrative law and practice are sufficient. Different challenges are posed when we move from a series of individual human decisions to potentially large-scale automated processes of decision-making, where single errors will be multiplied many times over. The shift to rapid decisions at scale demands reform of administrative law, and (as noted above in point 1) innovation innovative methods for and approaches to enforcement.



Comments on the Consultation Paper

Public interest values

We are pleased to see the Consultation Paper recognises competing policy considerations and expresses important democratic and rule of law values that should guide government use of ADM. We think that in noting these values and their application, the paper will play an important normative and educative role for public servants and the public service generally. The importance of this should not be underestimated, as highlighted by the Royal Commission's core emphasis on culture. An ADM framework should contribute to achieving a change in attitudes in society and the public service in particular; and the need to ensure, at all levels, that government designs policies and processes with a primary emphasis on those whom it is meant to serve (Recommendation 10.1).

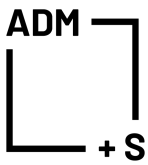
Critique: how to achieve these aspirations

However, the Consultation Paper offers few 'guarantees of protection' against ADM harms in the variety of administrative settings in which they may arise. The research and recent experience in this field show that we cannot assume that general guidance, combined with existing administrative law, will address the novel risks of harm caused by ADM. We think such an approach will fail.

An overview of the existing approaches at the federal level for AI illustrates the limitations of general policies and language. These documents also espouse positive values but lack guarantees of protection against ADM harms. The *National Framework for the Assurance of AI in Government* has been agreed by all Data and Digital Ministers. **However**, it has no binding force. It is a statement of intent, weakened by its grounding in Australia's partially outdated AI Ethics Principles, whose content predates the broader use of generative AI and the development of risk-based regulatory frameworks in Europe, Canada and other countries. Risk-based regulation is mentioned in the Framework, but the framework itself does not set out a basis for a risk-based approach.

The Digital Transformation Agency's *Policy for the responsible use of AI in government* appears to provide more protection because it is binding, requiring both the appointment of accountable officials (by 30 November 2024) and a public statement on the entity's 'approach to AI adoption and use' (by February 2025). However, the Policy is not aligned with the Australia's AI Ethics Principles, the National Framework or the proposed Mandatory Guardrails.

More importantly, the policy's only mandatory elements relate to (1) the appointment of accountable officials, allocating to these officials very limited responsibilities, and (2) the requirement to publish transparency statements outlining the general approach of the agency towards the use of AI. An Australian harmed by the use of AI by an agency is, by these actions, not provided with additional avenues for redress; no standards which can be enforced against the government or argued in any complaint; no significant additional information that would



enable them to contest the use of AI or its application in their particular case; and no information which would enable them to determine whether the system used in making decisions about them is legal, or fair at an individual, or systemic level. To address ADM harms identified by the Robodebt Royal Commission, any framework applied to ADM (including AI-powered ADM) must do more.

Most importantly, as outlined further below, any framework must address the need for effective accountability, provide both more effective individual enforcement and better ways to address systemic issues as envisaged by recommendation 17.2 of the Royal Commission, and must address the systemic nature of ADM harms.

A focus on automation

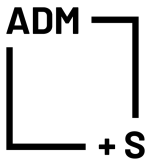
We commend the Paper for its focus on automation of government decision-making, including automation of preliminary steps in decision-making, rather than AI only. Researchers at the ADM+S Centre have long argued for the need to attend to automated decisions, as outlined recently in the consultation for the Joint Audit Committee.¹ Public sector use of new decision-making technologies (including AI and machine learning (ML)) should be understood as part of a broader move to automate government activities. The distinction between automation and AI is neither clear, nor necessarily important. Whether it involves AI or not, public sector automation can significantly affect citizen's rights and good public sector administration.

The Robodebt scheme did not involve AI. Nevertheless, it highlighted the impact of automation on the public; and that systems need not involve AI to be inscrutable and opaque to the general public. Robodebt also demonstrated the challenges for public administrators when designing, managing, applying and governing automated systems; and the difficulties of ensuring accountability for, and remediating harms imposed by systems that are not understood by affected individuals or the wider public.

The governance framework for government ADMs should focus on what governments do with technical systems, not the particular kinds of technical system involved.

An example of an automated government practice is risk profiling. The Australian government has long deployed risk profiling of jobseekers: allocating levels of funding to Job Network providers according to jobseekers' assessed risk of long-term unemployment and conditioning social security payments on the ability or otherwise of the person to meet expectations of those programs. The original Job Seekers Classification Index (JCSI) was developed following standard statistical (multiple regression) analysis coded into the tool. We understand that its replacement, the Job Seeker Assessment Framework, uses machine learning techniques for the

¹ Weatherall, K., Bello y Villarino, J-M., Goggin, G., Goldenfein, J., Henman, P., Matulionyte, R., Parker, C., Sleep, L., & van Toorn, G. (2024). 'Joint Parliamentary Committee of Public Accounts and Audit inquiry into public sector AI use' <<https://apo.org.au/node/328839>>.



same purposes. In both systems, job seeker profiling has very significant potential consequences effects for individuals (loss of social security payments and/or lack of sufficient assistance to re-enter the labour market). And both systems raise other notable issues, such as the opacity of profiling calculations, which have implications for the public at large (the appropriateness or otherwise of employment services settings), for jobseekers (their livelihood) and even for administrators (their understanding of the nature and effects of such tools). Therefore using “ADM” as the regulatory target is, in our view, a preferred approach for government regulation.

Critique: ADM in the Australian public sector

Thanks to recent work, we now know far more about the uses of ADM in Australian governments than we did in the era of Robodebt. At the ADM+S Centre, researchers have conducted the most systematic mapping of the field, working with NSW Government agencies to produce one of the few global examples of a comprehensive analysis. We describe some of the results of that research below.² Much more data, including examples of systems, case studies and more are available in our larger research report. We refer the AGD in particular to the Executive Report of our project mapping ADM use in NSW for a summary of our findings.³

The NSW research directly addresses several key questions which arise in the Consultation Paper.

- The research supports a focus on automation in government rather than AI. However, it also proved evidence of the significant challenges involved in (a) clearly identifying automation, (b) clearly identifying the boundaries of an ADM system, and (c) identifying the point in time when automation occurs.
- These three elements will make it more challenging (or impossible) to simply allocate responsibility to the ‘person’ who ‘controls the system’. There might be many people who can be so described, or there could be significant debate about who precisely that person

2 Weatherall, K., Henman, P., Bello y Villarino, J-M., Matulionyte, R., Sleep, L., & Trezise, M., (2024) ‘Automated decision-making in New South Wales: mapping and analysis of the use of ADM systems by State and Local governments’ <<https://apo.org.au/node/325901>>. See also Sleep, L., Coco, B.A & Henman, P. (2022) ‘Mapping ADM in Australian social services’ <<https://apo.org.au/node/321337>> which analyses over 40 uses of ADM systems in Federal and State social services.

3 Weatherall, K., Bello y Villarino, J-M., Goggin, G., Goldenfein, J., Henman, P., Matulionyte, R., Parker, C., Sleep, L., & van Toorn, G. (2024). ‘Joint Parliamentary Committee of Public Accounts and Audit inquiry into public sector AI use’ <https://apo.org.au/sites/default/files/resource-files/2024-03/apo-nid325901_3.pdf>.

is. Is the 'control' with the person who decides to automate or with the person using the system? Differentiation and uncertainty needs to be accommodated in the Framework if it is to be fully effective. Otherwise, individual responsibility for automation could be severely diluted. A practical application of these issues at the Federal level was described in the research on merits review by Terry Carney,⁴ when he noted the evolution of the Job Seeker Classification Instrument (JSCI), from a questionnaire to an automated assessment system based on that data and trusted by screen-driven civil servants. Is the decision to allow an automation of that assessment the one that involves 'control' (an IT and operational decision) or is it the civil servant who accepts and transmits that computer generated recommendation to the job seeker who should be responsible?

- The research also produced a list of systems identified in use in NSW state governments. The Compendium finally published by the NSW Ombudsman's Office provides a list of 'systems' against which any definition (and any attempt to decide which ones need regulation) could be tested.⁵ The proposed Framework needs to be tested for robustness and coverage against lists such as these if it is to be effective. We also understand that the Australian Digital Transformation Agency conducted a survey of ADM in Federal agencies in 2023. To our knowledge, their findings have not been published, but we believe that careful assessment of those findings and public release of the work underway is likely to enhance the trustworthiness and therefore effectiveness of any future regulatory framework.

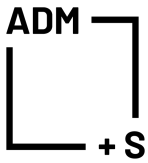
We also encourage the AGD to look to ADM-specific regulation in other comparable jurisdictions, such as Canada.⁶ We draw attention to our explanation of the Canadian Directive in addressing an Australian audience,⁷ where we note that, even with its limitations, the Canadian Directive has many positive attributes that could inform a regulatory framework that adapts as the technology progresses. In particular, for the Australian context, we note that the processes (algorithmic impact assessments) envisaged in the Canadian Directive are compulsory and focus the regulatory burden on those who take the decision to automate (6.1.1); that they establish compulsory monitoring (6.1.3); and are developed from a risk-based

4 Carney, T. (2020). 'Artificial intelligence in welfare: Striking the vulnerability balance?' *Monash University Law Review*, 46(2), 23–51.

5 NSW Ombudsman. (2024). 'Compendium of ADM systems' <<https://www.parliament.nsw.gov.au/tp/files/187691/Compendium-of-ADM-Systems.pdf>>.

6 Government of Canada. (2023). 'Directive on Automated Decision-Making' <<https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592>>.

7 Fraser, H. (2024) 'Decoding Canada's Directive on Automated Decision-Making' <<https://medium.com/automated-decision-making-and-society/decoding-canadas-directive-on-automated-decision-making-08124bcdf250>>.



approach. One key aspect is that they start with an assessment of what the risks are: in general, for the public sector institutions, for individuals and for more vulnerable or particularly affected segments of the population. As the Directive notes, its objective is to ‘reduce [...] risks to clients, federal institutions and Canadian society, and lead [...] to more efficient, accurate, consistent and interpretable decisions made pursuant to Canadian law.’⁸

Applying principles of administrative law and practice to automation

The paper provides a useful and accessible summary of the relevant administrative law principles for ADM. As the paper notes: ‘These principles may be technical in nature, but in general go to “openness, fairness, participation, accountability, consistency, rationality, accessibility of judicial and non-judicial grievance procedures, legality and impartiality”’.

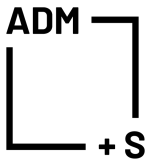
Critique: The impact of automation and the need for systemic, preventative measures

Solutions developed up to this point to implement and enforce the principles of administrative law have relied on individual litigation. Continuing this approach fails to take account of the impacts of large-scale automation, where it is likely that consistency may be guaranteed and rationality apparent, but where individual remedies, especially through court-related procedures, will arrive too late or not guarantee that the issues are addressed – as ultimately they must be – at the systemic level. Furthermore, automation can cause harms that are small in their individual effects – and so perhaps not amenable to individual litigation at all – but with broad impacts which can be large in their net effects across society. Finally, there is the problem of detection: especially where effects are small and seemingly not worth pursuing individually, violations of the administrative law principles may go undetected unless adequate monitoring of both systems outputs and complaints is implemented.

We argue that systemic and preventative measures require the highest priority attention. Individual administrative remedies (such as the Administrative Review Tribunal (ART) in high volume human services/impact areas have miniscule catchment limiting access to justice. The Royal Commission revealed that only 1 in every 2,000 Robodebt decisions got to the AAT for review.

Addressing systemic issues could be done in different ways, but some degree of ex-ante control before deployment and active monitoring once the system is in use is likely to be required. In our view, it is critically important that there should be an entity looking at these

⁸ Government of Canada (n 6) 4.1.



issues in a cross-agency way and developing and applying standards for them. Robodebt Recommendation 17.2 provides the crux of a method for addressing systemic issues. A common, oversight body that monitors the risk management and testing across agencies, and perhaps the trends in disputes in the ART, could address systemic issues in a way that no one now can, because there is no entity or agency with overall responsibility or power.

Critique: Consistency with related reforms, including new privacy rights and the proposed mandatory guardrails for AI in high-risk settings

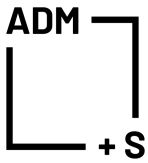
A range of laws require clarification and updating. Such updating should be consistent across the board, first to ensure predictability and reduce compliance costs, and second, and perhaps more importantly, to facilitate the development of a case-based jurisprudence that could inform future decisions in other areas. We refer the AGD to our earlier submissions, including in our submission on Safe and Responsible AI from 2023.⁹ We do not see this submission as the appropriate context to elaborate on these points, but stand ready to refer the AGD to features of that research discussing the interactions between the different legal regimes and noted limitations, if of relevance.

Issues beyond privacy and data collection

Collecting and aggregating data about persons, companies and spaces held by the public sector or by other entities carries inherent and serious risks. The work of the NDIS in this space, to which members of the Centre have contributed, has illustrated several of these potential risks and explored adequate mechanisms to mitigate them.¹⁰ Equally, establishing a 'right to a human decision' as envisaged in the EU GDPR is not a bullet proof protection against such risks. Concretely, as we discovered in NSW, the most common type of ADM system reported was hybrid ADM/human decision-making, where ADM systems are used to collect, filter and present information, which then guides or structures human decision-making. The interaction between humans and machines, the circumstances in which humans deviate from machine-driven recommendations (and the implications of being reluctant to do so) are of particular relevance

⁹ Weatherall et al. (2023). 'ADM+S Submission to the Safe and Responsible AI in Australia Discussion Paper' DOI: 10.25916/catx-q405 <<https://apo.org.au/node/323896>>.

¹⁰ NDIS Research and Evaluation Branch. (2023). 'Terms and conditions for the provision of aggregated or summary data for the purposes of research' <https://dataresearch.ndis.gov.au/media/3775/download?attachment>; Van Toorn, G., Carney, T. 'Decoding the Algorithmic Operations of Australia's National Disability Insurance Scheme' (2024) *Australian Journal of Social Issues* 1-19 ahead of print <https://doi.org/10.1002/ajs4.342>.



here. It would be unwise for the Framework to put undue weight on glib or superficial remedies to complex sources of ADM risk, that depend on many aspects of that interaction, some technical, some psychological, some behavioural. Our research project on this topic could inform the design of the proposed regulatory response.¹¹ We note that considerable research demonstrates automation bias – that is, for humans to defer to machine analysis – in similar settings.¹²

Critique: We should not ignore the role of Government as an exemplar

There is no need to create confusion with many overlapping regimes for automation. We may not need, at the current stage, the basic rules for the public sector to be substantially different from the general rules for AI and automation deployment in Australia. The principles and approaches to make automation safe could be the same for the private and public sector. The proliferation of guidelines and frameworks to date fail to provide consistent, coherent, clear rules for AI (or other types of advanced automation) development, deployment or use. There are too many rules; too many of them are optional or weak; and too many differences between the various guidelines and framework. The goal of consistency, coherence, and building trustworthiness and therefore public trust cannot be met as things stand.

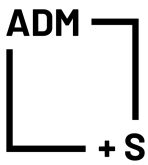
However, there is a public interest in the Government's role as an exemplary user of new technological systems, showing how the key principles and rules can be applied and enforced to a higher standard.

We have previously made submissions in relation to the government's proposed **Mandatory Guardrails for AI**. Our view is that the Mandatory Guardrails offer an opportunity to work towards 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 the government is bound by higher standards and ought to be an exemplar of safe and responsible use. The government has the exclusive authorisation to use force, to take away rights and is the vanguard against destitution.

¹¹ ADM+S. 'Humans, Machines, and Decision Responsibility'
<<https://www.admscentre.org.au/humans-machines-and-decision-responsibility/>>.

¹² Alon-Barkat, S., & Busuioc, M. (2023). Human–AI interactions in public sector decision making: “automation bias” and “selective adherence” to algorithmic advice. *Journal of Public Administration Research and Theory*, 33(1), 153-169.



- 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.

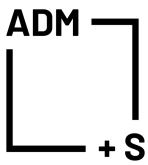
Following international practice (*Council of Europe Framework Convention on Artificial Intelligence and human rights, democracy and the rule of law*),¹³ we strongly believe that if the same rules are developed, the standards for allowing deployment of automation for government (Commonwealth and State) could and should be higher, as the public sector has an enhanced mandate of protection. Furthermore, even within government decision-making it may be necessary that the same general standards are applied in a stricter manner in some contexts. For example, welfare decisions (or decisions about imprisonment) may have such significant human rights impacts, that the strictest standards of monitoring should be applied

The regulation of automation cannot be restricted to ‘decisions’ under the Administrative Decisions (Judicial Review) Act 1977 (ADJR Act)...

Automation of preliminary or preparatory steps along the way to a decision may have such a significant impact on the final decision that such automation should be subjected to the same level of safeguards as if the automation were used to make the final decision. Some automation in the public sector is commonplace, and so routine that people overlook it. For example, some voice to text natural language processing systems (NLP) — such as those used in Amazon’s Alexa — have for many years been in use without being considered as automation or AI. Services Australia has used such systems for automated call routing and identification for over five years. Rules-based automation of administrative processes that progressively reduces human discretion or intervention — such as calculating leave entitlements — is omnipresent without being seen as an administrative automation process. Some of these decisions have limited impact on a final decision, but, for some others, outcomes may determine, or come close to determining final decisions, even those ostensibly made by humans.

The Consultation paper rightly identifies the constraint on access to merits review due to the HCA’s *Bond case* interpretation of the meaning of a ‘decision’ as being only the end product of decision-making processes. It reinforces the case for taking a broader view of what constitutes

¹³ Council of Europe. (2024). ‘Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law’ (Treaty Series No. 225) <<https://rm.coe.int/1680afae3c>>.



a decision in the present setting, because Robodebt AAT review numbers (and statutory internal reviews as a precondition to AAT jurisdiction [see below]) were deliberately depressed by administrative instructions to treat concerns as preliminaries to decisions, rather than the required 'decision' that opened the door to internal and then AAT review.

In the welfare context, only a small proportion of precursors to decisions (algorithms/ADM elements/digital processing of information) will generate sufficiently significant consequences for social security recipients so as to warrant the availability of individual merits review under any ADM protective framework. This creates several levels of complexity.

For example, as we have noted, it would be reasonable to make reviewable the use of the Job Seeker Classification Instrument (*JSC*) to assign people to particular employment services schemes, as it impacts their capacity to comply with JobSeeker payment conditions. On the other hand, the use of profiling tools for prioritising investigations into overpayments or possible fraud arguably may not be reviewable, if the decision remains with the human when several options are offered.¹⁴ Nonetheless, in our view, the lack of access to or appropriateness of individual merits review remedies, reinforces the need to ensure adequate assessment and monitoring of ADM at the systemic level for identification and elimination of unjustified biases, or other undesirable effects.

Critique: The decision to automate

If we focus our regulatory efforts on outputs of an administrative decision-making process with effects outside of government, most existing uses of AI as a technology are likely not to be covered. As explained in our previous submission to the Joint Audit Committee, 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 coded laws to cases (the 'if X, then Y' model). AI/ML operate more like "judgement machines" — making assessments within ambiguity and complexity. As noted in our report on automation in NSW, 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 or entirely autonomously. This means that current uses of AI/ML are likely to be part of a decision-making process, but not necessarily exclusively making decisions. However, this "upstream" automation use presents challenges and requires good governance:

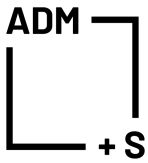
¹⁴ Wilcock, S. (2024). *Policing Welfare Fraud: The Government of Welfare Fraud and Non-Compliance* (Routledge).

- 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;
- Computerisation can progressively reduce administrative discretion, not deliberately or consciously and in such a way that may not be possible to detect. Evidence suggests that where computer 'recommendations' are perceived to be 'right' sufficiently often, a cycle may be created where over time humans trust the system more and more, until a point is reached where they do not sufficiently interrogate or challenge automated suggestions. Going back to the JSCI, research has noted how the percentage of frontline staff who agree with the statement 'Our computer tells me what steps to take with clients/jobseekers and when to take them' was at 17 per cent in 1998 but soared to half of respondents by 2012.¹⁵
- Furthermore, it is possible that officials at higher levels of seniority may be unaware when a higher level of automation in decision-making is adopted, owing to 'creeping changes' in systems. As our research focused on systems in NSW as of 2023 showed, humans were mostly making final decisions for now, but we also found many systems that were **technically capable of autonomous operation with little or no change to the technology being used**. This means officials at a lower level could choose to automate a process or part of a process, without there being a system change or procurement. This may even be minimally noticed, if automation often proceeds along a long-term incremental processes of codification, calculation to automation, with each change seen as mundane and uncontroversial

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**. Any decision to automate a process, and not only the outputs generated by the automation, should be considered as a potential regulatory target. This is already the approach in Canada.

A critical implication of this, and one the Consultation paper passes over, is that pre-deployment audits or co-design safeguards prior to rolling out ADM or AI elements of government administrative decision-making – valuable though such safeguards are – will not be sufficient on their own. The dynamic and often incremental character of the way administrative systems and their associated cultures evolve over time calls for the Framework to incorporate an equivalently dynamic ongoing accountability and audit capacity. Again, future development of the Framework will need to address Recommendation 17.2 of the Robodebt Report.

¹⁵ Considine, M., Nguyen, P., & O'Sullivan, S. (2018) 'New Public Management and the Rule of Economic Incentives: Australian Welfare-to-Work from Job Market Signalling Perspective', 20(8), *Public Management Review* 1186, 1199.



Not all ADM systems are the same

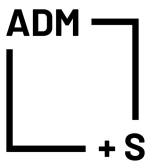
The Consultation Paper clearly explains that not all automation should be treated in the same manner and that whatever regulation is developed should be proportionate, to allow the public sector to take advantage of the benefits of automation and AI. Emergent technologies, such as generative AI, are already being used in mundane, low-risk ways, such as producing transcripts of meetings and editing internal and external communications from drafts, as evidenced in the recent Federal government Microsoft Copilot trial. Chatbots for customer-public agency interaction have rapidly grown, and have accelerated with Generative AI, as have tools assisting public sector employees within agencies for tasks such as in navigating complex legislation or identifying overlaps in patents. But when vulnerable populations are impacted – as for instance in the rightly abandoned development of the Nadia chatbot for NDIS – even seemingly innocuous tools may pose an unacceptable risk for a particular cohort or in particular sectors of government administration.

Critique: Impacts of large-scale automation in administrative functions

The paper notes that '[t]here is no need for legislative authority to automate general administrative functions of a department, such as allocating correspondence for action in a ministerial correspondence system.' However, linking the absence of need for legislative authority to lesser (or no) legal relevance from the automation and its effects is very problematic.

Allocation of correspondence is a form of triage that determines, for example, levels of priority and at which level (or department) matters are addressed. This could give rise to maladministration if it affects some groups in a distinct manner. For example, correspondence about water rights, uses and procedures that is classified and addressed in a preliminary manner automatically, sending some information to the person addressing the correspondence, may generate unnecessary delays or not adequately treat different needs appropriately. Even in single instances, triage of correspondence may be problematic. For example, information about the risk to cultural artifacts of a mining activity that is automatically diverted to a department of mines but not to officials responsible for protecting cultural heritage could change entirely how the matter is addressed.

Even routine uses of ADM and AI require care, attention, and good governance. Chatbots may appear to provide a low-risk example of use of automation, but considerable attention is needed to ensure accurate and safe information is provided. Furthermore, in this example, accuracy should also be accompanied by assessment and monitoring to ensure that the information is appropriate to people's needs. Misleading information at a large scale can have grave consequences or some particularly vulnerable individuals may be particularly affected by the automated system (e.g., cases of traumatic experiences). One human can only make a limited number of mistakes affecting a limited number of people, but an automated system could cause much larger damage and distinctly affect one group. There is now a litany of cases



of human service chatbots causing harm, such as giving advice to people with depression to suicide or to diet to people with anorexia nervosa. These cases do raise issues beyond correct decisions under administrative law.

The key element here is to assess the risk of deploying automation and correctly understanding the setting in which it is deployed. The expert application of chatbots and natural language processing are always very different from large scale citizen-facing deployment. Using 'legislative authority to automate' as a proxy for risk and the need of enhanced checks can be a first step, but it must not be the only one.

Balancing transparency requirements with the need to protect sensitive information

The paper provides a good overview of the risks of full transparency. It identifies four meaningful considerations to be addressed in relation to the recommendations of the Robodebt Royal Commission and the Ombudsman's Better Practice Guide. These bodies recommended that information about automated systems should be publicly available and that business rules and algorithms to be made available to enable independent expert scrutiny.

The paper shows how a simple application of these recommendations could create challenges. We agree on the need for meaningful transparency, not just 'transparency', and the seriousness of increasing risks of adversarial attacks, exploitations of cybersecurity vulnerabilities or, more simply, fraud based on disclosed information. We note the limitations already existing in FOI legislation.

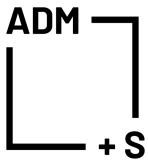
For reasons outlined below, however, in our view considerable caution should be exercised in our estimations of the risks of greater transparency to the intellectual property rights of third parties, and the commercial confidentiality of agreements with them.

Critique: Commercial confidentiality concerns may distract us from more pressing issues

The issue of the need for secrecy for proprietary software algorithms, or 'commercial in confidence information' that could not be reviewed in courts, has been widely covered in the academic literature, based on a series of cases in the United States, notably the COMPAS saga and, concretely, the Loomis case.¹⁶ This problem also arose in NSW in the O'Brien case,¹⁷ where

¹⁶ See a summary here <<https://harvardlawreview.org/print/vol-130/state-v-loomis/>>

¹⁷ [2022] NSWCATAD 1000.



an individual could not get information about the ADMS that denied him a social housing subsidy because the NSW agency did not ‘hold’ the algorithm developed by third party and the algorithmic information was considered commercially sensitive information. Protection of commercially sensitive information should not in all cases prevail over other legitimate interests, such as individual’s right to understand on how the decision affecting their legitimate interests has been made.

However, there is a risk that discussion of commercial confidentiality distracts us from broader problems. For example, focusing on this issue would be of limited use for addressing Robodebt. As far as we know, Robodebt did not involve commercially sensitive information but even so lacked transparency and accountability.

For ADM systems developed in-house, commercial secrets should not be a problem, and transparency should be appropriately prioritised.¹⁸ For ADM systems commissioned from external providers, but fully owned by government, access to confidential information for the purposes of accountability and transparency should be dealt with in contracts. For systems acquired through competitive procedures, it is important that procurement processes incorporate transparency requirements into requests for proposals.

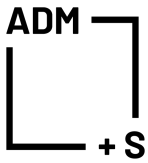
We note that recent research from the UK shows that current procurement practice falls well short of best practice, perhaps because of a lack of capacity in the public sector to appropriately buy sophisticated systems.¹⁹ The ADMS Centre is currently starting a project with the Office of the Information Commissioner in Victoria on this topic. This research could inform future developments in this space. However, for the time being, it is important that some key transparency requirements be incorporated into the acquisition of ADM systems, especially when envisaging decisions that may be challenged in court. Judicial review of individual rights should not be limited by third party claims of ‘commercial in confidence’. Such an approach could be incorporated into regulation.

Ensuring transparency around ADMs

The Consultation Paper discusses three solutions to ensure transparency: publication of general information about ADMs (on agencies’ websites), regular reporting requirements, and

¹⁸ Although we are aware of cases of systems that were developed in-house and then marketed to other users.

¹⁹ Ada Lovelace Institute. (2024). ‘Buying AI: Is the public sector equipped to procure technology in the public interest?’ <<https://www.adalovelaceinstitute.org/wp-content/uploads/2024/09/Ada-Lovelace-Institute-Buying-AI.pdf>>.



publications of business rules and algorithms, as well as limitations and risks of each of these solutions (such as disclosure of sensitive information).

The proposal for agencies to publish general information about their ADM use and how the process works should be welcome. However, it should be clear that such general information would target the general public but will not enable external expert scrutiny. Experts will need access to detailed information to provide assessment of these systems, and a mere plain-language summary will not be sufficient.

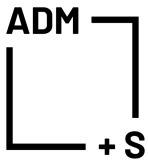
Also, while the FOI Act contains some provisions relevant for the publication of such general information about ADM use, they are not sufficient and detailed guidance is required as to what information agencies should publish about ADMs and how this requirement will be audited and enforced. Our research has shown that NSW GIPA Act has similar information publication requirements and NSW Information Commission previously provided guidance to agencies on what algorithmic information should be published on agencies' websites.²⁰ Despite this, both NSW Ombudsman and our studies have demonstrated that there is almost no publicly available information about ADMS use by NSW agencies.²¹ General FOI Act requirements are therefore likely to be insufficient.

While we agree that agencies should regularly report certain information about the use of ADMS, further detail is needed to understand what is meant by 'regular reporting requirements', as proposed by AGD. Potentially, regular reports about ADMS use could be published along with general information about ADMs, on agencies' websites. A key issue is likely to be definitional, whereby agencies do not appreciate what types of automation meets the threshold for reporting.

The most controversial proposed solution is the publication of business rules and algorithms. Indeed, in order to ensure independent expert scrutiny of government ADM, detailed information about ADM systems and the rules they apply would be needed. At the same time, as the paper has noted, this information may be sensitive from the agency's perspective or considered confidential commercial information by businesses. FOI Act has very broad exceptions in this regard, and, currently, such sensitive information would not be made publicly available for expert review.

²⁰ Matulionyte, R. (2024). 'Government AI, Transparency and Trade Secrets', 47(3) *Melbourne University Law Review*.

²¹ NSW Ombudsman. *The New Machinery of Government: Using Machine Technology in Administrative Decision-Making* (Report, 29 November 2021) 7 ('*The New Machinery of Government*')



Critique: We need to protect sensitive information

To address this conflict between transparency demands and sensitive information, two solutions could be explored:

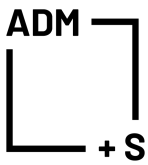
- 1) There is no need to release detailed business rules and algorithms publicly. Rather such information could be provided to external assessors under confidentiality duties. Such external review reports, or at least their summaries, should be made publicly available. Since public members themselves are usually not able to conduct reviews themselves, access to summaries of audits conducted by external independent experts would be much more useful in ensuring public trust than publicly disclosing all information about complex algorithms which public members will anyway not be able to understand.
- 2) Some information, even if commercially sensitive, will need to be disclosed to the public. For instance, even if an ADM developer argues that they cannot disclose criteria/parameters that the algorithm considers when making a decision due to its commercially sensitive nature, if this information is essential for individuals to understand how the decision was made, it should be disclosed despite its potentially sensitive nature. The FOI Act provides very strong protection for sensitive information (including commercial confidential information). The Government might want to consider revising the protection provided as private-public partnerships for developing and using ADMS are increasing, especially where core government services and welfare are involved. We note that the Royal Commission's recommendations were not qualified by reference to commercial confidentiality. More importantly, Government will need to improve their procurement processes to ensure that agencies get access to certain essential information about ADMs they acquire, which they could share with public members who are affected by these systems.
- 3) More generally we note that transparency and the protection of commercial in confidence information are not a binary: it is not the case that either we protect information or disclose it to the world. Increasingly the direction of reform overseas is to allow for qualified transparency, using mechanisms such as those discussed here.

The right to review

The paper very appropriately notes the importance of the right of individuals to challenge a decision after it has been made. In the ADM context a right to merits review, that is a right to have an internal or external reviewer reconsider a decision on its merits, is an essential avenue to address the most egregious mistakes.

Critique: Ex-ante and active monitoring procedures will be vital

The Robodebt experience demonstrated how *inaccessible* the avenues of review are. Prior to the class action and other processes there was no baseline statistic of administrative decision-making error rates. We usually knew how many decisions were made (often in the millions) but no one anywhere ever knew how many of those were erroneous decisions. Therefore there was



no way of contextualising figures about the number of those errors that reached a merits review body – in the case of Robodebt a not entirely ‘dusty’ figure of a few hundred people.

What Robodebt enabled us to discover and reflect on, was that only 1 in every 2,000 unquestionable ‘errors’ ever reached the AAT. More than sobering, such a figure calls into question our faith in principles of access to justice and the adequacy of relying on individual remedies for what may be (and in welfare often are) systemic issues.

Oddly the paper seems to skate around the core Robodebt issue of ‘legality’ of ADM (touched on in reference to ‘coding’ at page 10 and more directly at page 20 in relation to ‘pre-implementation’ stages). In welfare/human services contexts this is no small issue, because mass decision-making necessarily spawns several different levels of granularity of translation of ‘law’ (as understood by lawyers) into ultimate decision-making systems.

The first level of granularity, the publicly available statement of policies attracting ‘Drake #2’ deference at ART review, such as the *Social Security Guide* is a less than perfect distillation of court and tribunal jurisprudence (e.g. on what constitutes ‘special circumstances’).²² Those discrepancies widen further as we drill down into the ‘business rules’ and any ADM coding of those rules.

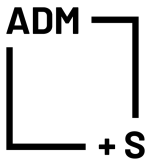
As revealed in the as yet unresolved ‘apportionment debt’ under- and overpayments affecting up to 4 or more million debts, where (unintentionally) Centrelink misapplied the law,²³ errors can involve both human and automated decision-making.²⁴ Fidelity to legality, then, calls for much more than merely ensuring that ADM mirrors the ‘business rules’.²⁵

²² Australian Government. (2024). ‘Guides to Social Policy Law: Social Security Guide’ <<https://guides.dss.gov.au/social-security-guide>>.

²³ Commonwealth Ombudsman. (2023). ‘Accountability in Action: identify, owning and fixing errors’ <https://www.ombudsman.gov.au/_data/assets/pdf_file/0019/302059/FINAL-Income-Apportionment-OMI2-Report.pdf>; Re FTXB [2024] AATA 3021.

²⁴ Morton, R. (2024). *Mean Streak*. HarperCollins.

²⁵ Attorney-General’s Department. (2024). ‘Use of automated decision-making by government’ <https://consultations.ag.gov.au/integrity/adm/user_uploads/consultation-paper-use-of-automated-decision-making-by-government.pdf> 10.



The limited capacity of the public sector to manage complex automation

The Administrative Review Council noted as far back as 2004 that ‘the complexity of modern legislation creates an environment where errors in decision-making are likely’, which could be exacerbated by ‘high staff turnover; pressure of work; limitations on training, and reliance on staff keeping up to date with the legislation’.²⁶ All these issues are still relevant in 2025.

Critique: Critical engagement with stakeholders

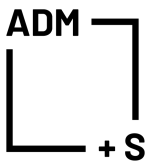
Procedural steps to meaningfully engage with a range of stakeholders to ensure technical and democratic control over, and benefits from AI use. Notably, our research has stressed the importance of consultation on areas often disregarded in the administrative decision to automate:

1. **Digitally excluded Australians**, including older people, First Nations communities, and regional and remote communities, rely to a high degree on an array of government (and commercial) services which are rapidly being digitised and automated.
2. **Disability and accessibility**: ADM+S research raises doubts whether accessibility is sufficiently considered in ADM 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.
3. **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.
4. **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.²⁷ Tribunals such as the Veterans Review Board already explicitly apply such considerations when reviewing decisions and considering ADM elements.

Critique: Mechanisms for allocating responsibility for automated systems

²⁶ Administrative Review Council. *Automated Assistance in Administrative Decision Making* (Report to the Attorney-General, No 46, November 2004) 34.

²⁷ Maitra, S., Sleep, L., Fay, S. & Henman, P. (2024) ‘Building a trauma-informed algorithmic assessment toolkit’, < <https://apo.org.au/node/326724>>.



We agree that ‘the person who authorises or controls the system’ should be required ‘to take reasonable steps to ensure that automated decisions are consistent with the objects of the Act and based on grounds or information on which an officer could have made that decision, or ... to be satisfied that the system would be programmed to comply with administrative law standards’.

The Consultation Paper does not, however, suggest how a responsible person could ensure that they were ‘reasonably satisfied’. The steps to ensure consistency with the Act are crucial here. The right starting point may be a set of common rules for all advanced automation, when the steps or logic of the automation are not easily understandable by humans. This would allow the development of a coherent jurisprudence and the progressive refinement of the rules. Furthermore, as we have noted in this submission, the existence of an entity considering systems horizontally, across government, and developing/applying consistent standards, would enormously facilitate this task.

We have also raised the challenge of identifying the responsible person. Consideration should be given to allocating responsibilities to both the person who authorises the use of the system, and the person who controls its use on an ongoing basis. Both could be required to take steps to ensure the system is operating consistently with law. The Mandatory Guardrails paper applies differentiated obligations to the developer and deployed: a similar approach might be appropriate here.

An Australian response to public sector automation is urgently needed

The consultation paper represents an important advance. It addresses in the Australian context many of the issues of ADMs that are currently being explored in other jurisdictions. There is now a need to develop an Australian-specific perspective. The need to comply with Australian administrative law requires a degree of specificity.

Critique: We can learn from the work of other jurisdictions

For the use of ADM systems at a federal level, the Canadian experience in particular offers important lessons. Their experience could well further inform the Australian approach. The Canada-Australia Public Policy Initiative (CAPPI), which brings together Australian and Canadian senior public sector officials biennially, could be an adequate forum.

Related to this, it would be important to consider how ADM systems developed in other like-minded jurisdictions could be deployed in Australia. The scalability and adaptability of ADM systems, especially those based on AI, could reduce the costs of deploying in Australia systems already tested and deployed in other jurisdictions; and, in reverse, for systems deployed in Australia to be sold abroad. Any regulation of public sector use of ADM should envisage how this could be achieved (accreditation, mutual recognition, simplifying testing, etc.)