



Australian Government
Attorney-General's Department

Use of automated decision-making by government

Consultation paper

NOVEMBER 2024

© Commonwealth of Australia 2024

With the exception of the Coat of Arms, the department's logo, content supplied by third parties and where otherwise stated, all material presented in this publication is provided under a Creative Commons Attribution 4.0 International licence (www.creativecommons.org/licenses).

For the avoidance of doubt, this means this licence only applies to material as set out in this document.



The details of the relevant licence conditions are available on the Creative Commons website as is the full legal code for the CC BY 4.0 licence (www.creativecommons.org/licenses).

Use of the Coat of Arms

The terms under which the Coat of Arms can be used are detailed on the Department of the Prime Minister and Cabinet website (www.pmc.gov.au/honours-and-symbols/commonwealth-coat-arms).

Attribution

Material obtained from this publication is to be attributed as:

© Commonwealth of Australia 2024, *Options for reform of automated decision-making legislation Consultation paper November 2024*.

Acknowledgement of Country

We acknowledge the traditional custodians of Australia and their continuing connection to land, sea and community. We pay our respects to the people, the cultures and the Elders, past and present, and thank them for their ongoing contributions to the culture and prosperity of the Australian nation.

Contents

- Glossary 4**
- Reports 4**
- Introduction..... 5**
 - What is ADM? 7
 - What are ‘decisions’ and ‘administrative action’? 8
 - Benefits and risks of ADM..... 9
 - Scope of this paper 10
 - Interaction with other government reforms..... 10
 - Making a submission 11
- Types of Automation 12**
 - Application to different automation technologies 12
 - Using ADM for decisions and administrative actions 13
- Transparency and safeguards..... 16**
 - Transparency 16
 - Safeguards..... 17
 - Issues for consideration 18
 - Questions..... 31

Glossary

The language used in this paper is set out below.

Abbreviation	Definition
ART	Administrative Review Tribunal
ADJR Act	<i>Administrative Decisions (Judicial Review) Act 1977</i>
ADM	Automated decision-making
Agencies	Australian Government departments and agencies
AI	Artificial intelligence
ARC	Administrative Review Council
FOI Act	<i>Freedom of Information Act 1982</i>
PGPA Act	<i>Public Governance, Performance and Accountability Act 2013</i>
Privacy Act	<i>Privacy Act 1988</i>
Recommendation 17.1	Recommendation 17.1 of the Robodebt Report
Recommendation 17.2	Recommendation 17.2 of the Robodebt Report
Regulatory Powers Act	<i>Regulatory Powers (Standard Provisions) Act 2014</i>
Robodebt Royal Commission	The Royal Commission into the Robodebt Scheme
The Attorney-General's Department	The Australian Attorney-General's Department

Reports

Short title	Long title
Robodebt Report	Report of the Royal Commission into the Robodebt Scheme, Final Report, 7 June 2023
Commonwealth Ombudsman's Guide	Automated Decision-making Better Practice Guide, Commonwealth Ombudsman, 2024
ARC Report	Report 46 – Automated Assistance in Administrative Decision Making, Administrative Review Council, 1 January 2004

Introduction

The Royal Commission into the Robodebt Scheme (Robodebt Royal Commission) called for the Australian Government to consider legislative reform to introduce a consistent legislative framework in which automation in government services can operate.

The use of automated decision-making (ADM) by government is long standing practice. For example, section 6A of the *Social Security (Administration) Act 1999*, which provides that the ‘Secretary may arrange for use of computer programs to make decisions’, was inserted into that Act nearly 25 years ago in 2001.¹ ADM provisions have since been adopted and currently operate across major Commonwealth service delivery contexts, including social services, migration, biosecurity, aged care and veterans’ entitlements. However, because these provisions have been implemented individually over a lengthy period of time, they are not entirely consistent across the statute book.

The piecemeal nature of this approach was strongly criticised by the Robodebt Royal Commission. A desktop review conducted by the Attorney-General’s Department has identified 46 instances of primary legislation containing provisions which authorise the use of ADM. These provisions differ in their complexity. Some provisions only provide authority for ADM use, while others are more detailed and include safeguards. Recent practice has been to include a range of safeguards to accompany standard provisions authorising ADM use. An example of legislation with a more complete set of safeguards is the *Telecommunications Amendment (SMS Sender ID Register) Act 2024*, which recently received assent on 5 September 2024.²

The Robodebt Royal Commission outlined the importance of a cohesive and accessible legislative legal framework aimed at ensuring that algorithms and automated decision systems are fit for purpose, lawful, fair and do not adversely affect human and legal rights.

To address these issues the Robodebt Royal Commission made the following recommendations in relation to ADM:

Recommendation 17.1: Reform of legislation and implementation of regulation

The Commonwealth should consider legislative reform to introduce a consistent legal framework in which automation in government services can operate.

Where automated decision-making is implemented:

- there should be a clear path for those affected by decisions to seek review
- departmental websites should contain information advising that automated decision-making is used and explaining in plain language how the process works
- business rules and algorithms should be made available, to enable independent expert scrutiny.

¹ This provision was inserted by the *Family and Community Services and Veterans’ Affairs Legislation Amendment (Debt Recovery) Act 2000*, which received Royal Assent on 12 June 2001.

² *Telecommunications Amendment (SMS Sender ID Register) Act 2024*
<www.legislation.gov.au/C2024A00083/asmade/text>.

Recommendation 17.2: Establishment of a body to monitor and audit automated decision-making

The Commonwealth should consider establishing a body, or expanding an existing body, with the power to monitor and audit automated decision-making processes with regard to their technical aspects and their impact in respect of fairness, the avoiding of bias, and client usability.

The Australian Government is committed to ensuring services and policies are designed with people at the centre; therefore the use of ADM in delivery should be clear, transparent, accessible, and support improved accuracy and timeliness in decision-making. The use of ADM should not displace review rights for individuals. The following design and delivery features will be critical to support trust in government:

- ethical considerations and principles of administrative law must be taken into account in the design of automated decision-making processes, and there should be proper oversight of these arrangements
- legislative reforms should ensure clear review pathways for those affected by decisions, and
- transparency should extend to both the use of ADM, and how decision-making processes operate.

The Australian Government has committed to considering opportunities for legislative reform to introduce a consistent legal framework for ADM and agreed to consider the establishment, or expansion, of a body to monitor and audit ADM. The development of the new ADM framework will inform the Government's consideration of Robodebt Recommendation 17.2.

The Australian Attorney-General's Department is leading the development of this framework. In the interim, existing Commonwealth ADM provisions will continue to operate while the new framework is developed.

Development of the new framework will be informed by existing reports and best practice guidelines regarding government implementation of ADM.³ These include the Administrative Review Council's (ARC) report on *Automated Assistance in Administrative Decision making*⁴ and Commonwealth Ombudsman's *Automated Decision-Making Better Practice Guide* (Commonwealth Ombudsman's Better Practice Guide).⁵

³ See, for example, *Data-Matching Program (Assistance and Tax) Act 1990* sub-s 12(2); Office of the Australian Information Commissioner, *Guidelines on data matching in Australian Government administration* (Web Page) <<https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/government-agencies/guidelines-on-data-matching-in-australian-government-administration>>.

⁴ Administrative Review Council, *Automated Assistance in Administrative Decision Making* (Report to the Attorney-General, No 46, November 2004).

⁵ Commonwealth Ombudsman, *Automated Decision-making Better Practice Guide* (Better practice guide, 8 November 2019) 5.

More broadly, the making of an administrative decision or the taking of an administrative action should comply with administrative law principles. 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’.⁶

The requirements of administrative law apply regardless of whether a human or ADM system makes a decision. At a minimum, the use of ADM to make an administrative decision or take administrative action would need to comply with these principles.

What is ADM?

ADM refers to the use of technology, commonly referred to in Commonwealth legislation as a ‘computer program’, to automate a decision-making process. Legislative provisions authorising the use of ADM to make decisions are generally expressed to apply to situations where the power or authority to take that action is contained in legislation, rather than decisions or actions more generally. There 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. These kinds of automations are also lower risk, because while they improve administration, they are unlikely to significantly affect the rights of individuals.

Most Commonwealth ADM is delivered through relatively simple rules-based systems. An example of this approach is a system which calculates a rate of payment in accordance with a set of pre-determined rules, based on a formula set out in legislation.

However, ADM may involve the use of more specialised systems which use automated tools to predict and deliberate, such as Artificial Intelligence (AI) systems and including machine learning. AI refers to an engineered system that generates predictive outputs for a given set of human-defined objectives or parameters without explicit programming.⁷

ADM and AI are related, but distinct concepts. AI can be used for a range of purposes by government, which may include assisting decision-making. ADM refers to the use of automated systems (which may include AI) to carry out administrative actions and decisions. An example of AI use to assist an administrative decision might be a system using machine learning to assess elements of a decision, such as biometric assessment of a person’s identity. AI systems are machine-based systems that, for explicit or implicit objectives, infer from the input they receive how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.⁸

⁶ Mark Aronson and Matthew Groves, *Judicial Review of Administrative Action* (Thomson Reuters, 5th ed, 2013) ch 1.

⁷ Department of Industry, Science and Resources, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 5.

⁸ OECD, ‘OECD AI Principles overview’ (Web Page) <<https://oecd.ai/en/ai-principles>>.

This paper also refers to the broader and distinct concept of ‘automated systems,’ as used by the Commonwealth Ombudsman.⁹ Automated systems can be used in the administrative decision-making context to assist decision-making. They can do this by recommending a decision to the decision-maker, providing preliminary assessments for individuals and internal decision-makers, and automating aspects of fact-finding processes in a way that may influence subsequent human-made decisions.¹⁰

What are ‘decisions’ and ‘administrative action’?

A ‘primary decision’ in the Commonwealth administrative law system broadly refers to an ‘original’ decision affecting an individual, that is administrative in character¹¹ and final or operative and determinative on the matter being considered.¹² Examples of decisions may include making, issuing, or failing to make or issue a determination, certificate, approval or a licence.¹³ The use of ADM to make decisions is generally limited to situations where the power or authority to take that action is contained in legislation.

The expression ‘administrative action’ extends beyond decisions to also include exercising a power, or performing a function or duty (or failing or refusing to do so). Administrative actions may have an effect on an individual, either directly or indirectly. Examples of administrative actions may include exercising a power to conduct a hearing, or performing a function to monitor or oversee certain conduct.

A decision-making or administrative action process may also include automated processes that contribute to or are part of the process, but are not ‘decisions’ because they are not determinative of the final outcome or action. These processes are not typically set out prescriptively in legislation. The use of automated processes could include, for example, where a computer extracts relevant data from a system with no analysis, for a human decision-maker to review in making a decision.

The meaning of a ‘decision’, and discussion of other administrative actions that do not involve decisions, is discussed further under **types of automation below**.

⁹ See, Commonwealth Ombudsman, *Automated Decision-making Better Practice Guide* (Better practice guide, 8 November 2019).

¹⁰ *Ibid* 5.

¹¹ *Administrative Decisions (Judicial Review) Act 1977* sub-s 3(1).

¹² *Australian Broadcasting Tribunal v Bond* (1990) 170 CLR 321 at 337 (Mason CJ, Brennan and Deane JJ agreeing).

¹³ *Administrative Decisions (Judicial Review) Act 1977* sub-s 3(2).

Benefits and risks of ADM

A key benefit of ADM is its potential to improve the quality and consistency of administrative decision-making. In its 2004 report, the ARC found that the complexity of modern legislation creates an environment where errors in decision-making are likely. The ARC considered that this legislative complexity was exacerbated by:

- ambiguity or uncertainty in legislation
- frequent amendments to legislation
- the need to administer different provisions for different cases at different times
- high staff turnover
- pressure of work
- limitations on training, and
- reliance on staff keeping up to date with the legislation.¹⁴

Despite its age, the observations of the 2004 ARC report remain relevant and are frequently relied on by the Senate Standing Committees for the Scrutiny of Bills and Delegated Legislation when scrutinising ADM proposals in legislation.

A properly designed ADM system can assist agencies to manage the complexity of modern legislation, and allow high volumes of decisions to be made more efficiently, accurately, and consistently than human decision-makers. In some high-volume contexts, ADM systems may be critical to enabling government to meet industry expectations. This is particularly important in the social services, immigration, taxation and international trade contexts, where decision-making delays or human error could have significant impacts on applicants.

Another key benefit is the ability for ADM to improve levels of service and operational effectiveness of agencies, including near instantaneous decisions on some matters. This can be particularly important where demands on resources are increasing, or fluctuate in a way that makes meeting variations in demand with appropriately trained staff difficult. If appropriately designed, ADM may also increase fairness for applicants by ensuring similar cases are treated identically, reducing human bias and minimising the risk of irrelevant factors being considered.¹⁵

Some examples of decisions which could be subject to automation or are currently automated include Business Name Registrations,¹⁶ Smartgates at Airports, and managing noise certificates for Remote Piloted Aircrafts (RPA) – where there are a large number of RPA operators and low risk applications can be automatically approved.¹⁷

¹⁴ Administrative Review Council, *Automated Assistance in Administrative Decision Making* (Report to the Attorney-General, No 46, November 2004) 34.

¹⁵ Ulrik BU Rohel, 'Automated Decision-Making and Good Administration: Views from Inside the Government Machinery' (2023) 40(4) *Government Information Quarterly* 3; citing Marion Oswald, 'Algorithm-Assisted Decision-Making in the Public Sector: Framing the Issues using Administrative Law Rules Governing Discretionary Power' (2018) 376(2128) *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 1.

¹⁶ *Business Names Registration Act 2011*, ss 62F, 66.

¹⁷ Senate Standing Committee for the Scrutiny of Delegated Legislation, Parliament of Australia, Delegated Legislation Monitor (Monitor No 5 of 2022 - Ministerial Responses, 7 September 2022) 2 <Ministerial_responses_5_of_2022.pdf (aph.gov.au)>.

However, there are a number of risks which agencies using ADM must manage, including algorithmic bias, coding errors, inaccuracy and inconsistency with administrative law requirements.

- **Algorithmic bias** poses risks to individuals when systems are ‘trained’ using data that is outdated, unreliable, unchecked or discriminatory. Automated decisions made using flawed data have the potential to be inaccurate or to generate discriminatory outcomes. In addition, where there is a lack of transparency about how an automated decision has been made, a person whose rights are affected by that decision may be unable to identify the source of the bias to challenge the automated decision.
- **Coding errors**, including algorithms that misrepresent, oversimplify, or misunderstand the law, may be built into ADM systems and lead to legal error. These can arise from a range of errors in code, ranging from typographic errors through to errors in the logic of code that lead to incorrect implementation of business rules.
- Use of ADM for decisions which require a decision-maker to exercise discretion may result in the **inflexible application of pre-determined criteria** to decisions that should have been made on the merits of an individual case.
- Use of ADM can also contribute to the **de-skilling** of human decision-makers and increase reliance on, and default to, ADM.¹⁸

Where these risks are not managed appropriately, ADM may disadvantage individuals and communities, particularly marginalised and vulnerable groups, and undermine public trust and confidence in government administration.

Any new framework should balance the benefits and risks of ADM, by providing clear authority for the use of ADM, while ensuring that it is implemented lawfully and with appropriate safeguards.

Scope of this paper

This paper discusses types of ADM and ADM technologies that support the Australian Government to deliver services, and seeks public views on transparency and safeguards that should be put in place to protect the human and legal rights of individuals in the new ADM framework.

Interaction with other government reforms

Government Response to the Privacy Act Review Report

On 12 September 2024, the Privacy and Other Legislation Amendment Bill 2024 was introduced. The Bill implements proposals 19.1 and 19.2 from the Government Response to the Review of the *Privacy Act 1988* by requiring entities to include in their privacy policies the types of personal information that may be used in automated decisions which have a legal or similarly significant effect on an individual’s rights.

¹⁸ Administrative Review Council, *Automated Assistance in Administrative Decision Making* (Report to the Attorney-General, No 46, November 2004) 40.

The Australian Government has also agreed to introduce a right for individuals to request meaningful information about how substantially automated decisions with legal or similar effect are made (Proposal 19.3 from the Government Response). This proposal will be further considered as part of broader privacy reforms, and in the context of this project, to ensure consistency in approaches.

AI reform proposals

The ADM reform project is intended to complement related AI reforms, including the Australian Government's proposals paper for introducing mandatory guardrails for AI in high-risk settings.¹⁹

The ADM-focused approach proposed in this paper is closely aligned with the proposed regulatory guardrails for AI. Further information on the proposals paper is available on the [Department of Industry Science and Resources' website](#). Feedback received on that proposals paper will inform the Government's regulatory response to help mitigate the potential risks of AI and increase public trust and confidence in its development and use.

The ADM reform project will consider the use of AI in the context of its use in ADM in government services, but will not extend to the use of AI by government or across the economy more broadly.

Making a submission

The department invites feedback on the consultation paper through our Citizen Space consultation page.

You are welcome to answer only those questions that are relevant to you or your organisation.

This consultation closes on 15 January 2025.

To promote transparency in the consultation process, we intend to publish submissions online as appropriate. If you would like us to publish your submission, please indicate this when submitting it. We will not publish your submission without your consent or if we consider, for any reason, that it should not be made public. We may redact parts of published submissions, as appropriate. You can also ask for your submission to be published anonymously.

Submissions may be subject to freedom of information requests, or requests from the Parliament. Personal information shared through the consultation process will be treated in accordance with the *Privacy Act 1988*. For more information on how the Attorney-General's Department collects, stores and uses personal information, please visit the department's Privacy Policy at www.ag.gov.au/about-us/accountability-and-reporting/privacy-policy.

¹⁹ Department of Industry, Science and Resources, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024).

Types of Automation

Automation is a useful way to streamline and guide simple administrative decisions through the use of a consistent decision-making process.²⁰ However, automating decisions raises complex policy issues. As with non-automated decisions, principles of administrative law must be considered in the design of ADM processes. These principles include lawfulness, procedural fairness, independence and impartiality, transparency, efficiency and rationality. Compliance with these principles will depend on the way in which automation is used by an agency. The risk profile of ADM would vary depending on the way it is used – not all uses of ADM would be considered high risk.

Agencies currently use ADM for a range of decision types, and in ways that are essential to promote accuracy and consistency in decision-making and reduce processing delays, particularly in areas where high volumes of decisions are made.

Currently, a range of Commonwealth legislation providing for ADM applies to the use of a ‘computer program’ to make decisions.²¹ Other frameworks refer to the use of ‘electronic means,’²² ‘methods (including technologies)’²³ to make decisions, or ‘processes to assist decision making (such as computer applications and systems)’.²⁴

ADM may occur through various different ways across Commonwealth agencies, including through enabling different automation technologies and through making part or full decisions.

Application to different automation technologies

Current automated systems could include both pre-programmed rules-based processes, and machine-learning processes that learn from patterns and correlations in historic data.²⁵

Pre-programmed rules-based processes rely on a set of facts or sources of data and rules based on ‘if this then that’ logic to choose an action.²⁶ In the decision-making context, there may be, for example, a framework which provides that a payment must be approved if the applicant is over 18 years of age. The process could be programmed to determine the applicant’s age based on identification documents provided as part of the application process, and approve the payment if the applicant is over 18.

²⁰ Administrative Review Council, *Automated Assistance in Administrative Decision Making* (Report to the Attorney-General, No 46, November 2004) 11.

²¹ See, for example, *A New Tax System (Family Assistance) (Administration) Act 1999* s 223, *Agricultural and Veterinary Chemicals Code Act 1994* cl 5F of the Schedule, *Anti-Money Laundering and Counter-Terrorism Financing Act 2006* s 228A, and *Australian Citizenship Act 2007* s 48.

²² *Australian Film, Television and Radio School (Council Elections) Regulations 2017* s 9.

²³ *Australian Passports Act 2005* s 47.

²⁴ *Business Names Registration Act 2011* s 62F, *Commonwealth Registers Act 2020* s 11.

²⁵ Monika Zalnieriute, Lyria Bennett Moses and George Williams, ‘The Rule of Law and Automation of Government Decision-Making’ (2019) 82(3) *Modern Law Review* 425, 432–5.

²⁶ Anna Huggins, ‘Addressing Disconnection: Automated Decision-Making, Administrative Law and Regulatory Reform’ 2021 44(3) *UNSW Law Journal* 1048, 1060.

Different AI systems vary in their levels of autonomy and adaptiveness after deployment.²⁷ This includes machine learning, which involves inferences or predictions based on rules a computer program has learned from patterns and correlations in historical data.²⁸ In the ADM context, a framework could provide that a payment must be made if it is for a charitable purpose. A machine learning process could be designed to assess previous applications and decisions about charitable purposes to infer whether a specific application meets the relevant test, in order to decide whether a payment should be approved.

Use of AI technologies may allow ADM systems to consider a wider range of factors, and more complex factual circumstances, than a decision made by simpler rules-based systems. On the other hand, AI amplifies existing risks such as bias. For example, AI systems may inherit human biases from the data that they are trained on. This can lead to systemic impacts on groups in relation to attributes protected under Australian law, such as race, gender or age. The use of AI in the public service requires specific considerations, such as how it can be used to improve trustworthiness of the public service, do 'more with less' to meet public expectations, and provide better experiences and outcomes for the whole community.²⁹

Although the outcomes of machine learning processes are not pre-determined, their probabilistic results are still shaped by human choices made in the process of designing and training algorithms. However, due to their self-learning properties even programmers may not be able to explain the particular reasons for a specific machine-learning outcome.³⁰

As technology used in ADM is rapidly developing, including both pre-programmed rules-based processes and AI systems, it will be important to ensure that the new framework is applicable and useable into the future and supports the consistent application of future ADM legislation.

It will also be important to ensure that the framework applies appropriately to current use of pre-programmed rules-based processes. A tailored approach would enable the framework to address the unique risks and opportunities presented by each specific ADM process and technology.

Using ADM for decisions and administrative actions

Existing Commonwealth ADM frameworks authorise automation of various actions including:

- specific decisions only
- decisions made under a certain Act or anything related to making those decisions, and
- decisions and other administrative actions under an Act.

²⁷ OECD, 'OECD AI Principles overview' (Web Page) <<https://oecd.ai/en/ai-principles>>.

²⁸ Zalnieriute, Moses and Williams (n 1) 432–5.

²⁹ Department of the Prime Minister and Cabinet, 'How might artificial intelligence affect the trustworthiness of public service delivery?' (27 October 2023) <<https://www.pmc.gov.au/resources/long-term-insights-briefings/how-might-ai-affect-trust-public-service-delivery/summary#:~:text=This%20Long%2Dterm%20Insights%20Briefing.now%20and%20in%20the%20future>>.

³⁰ Anna Huggins, 'Addressing Disconnection: Automated Decision-Making, Administrative Law and Regulatory Reform' 2021 44(3) UNSW Law Journal 1048, 1061.

There is inconsistency and a lack of clarity across Commonwealth legislation in terms of whether the authority to automate applies to decisions only or extends to administrative actions as well. Clarity of this authority would ensure those automations are subject to appropriate regulation and safeguards.

As noted in the introduction, a ‘primary decision’ in the Commonwealth administrative law system broadly refers to an ‘original’ decision affecting an individual. For example, a decision may include making, issuing or failing to make or issue:

- a determination
- a certificate
- a direction
- an approval
- a consent
- a permission
- a licence
- an authority, or
- other instrument³¹

that is administrative in character³² and final or operative and determinative.³³

However, this may not capture all types of administrative actions that may have an effect on an individual, either directly or indirectly. Administrative actions may include making a decision, as well as exercising a power, or performing a function or duty (or failing or refusing to do so), such as:

- exercising a power to conduct a hearing
- exercising a power to investigate a matter
- performing a function to monitor or oversee certain conduct
- performing a function to administer a certain legislative scheme, or
- performing work health and safety duties.

Including administrative actions in the new framework would cover most government actions that have an effect on an individual, including decisions. This would ensure that authorisation for automation, and applicable safeguards, apply broadly. However, this may also include actions that do not have a direct effect on the rights of an individual, which may impose additional obligations and costs on government agencies to comply with the framework without a clear benefit for individuals.

³¹ *Administrative Decisions (Judicial Review) Act 1977* sub-s 3(2).

³² *Ibid* sub-s 3(1).

³³ *Australian Broadcasting Tribunal v Bond* (1990) 170 CLR 321 at 337 (Mason CJ, Brennan and Deane JJ agreeing).

Automated processes that contribute to a decision or administrative action

Automation could be used in a range of different ways in the decision-making process. For example, it could be used to:

- assist decision-makers, by providing guidance on how decision-making parameters apply in a particular case
- make part of a decision, such as to determine factual or objective elements of a discretionary decision, or
- make a full decision.

A decision-making or administrative action process may include automated processes that contribute to or are part of the process, but are not alone determinative of the final outcome or action. A human may need still to be involved in the process, either by reviewing the output of the computer system and making a decision based on that output, or through reviewing a proposed decision by the computer program and signing off on the decision. Examples of this type of automated process include:

- where a computer extracts relevant data from a system with no analysis, for a human decision-maker to review in making a decision
- where a computer analyses data and produces prompts and analysis to inform the decision-maker, and a human then analyses the information and makes a decision
- where a computer analyses data and produces a recommended decision that a human decision-maker agrees with and signs off on, with the human decision-maker either conducting no analysis of their own, or no significant analysis of the data
- where a computer does the analysis and prepares a proposed decision and a human conducts their own analysis of the computer's work and signs off on the final decision
- where a human enters data into a system and the computer uses that input to generate a pre-filled decision letter that a human reviews and signs off on the final decision.

Courts have found that decisions under the *Administrative Decisions (Judicial Review) Act 1977* (ADJR Act) are those that are final or operative and determinative.³⁴ 'Decisions' are not generally considered to include conclusions reached as a step along the way in a course of reasoning to an ultimate decision, unless the relevant Act has a separate provision for it, and it resolves a 'substantive issue.'³⁵

However, there may be automations used as part of a process that have such a significant impact on the final decision that may mean it is appropriate for the automation to be subject to the same level of safeguards as if the automation were used to make the final decision.

³⁴ *Australian Broadcasting Tribunal v Bond* (1990) 170 CLR 321 at 337 (Mason CJ, Brennan and Deane JJ agreeing).

³⁵ *Ibid.*

Transparency and safeguards

This section considers transparency and a range of safeguard options that could be implemented at various stages of the decision lifecycle to ensure that automated decisions by government agencies are fair, transparent, accountable and timely.

The making of an administrative decision or the taking of an administrative action must comply with administrative law principles. As noted in the introduction, these principles generally go to matters including 'openness, fairness, participation, accountability, consistency, rationality, accessibility of judicial and non-judicial grievance procedures, legality and impartiality'.³⁶ Administrative law requirements apply regardless of whether a human or ADM system made the decision, and a decision may be invalid if it does not comply with them. However, as Robodebt demonstrated, considerable harm can occur before complaints-based processes such as merits review and judicial review can put a stop to further unlawful action.

Legal academics, including Dr Yee-Fui Ng and Professor Stephen Gray, have argued that safeguards should be implemented at each stage of the automation process, including in system design and implementation, during decision-making and after automated decisions have been made.³⁷ Adopting a lifecycle approach to safeguards may assist with addressing problems with automations before significant harm can occur, reducing the risk of substantial legal liabilities, and more fairly sharing the burden of identifying and rectifying problems between the agency administering the automation and the recipient of a decision.

Transparency

Transparency has been a particular theme of the debate surrounding ADM. The Robodebt Royal Commission emphasised the need for transparency and accountability in ADM to prevent similar failures of public administration. The Commonwealth Ombudsman's Better Practice Guide outlines a range of transparency and accountability safeguards that should form part of automated system design. Commentators have also argued that transparency relating to the use of ADM in government activities and processes is essential to promote trust in government, government accountability and democratic society more generally.³⁸

³⁶ Mark Aronson and Matthew Groves, *Judicial Review of Administrative Action* (Thomson Reuters, 5th ed, 2013) ch 1.

³⁷ Yee-Fui Ng and Stephen Gray, 'Disadvantage and the Automated Decision' (2022) 43(2) *Adelaide Law Review* 641, 641.

³⁸ See, for example, Michele Loi, *Automated Decision Making in the Public Sector: An Impact Assessment Tool for Public Authorities* (Report, Algorithm Watch, 2021); see also Lyria Bennett Moses and Louis de Koker, 'Open Secrets: Balancing Operational Secrecy and Transparency in the Collection and Use of Data by National Security and Law Enforcement Agencies' (2017) 41 *Melbourne University Law Review* 530–570.

However, the design of any transparency measures will require careful consideration to balance the benefits and risks of transparency. For example, transparency measures may not be appropriate if they would affect national security or reveal trade secrets or commercially valuable information that a person could reasonably expect would be protected. Broader practical constraints may also need to be considered in implementing transparency measures, including personal privacy, cyber security risks and data governance considerations.

Further, in some circumstances it may not be possible or useful to provide full transparency about the specific coding of the ADM system or algorithmic model due to difficulties explaining complex algorithmic processes in a way that is meaningful to general members of the public.³⁹ Alternative options might include publication of business rules or principles on which the ADM algorithm is based, rather than the algorithmic model itself, or more general case studies about how the automated system works.

Safeguards

This section canvasses options for potential safeguards:

- before implementation of an automation
- at the system level, during the lifecycle of an automation, and
- in respect of individual decisions through the lifecycle of a particular decision-making process.

Most of the safeguards outlined, in particular those that could be implemented before implementation and at the system level, apply without individual action. The safeguards outlined below impose obligations on an agency when using an ADM system. These differ from safeguards which require action from an individual, such as a request from an individual to review certain decisions. Those safeguards would provide for review rights but it would be ultimately be the individual's responsibility to seek review of an ADM decision, where they consider it necessary. Safeguards need to be designed carefully to appropriately balance the considerations of managing regulatory impacts on agencies with ensuring that safeguards apply consistently across agencies.

Under each option, agencies may be exempt from transparency mechanisms and safeguards if certain criteria are met. For example, if publication or disclosure would affect national security or reveal trade secrets or commercially valuable information.

As discussed earlier, agencies may use ADM in different ways. It is also important to consider how safeguards may need to be adjusted based on how an agency is using ADM. ADM should be implemented in a way that improves the efficiency and accuracy of government decision-making processes, while embedding legality, fairness and respect for human rights.

³⁹ See generally Jenna Burrell, 'How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms' (2016) [January-June] *Big Data & Society* 1, 3-5.

Issues for consideration

Options for transparency about how ADM is used

There are a range of possible ways to provide transparency about the use of ADM in decision-making that have been included in existing legal frameworks or that have been suggested by bodies, such as the Robodebt Royal Commission. These include requirements to publish general information about the use of ADM, report statistics about an agency's use of ADM, and publish business rules and algorithms.

Publication of general information about the use of ADM

The Robodebt Royal Commission recommended that where ADM is used, departmental websites should contain information advising that ADM is used and explaining in plain language how the process works.⁴⁰ This would allow for external scrutiny on an agency's use of ADM and improve public confidence in an agency use of ADM. An example of such a provision is section 484K(5) of the *Telecommunications Amendment (SMS Sender ID Register) Act 2024*, which provides that when ADM is used, the chair of the Australian Communications and Media Authority (ACMA) must publish on ACMA's website that an ADM arrangement has been made, and the particular provisions that the ADM arrangement has been made in relation to.⁴¹

Most agencies have an existing obligation to publish operational information under the *Freedom of Information Act 1982* (FOI Act).⁴² Under the FOI Act, the information to be published includes information held by the agency that assists in making decisions or recommendations affecting members of the public.⁴³ There might be other mechanisms that could facilitate transparency around ADM use, such as requiring information to be published in a notifiable instrument or in the Government Gazette. Publication of general information about the use of ADM would be consistent with the Australian Government's policy for the responsible use of AI in government, which requires government agencies to publish transparency statements about their approach to AI adoption and use.⁴⁴

Public reporting on the use of ADM

Another option to increase government transparency around the use of ADM could include regular reporting requirements, including statistics about an agency's use of ADM. Reporting may be done through a number of mediums, including on an agency's website, in an agency's annual report or to a central registry for publication.

⁴⁰ *Royal Commission into the Robodebt Scheme*, (Final Report, July 2024) vol 1, xvi.

⁴¹ See, for example, *Telecommunications Amendment (SMS Sender ID Register) Act 2024* s 484K(5).

⁴² *Freedom of Information Act 1982* sub-s 8(2)(j).

⁴³ *Ibid* sub-s 8A(1).

⁴⁴ Digital Transformation Agency, *Policy for the responsible use of AI in government* (1 September 2024) 13.

There are advantages and disadvantages to these options for reporting, including frequency of data, ease of access to the data by members of the public, and ease of aggregation of data across government for analytical purposes. Providing reporting requirements would be consistent with some aspects of the Australian Government's proposed AI guardrails which include requirements to publish accountability processes and inform people when AI is used to make or inform decisions relevant to them.⁴⁵ It would also be consistent with:

- the Privacy and Other Legislation Amendment Bill 2024,⁴⁶ which would increase transparency about substantially automated decisions which significantly affect individuals' rights or interests. Entities would be required to include information in their privacy policy about the kinds of decisions and kinds of personal information used in these decisions. It is not expected that the information required to be included in privacy policies would be at a level of detail which would compromise commercially sensitive information about automated decision-making systems, or confidential information about ongoing detection of wrongdoing using automated decision-making systems, and
- the Australian Government's policy for the responsible use of AI in government, which requires government agencies to publish transparency statements about their approach to AI adoption and use.⁴⁷

Publication of business rules and algorithms

The Robodebt Royal Commission recommended that where ADM is implemented, business rules and algorithms should be made available to enable independent expert scrutiny.⁴⁸ The Ombudsman's Better Practice Guide also recommends that the underlying business rules of an automated system must be readily understandable, and information about automated systems should be publicly available.⁴⁹

Business rules and algorithms may be very complex, and if published as code may not be comprehensible to general members of the public. To facilitate transparency and improve access to justice it may therefore be necessary to provide rules about how rules should be published in a way that may be easily understood and explained to members of the public.

It is also necessary to consider risks around the publication of full business rules. One example of a risk is that providing a complete specification of how certain processes work may open opportunities for fraud, by enabling fraudsters to engineer applications or claims to navigate the business rules to a desired result. Full publication of business rules may also raise cyber security risks if careful consideration is not given to the level of information published. It may allow malicious actors to circumvent regulatory requirements or understand operational sensitivities relating to law enforcement activities.

⁴⁵ See, guardrail 1 and 6, Department of Industry, Science and Resources, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 34, 39.

⁴⁶ Part 15, Privacy and Other Legislation Amendment Bill 2024.

⁴⁷ Digital Transformation Agency, *Policy for the responsible use of AI in government* (1 September 2024) 13.

⁴⁸ Royal Commission into the Robodebt Scheme (Final Report, July 2024) vol 1, xvi.

⁴⁹ Commonwealth Ombudsman, *Automated Decision-making Better Practice Guide* (Better practice guide, 8 November 2019) 25.

Information relating to certain operation of agencies may already be exempt from publication under FOI laws. The exemptions and conditional exemptions under the FOI Act⁵⁰ provide for the protection of sensitive information. Exemptions can apply to a range of information, including information that is commercially valuable or would disclose trade secrets⁵¹. Conditional exemptions may also apply to information, for example, section 47E of the FOI Act provides that ‘certain operations of agencies’ are conditionally exempt if their disclosure would reasonably be expected to affect certain operations of the agency. This conditional exemption applies to a number of circumstances to safeguard the operation of agencies, including documents that if disclosed would prejudice the effectiveness of procedures or methods for the conduct of tests, examinations or audits by an agency, or if disclosed would have a substantial adverse effect on the proper and efficient conduct of the operations of an agency. However, conditional exemptions are subject to a public interest test which involves determining whether, on balance, the benefit to the public resulting from disclosure is outweighed by the benefit to the public of withholding the information.

A further consideration is how publication of rules or algorithms may affect third parties, including intellectual property rights and commercial confidentiality. For example, in some circumstances, the importance of protecting commercial confidential information, such as trade secrets, could mean it is inappropriate to publish the business rules and algorithms associated with ADM in full.

In addition, the publication of business rules may not necessarily be possible with some technologies such as AI.

Options for safeguards during the pre-implementation phase

One potential approach to safeguards is to require appropriate action to be taken before implementation of an ADM system to ensure that the automation is appropriate on policy grounds, such as fairness and efficiency, and appropriately designed to address key risks such as legal compliance, errors and impacts on individuals. This is consistent with the Commonwealth Ombudsman’s advice that agencies must establish appropriate governance frameworks for automated systems and ensure that a range of areas are involved pre-implementation, including legal, policy and program areas.⁵²

⁵⁰ See *Freedom of Information Act 1982* part IV; Freedom of Information Guidelines Part V.

⁵¹ *Freedom of Information Act 1982* s 47.

⁵² Commonwealth Ombudsman, *Automated Decision-making Better Practice Guide* (Better practice guide, 8 November 2019) 17.

Risk Assessments

Given the risks inherent in using ADM, one possible pre-implementation safeguard is to include requirements for general risk assessments, or assessments of particularly important risks, in the framework. The Robodebt Royal Commission endorsed the advice of the Commonwealth Ombudsman's Guide that the design of automated systems should integrate administrative law, privacy and human rights obligations through appropriate planning and assessment.⁵³ A risk assessment would allow risks to be considered, identified and mitigated prior to the deployment of an ADM system, reducing the likelihood of any large scale detrimental impacts of these systems.

Accountable authorities have an existing obligation to establish and maintain an appropriate system of risk oversight and management for an entity.⁵⁴ In addition to design, implementation, and legal risks, there may be particular risks relating to ADM that should be assessed, such as human rights, privacy, and data security. The need to conduct risk assessments before implementing an ADM system is recommended in the Commonwealth Ombudsman's Guide.⁵⁵

The need to conduct a risk assessment prior to implementing an ADM system could be incorporated into existing processes. Consideration could be given to:

- explicitly providing that existing risk assessments processes apply prior to implementing an ADM system
- how existing risk assessments could be adapted for assessments prior to implementing an ADM system, and
- avoiding duplication with existing risk assessments.

Although risk assessment is an existing requirement for Commonwealth agencies, incorporating greater detail about risk assessments for ADM in a new framework may clarify the relevance of risk assessment to ADM processes. Including a requirement to undertake a risk management process prior to the implementation of an ADM system would also be consistent with the Australian Government's proposed AI guardrails, which would require entities to establish risk management processes for high-risk AI systems.⁵⁶

Options for system-level safeguards

A further area of possible safeguards is mechanisms to ensure that a particular automation system is governed properly and operates appropriately. Many of these system-level safeguards are already recommended for agencies to implement through the Commonwealth Ombudsman's Guide, and they could extend to requiring audit and compliance mechanisms, requirements to ensure that automations comply with law, or requirements to ensure that automations are kept appropriately up to date, amongst others.

⁵³ See *Royal Commission into the Robodebt Scheme* (Final Report, July 2024) vol 2, 486, referring to Commonwealth Ombudsman, *Automated Decision-Making: Better Practice Guide* (2020) 6.

⁵⁴ *Public Governance, Performance and Accountability Act 2013* s 13.

⁵⁵ Commonwealth Ombudsman, *Automated Decision-making Better Practice Guide* (Better practice guide, 8 November 2019) 8.

⁵⁶ Guardrail 2, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 34.

Audit and compliance mechanisms

The Robodebt Royal Commission recommended that where ADM is implemented, business rules and algorithms should be made available to allow independent expert scrutiny.⁵⁷ As the Commonwealth Ombudsman's Guide notes, scrutiny of ADM – including conducting audits of the automated system – contributes significantly to confidence in public administration.⁵⁸

To ensure disclosure and exposure to audit, the framework could include a requirement for agencies to undertake regular internal and external audits of the automated system, and maintain good record keeping.⁵⁹ Audits could help agencies monitor any updates to legislation, such as changes to eligibility criteria, to ensure the accuracy of information collected by the system used for ADM. Audit mechanisms are also intended to ensure agencies are taking reasonable steps to ensure that the automated system is designed, used and maintained in such a way that it accurately and consistently reflects the relevant law and policy.⁶⁰ Good record keeping would assist with audits and monitoring, and support accountability and transparency of ADM systems.

Audit and compliance mechanisms would align with the Australian Government's proposed AI guardrails, which provides that organisations developing or deploying high-risk AI systems are required to 'Protect AI systems, and implement data governance measures to manage data quality and provenance,' 'Test AI models and systems to evaluate model performance and monitor the system once deployed' and 'Keep and maintain records to allow third parties to assess compliance with guardrails'.⁶¹

Taking reasonable steps to ensure automated decisions are consistent with the objects of the Act, meet administrative law standards, and are governed appropriately

When making an administrative decision or taking administrative action, government agencies must comply with administrative law principles, whether a human or ADM system took the action or made the decision. These principles include ensuring that the administrative action was consistent with the legislation under which the action was made, all relevant considerations were taken into account, and irrelevant considerations were not taken into account.⁶²

The Robodebt Royal Commission found that the Robodebt Scheme fell short of principle 7 of the ARC report on ADM. Principle 7 is that the **construction** of an automated system must comply with administrative law standards if decisions made with the rule base are to be lawful. The principle also states that **decisions made by or with** the assistance of an automated system must comply with administrative law standards to be legally valid.

⁵⁷ *Royal Commission into the Robodebt Scheme*, (Final Report, July 2024) vol 1, xvi.

⁵⁸ Commonwealth Ombudsman, *Automated Decision-making Better Practice Guide* (Better practice guide, 8 November 2019) 26.

⁵⁹ See, principle 19, Administrative Review Council, *Automated Assistance in Administrative Decision Making* (Report No 46, 2004) 43.

⁶⁰ See, principle 10, Administrative Review Council, *Automated Assistance in Administrative Decision Making* (Report No 46, 2004) 5-6.

⁶¹ Guardrail, 1 *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 34.

⁶² *Administrative Decisions (Judicial Review) Act 1977* s 5.

The Commonwealth Ombudsman's Guide outlines best practice to ensure that an ADM system's outputs are consistent with statutory provisions. The guidance outlines requirements in relation to regular testing and processes to assess the quality of data being used for ADM.⁶³

If followed, this guidance and general administrative law principles would require an agency to take reasonable steps to ensure automated decisions are consistent with the objects of the Act, governed appropriately and meet administrative law standards. One option for the new framework could be to clarify this duty by:

- requiring the person who authorises or controls the system 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
- requiring the person who authorises or controls the automated system to be satisfied that the system would be programmed to comply with administrative law standards.

The framework could be accompanied by guidance material to clarify the relevant administrative law standards agencies must comply with.

These safeguards would also be consistent with the National Framework for the Assurance of AI in Government, by ensuring that government use of ADM complies with law, policy and guidelines that regulate such processes.⁶⁴ They would also facilitate the implementation of accountability processes for ADM, by ensuring a person is responsible for the system's programming in line with administrative law principles and legislative requirements.⁶⁵

Requirement to update the system throughout its lifecycle

If a system is not regularly maintained, it may over time become out of step with the current state of the law, or develop errors in its processing. Further, updating the system could ensure that any administrative action taken by the system is legally valid and consistent with administrative law principles. This safeguard would be consistent with principle 5.4 of the with the National Framework for the Assurance of AI in government which provides that government should ensure the use of AI is continuously monitored and evaluated to ensure safety.⁶⁶ It is also consistent with the continuous monitoring and evaluation of AI systems to ensure they remain fit for purpose.⁶⁷

⁶³ Ibid 23.

⁶⁴ Principle 7.1, Australian Commonwealth, State and Territory Governments, National Framework for the Assurance of Artificial Intelligence in Government (National Framework, 21 June 2024) 23 <<https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government>>.

⁶⁵ Guardrail 1, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 34.

⁶⁶ Principle 5.4, Australian Commonwealth, State and Territory Governments, National Framework for the Assurance of Artificial Intelligence in Government (National Framework, 21 June 2024) 20.

⁶⁷ Digital Transformation Agency, Policy for the responsible use of AI in government (1 September 2024) 14.

A number of existing guidance materials already provide that systems must be updated as matter of good practice. Principle 12 of the ARC report on ADM is that systems must be regularly updated and maintained.⁶⁸ This is consistent with the Commonwealth Ombudsman's advice that automated systems should be designed so that changes in the business rules can be easily updated across systems.⁶⁹

In addition, the Commonwealth Ombudsman Guide recommends that technical solutions should be found that maximise the interoperability of the automated system interface (with other IT systems), to minimise the cost, time and disruption caused by the update process.⁷⁰ As the Commonwealth Ombudsman notes, the update process is a vital determinant of the accuracy of the decisions made by an automated system. Depending on the complexity and frequency of legislative, policy or procedural change, updating could involve only simple changes in various fields, or the incorporation of large sets of new or changed business rules into the system.

The framework could consolidate and clarify how these recommendations apply into a single source. Depending on the complexity and frequency of legislative, policy or procedural change, updating could involve only simple changes in various fields, or the incorporation of large sets of new or changed business rules into the system.⁷¹

This safeguard would be consistent with the Australian Government's proposed AI guardrails by promoting transparency and information sharing with other organisations in the system's lifecycle, to help agencies effectively address risks.⁷²

Options for decision-level safeguards

A third potential area for safeguards are measures to ensure that individual decisions are made fairly and lawfully, that errors can be corrected, and that there is appropriate transparency to persons affected by decisions about the use of ADM. These safeguards could include pathways for human intervention in decisions, substitution of incorrect decisions made by the automated system, and notification requirements.

⁶⁸ Principle 12, Administrative Review Council, *Automated Assistance in Administrative Decision Making* (Report to the Attorney-General, No 46, November 2004) 37.

⁶⁹ Commonwealth Ombudsman, *Automated Decision-making Better Practice Guide* (Better practice guide, 8 November 2019) 21.

⁷⁰ *Ibid.*

⁷¹ *Ibid.*

⁷² Guardrail 8, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 39.

Pathways for human referral

The Robodebt Royal Commission considered that human intervention is one of the most effective safeguards against the failure of automated systems.⁷³ This view is supported by the Australian Human Rights Commission,⁷⁴ as well as a range of legal academics who have emphasised the importance of human oversight in ADM, including Associate Professor Ben Wagner and Professor Anna Huggins.⁷⁵ Principle 6.3 of the National Framework for the Assurance of Artificial Intelligence in Government also provides for the implementation of human oversight to promote accountability and ensure that decisions made by AI systems are explainable.⁷⁶

Another possible safeguard that could be included in the framework is an obligation for human referral pathways. For example, if the information entered into a computer system does not squarely fall within the pre-programmed rules, pathways could be created for ‘further information’ to be provided and for the particular decision to be referred to a human officer.

A different approach for particularly sensitive classes of decision could be to allow the system to make beneficial decisions only, and refer any potentially adverse decisions to a human for consideration. However, this approach may not be appropriate in a range of circumstances. For example, where there is no decision-making discretion, and referral would not change the outcome of a particular decision, it is likely that this approach would increase cost and delay without tangible benefit to a person affected by a decision.

This safeguard would align with the Australian Government’s proposed AI guardrails, by facilitating human control or intervention in an automated system to achieve meaningful human oversight.⁷⁷

Power to substitute a decision

Enabling human decision-makers to substitute their own decision in place of computer-made decisions could provide a separate avenue for errors to be corrected, in addition to any merits review process. This safeguard would align with the Australian Government’s proposed AI guardrails, by enabling human control or intervention in an automated process to achieve meaningful human oversight.⁷⁸

⁷³ *Royal Commission into the Robodebt Scheme* (Final Report, July 2024) vol 2, 487.

⁷⁴ Recommendation 8, Australian Human Rights Commission, *Human Rights and Technology* (Final Report, March 2021) 68 – 72.

⁷⁵ See, for example, Ben Wagner, ‘Liable, But Not In Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems (2019) 11(1) *Policy and Internet* 104, 115; Anna Huggins, ‘Addressing Disconnection: Automated Decision-Making, Administrative Law and Regulatory Reform’ (2021) 44(3) *University of New South Wales Law Journal* 1048, 1061.

⁷⁶ Principle 6.3, Australian Commonwealth, State and Territory Governments, National Framework for the Assurance of Artificial Intelligence in Government (National Framework, 21 June 2024) 21.

⁷⁷ Guardrail 5, Department of Industry, Science and Resources, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 37.

⁷⁸ Guardrail 5, Department of Industry, Science and Resources, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 37.

In determining the appropriate framing for a substitution power, it is important to consider the need to achieve finality in government decision-making, while ensuring individuals are not unfairly disadvantaged or advantaged by the use of ADM. Many existing frameworks facilitate substitution of primary decisions (including automated decisions) with a decision made by the relevant senior person or delegate. These frameworks apply differently in terms of when a decision may be substituted and the protections that apply.

- *Basis for substituting a decision*

The appropriate basis for allowing a decision to be substituted may depend on what type of decisions the framework applies to. For example, the substitution power could be:

- Available to substitute an incorrect automated decision with a *correct decision*, being the only possible decision available on the facts and the law.
 - This may be most suitable for a framework limited to non-discretionary decisions involving only objective criteria as these decisions only have one possible answer available on the facts and the law.
- Available to substitute an undesirable automated decision with a *correct or preferable decision*, where there are multiple possible decisions available on the facts and the law.
 - This may be most suitable for a framework applying to decisions involving discretion or evaluative criteria as these decisions have multiple options possible.
- Available for certain outcomes, for example where the substituted decision is more favourable.
 - This may be most suitable where there is a high risk of a substituted decision unfairly disadvantaging an individual and a low risk of inappropriate windfall gains to individuals arising from computer errors.

- *Requirement to substitute a decision*

Most existing ADM frameworks do not include an **obligation** to substitute a decision when an error has been identified, as there may be circumstances where that is unnecessary or inappropriate. For example, there may be some instances where the substitute decision would not change the outcome, such as when a decision has been made in regard to an event that has already occurred.

However, the framework could require a decision-maker to consider whether to substitute a decision when an error is identified. This would ensure the agency has appropriate flexibility in resolving administrative errors.

An applicant could then separately consider merits review if they would like the agency to take a different action to resolve the error.

- *Specifying a date of effect*

Most ADM frameworks are silent on when a substituted decision is taken to have effect. Where frameworks provide for a decision (automated or not) to be changed and specify the 'date of effect', that date often differs. Some do not expressly deal with the date a changed decision comes into effect, while others have detailed and complex backdating regimes that depend on the particular decision-making power and outcome.

A key challenge for a unified ADM framework will be promoting consistency, while recognising the different policy considerations that might affect backdating rules in particular legislative regimes.

- *Protections that could be included in a substitution power*

There may be some instances where substituting a decision may not be the preferred or appropriate approach.

For example, some frameworks in the non-ADM context provide protections for actions undertaken in reliance of an incorrect decision, where that incorrect decision is due to an administrative error by the agency. For example, some frameworks:

- require a debt for an overpayment to be waived if the debtor received the payment in good faith and would suffer severe financial hardship if not waived, or
- allow a person's non-compliant actions to be treated as compliant with the legislation.

A consistent ADM framework could put in place similar protections where there has been reliance on an incorrect decision. This would need to be balanced with consideration of not substituting the decision where it unfairly disadvantages a person, such as where the error detrimentally affects the interests of third parties or the general public.

Additional measures to ensure the substitution power neither advantages nor disadvantages an applicant could also be considered. For example:

- The ARC recommended that before an expert system is overridden, a primary decision maker should first discuss the decision with a senior officer. The requirement to consult a senior officer depends on whether authorised use of ADM is limited to decisions made by senior officers.
- Procedural fairness is an important requirement for administrative decision-making. Requirements for procedural fairness may include providing an applicant with the opportunity to respond to issues or provide further information in support of their application. Procedural fairness duties arise under the common law. However, some legislative frameworks include express procedural fairness requirements for certain decisions. The framework could include express procedural fairness requirements when deciding to substitute a decision, or provide additional guidance to decision-makers in relation to meeting their common law procedural fairness duty.

Notifying persons affected by decisions involving ADM

Given the potential for automated systems to affect the rights and interests⁷⁹ of individuals, as a general principle, individuals affected by these decisions should be notified about the role of those systems in making a specific decision, and have the opportunity to seek review of that decision. This is consistent with principle 6.1 under the National Framework for the Assurance of Artificial Intelligence in Government to ensure the use of AI systems are disclosed to users and people who may be impacted by it.⁸⁰

Without specific notification about the use of ADM, a person affected by a decision may be unaware that ADM has been used, even where that information is publicly available. For example, this could occur where information about an agency's use of ADM is contained in legislation, regulations or on a government website which are not sources of information member of the public typically look at.

One option to address this could be to require agencies to provide individuals affected by a decision with written notice about the use of ADM to make that decision. If this rule were adopted, it may be appropriate to include exceptions, for instance, for self-executing provisions which do not constitute a 'decision'.

If notification requirements are implemented, a further consideration is when notification should occur. Notification of the use of ADM could be pre-decision in some circumstances or provided at the same time that the decision is notified to a person. For example, notification could occur pre-decision if the applicant was asked to complete a declaration acknowledging that their application would be assessed with ADM, before starting the application.

A key risk of notification provisions is producing very large numbers of notifications in relation to minor or routine decisions, both in terms of the frequency of communication with end users, and in relation to what those end users are able to do with the information. Notification requirements may need to balance transparency with the meaningfulness of the information that is provided and notification fatigue.

Including notification rules would be consistent with the Australian Government's proposed AI guardrails by providing individuals affected by an automated decision with written notice about the use of ADM to make that decision.⁸¹

⁷⁹ The use of the language 'rights and interests' is intended to have broad coverage. Rights do not have the same application in Australian law as in Europe which has more developed rights-based frameworks. Interests may include things that are not rights under the Australian law – i.e. the provision of benefits under an Act or denial of significant services or support.

⁸⁰ Principle 6.1, Australian Commonwealth, State and Territory Governments, National Framework for the Assurance of Artificial Intelligence in Government (National Framework, 21 June 2024) 21.

⁸¹ See, guardrail 6, Guardrail 5, Department of Industry, Science and Resources, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 37.

Notification rules should avoid being complex, lengthy, legalistic and vague, which can undermine individuals' understanding of how automated systems affect their rights and interests. Consideration should be given to ensuring notification is clear, up-to-date, concise and understandable, with appropriate accessibility measures in place in line with reforms being progressed to collection notices under the Privacy Act.⁸²

Statement of reasons

The Robodebt Royal Commission stated that individuals subject to decisions made by an automated process should know or understand the reasons behind those decisions.⁸³ This is supported by principle 6.3 of the National Framework for the Assurance of Artificial Intelligence in Government which recommends government provide clear, simple explanations for how an AI system reaches an outcome.⁸⁴ The Australian Government has also agreed to reform the Privacy Act to give individuals more transparency about substantially automated decisions which significantly affect individuals' rights or interests, including through a right to request meaningful information about how such decisions are made.⁸⁵ While there is no common law right of individuals to receive explanations or reasons about how a decision is made, the High Court has recognised this as fundamental to good administrative decision-making.⁸⁶ Legislation also requires statements of reasons in particular circumstances. For instance, if a decision is subject to review by the ART or if a provision of the ADJR Act applies to a decision, a person who is entitled to review of the decision can seek reasons for the decision.⁸⁷

It would not be desirable for the framework to interfere with existing rights to request reasons under the ART Act and ADJR Act. If a person would have a right to request reasons for a human-made decision, they should (and likely already do) have a right to request reasons if that same decision was automated. Consideration is being given to avoiding duplicative requirements to provide reasons under privacy and administrative law frameworks. An additional requirement to provide a statement of reasons for automated decisions would address the Robodebt Royal Commission's views, but could be duplicative in many circumstances.

Options for post-decision safeguards

Finally, there is a potentially important role for safeguards that allow a decision to be challenged after it has been made. The most common example of such rights is a right to merits review, that is a right to have an internal or external reviewer reconsider a decision on its merits.

⁸² This reflects proposal 10.1 of the Privacy Act Review Report, which the Government agreed to in-principle (Government Response Privacy Act Review Report, September 2023).

⁸³ *Royal Commission into the Robodebt Scheme* (Final Report, July 2024) vol 2, 486.

⁸⁴ Principle 6.3, Australian Commonwealth, State and Territory Governments, National Framework for the Assurance of Artificial Intelligence in Government (National Framework, 21 June 2024) 21.

⁸⁵ The Government agreed to proposal 19.3 in the Privacy Act Review Report (Government Response Privacy Act Review Report, September 2023).

⁸⁶ See, for example, McHugh J in *Re Minister for Immigration and Multicultural and Indigenous Affairs; Ex parte Palme* (2003) 216 CLR 212.

⁸⁷ See, *Administrative Review Tribunal Act* s 111; *Administrative Decisions (Judicial Review) Act* s 13.

Merits review rights

Effective merits review is an essential part of the legal framework that protects the rights and interests of individuals. It promotes government accountability and plays a broader role in improving the quality and consistency of government decisions.⁸⁸

Decisions made by or with the assistance of a computer system can be subject to internal or external review. Internal review is a process of review on the merits of an agency's primary decision. It is undertaken by another officer within the same agency, usually a more senior officer.⁸⁹ External merits review is the independent review of government decisions by the ART. This process involves stepping into the shoes of the decision-maker to determine whether a decision was the 'correct or preferable' decision.⁹⁰

The Robodebt Royal Commission recommended that where ADM is implemented there should be a clear path for those affected by decisions to seek review. The Commission considered that the availability of review pathways are vital safeguards in the use of ADM.⁹¹ This is consistent with principle 7 of the National Framework for the Assurance of Artificial Intelligence in Government which provides that when an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.⁹²

Providing review mechanisms for automated decisions would also be consistent with the Australian Government's proposed AI guardrails which require organisations developing or deploying high-risk AI systems to establish processes for people impacted by AI systems to challenge uses or outcomes.⁹³

There are a number of ways that merits review could be embedded in the new framework for ADM, including by providing that:

- both internal and external review is available by right for all automated decisions
- external review is only available by right for all automated decisions, or
- external review is only made available if it would ordinarily apply to a decision made by a human.

Any approach would likely need to be modified in each legislative context to account for existing internal and external review mechanisms, to avoid overlap or duplication.

⁸⁸ Gerard Brennan, 'Canberra Opening Address' (Speech, The Administrative Appeals Tribunal Twentieth Anniversary Conference, 1 July 1996); *Minister for Immigration and Ethnic Affairs v Pochi* (1980) ALD 130, 154 (Deane J). See also Garry Downes, 'Structure, Power and Duties of the Administrative Appeals Tribunal of Australia' (Speech, Administrative Court of Thailand and Central Administrative Court of Thailand, 21 February 2006) [48]; *Re Drake & Minister for Immigration and Ethnic Affairs (No 2)* (1979) 2 ALD 634, 639.

⁸⁹ Administrative Review Council - Internal Review of Agency Decision Making – Report to the Attorney-General – Report No. 44, November 2000, paragraph 1.4.

⁹⁰ See *Drake v Minister for Immigration and Ethnic Affairs* (1979) 2 ALD 60.

⁹¹ *Royal Commission into the Robodebt Scheme* (Final Report, July 2024) vol 2, 486.

⁹² Principle 7, Australian Commonwealth, State and Territory Governments, National Framework for the Assurance of Artificial Intelligence in Government (National Framework, 21 June 2024) 23.

⁹³ Guardrail 7, Guardrail 5, Department of Industry, Science and Resources, *Safe and Responsible AI in Australia: Proposals paper for introducing mandatory guardrails for AI in high-risk settings* (5 September 2024) 38-39.

As outlined earlier in the **types of automation** section of this paper, there may be particular complexities and higher risk involved in automating some decisions, such as discretionary decisions. However, where the risk of automation is higher, increased safeguard mechanisms could be put in place or particular safeguards could be mandatory to address potential risks. For example, pathways for human referral and review rights could apply to higher risk automations, such as discretionary decisions.

Questions

1. How should the need for transparency about the use of ADM be balanced with the need to protect sensitive information about business processes and systems?
2. What transparency rules would be appropriate to build into the framework?
3. What pre-implementation safeguards should apply where ADM is intended to be used?
4. What system-level safeguards should be required to ensure that ADM operates appropriately?
5. What decision-level safeguards should there be for persons affected by decisions made using ADM (for example, review rights)?
6. What post-decision safeguards should there be to allow a decision to be challenged after it has been made?
7. Should individuals be notified of the use of ADM? If so, should notification be required at a specific point in the decision-making process, or should flexibility be provided to agencies about the appropriate time to make a notification?
8. Should there be any exemptions to ADM safeguards? If so, what exemptions should be included and why?
9. Should safeguards be different depending on the risks associated with the use of ADM for a particular decision or administrative action?

ag.gov.au