ACCESS TO WELFARE JUSTICE AFTER THE ROBODEBT ROYAL COMMISSION: THE CASE FOR SYSTEMIC MEASURES?

Terry Carney AO*

This paper stress tests recent legal measures and associated reforms contemplated as 'futureproofing' against any repetition of automated decision-making ('ADM') disasters such as Robodebt. I argue that the reformed accountability and regulatory machinery offers greater protection against violations to individual rights and some promising, if untested, potential for addressing systemic harms from the deployment of ADM in human services, but does little to address more powerful structural sources and reasons for such risks. To address the more powerful structural sources of ADM harm in human services, I make the case for a broader set of measures. I argue that there is an urgent need for more creative measures, better attuned to the nature of bureaucracy, the technology at hand and the power asymmetries associated with ADM deployment within government.

Contents

I Introduction 11
II Robodebt and Reflections on Its Causes and Prevention 14
A Robodebt's Legal and Administrative Failings 14
B Research Findings on Robodebt 15
C Framings Pertinent to Risks of ADM in Human Services 18
1 Infrastructural Relationships of Law, Bureaucracy and Politics 18
2 Administrative Burden 19
3 Algorithmic Justice 20
4 Differential Harms Specific to Welfare 21
III A Post-Robodebt Thought Experiment 22
A Background 23
B Stress Testing a Flawed Post-Robodebt ADM Initiative 25

* Emeritus Professor of Law, University of Sydney Law School; Associate Investigator, Australian Research Council Centre of Excellence for Automated Decision-Making and Society. I gratefully acknowledge the insightful contributions made by three anonymous referees and various colleagues who have read drafts of this article.

1 How Would Citizens Discover the Use of ADM as a Potential Flaw? ... 25
2 How Could Individuals Review Flawed ADM? ... 26
3 How Would an Individual Review Translate into Systemic Action? ... 28
4 Would Exposure of Flawed ADM Necessarily Be Remedied? ... 29
IV Individual Safeguarding Pathways ... 31
A Reforming Merits Review — Correcting Individual Error ... 31
1 Administrative Burden ... 33
2 The 'Decision' Gateway ... 35
3 Government Gaming of Review Avenues ... 36
B Refreshing the Normative Role of Merits Review ... 38
1 The Robodebt Reforms Addressing Systemic Issues ... 38
2 Managerial Reforms as an Answer to Systemic Issues ... 39
V Addressing Systemic Risks of ADM in Government ... 42
A The Royal Commission and Regulation of ADM in Government ... 42
B Transparency of ADM Use ... 45
1 The Case for Unqualified Transparency of ADM Algorithms ... 45
2 Third-Party Proprietary Rights ... 48
C Review Pathways and Decision-Level Safeguards ... 48
1 A New ART Pathway or a Purpose-Designed Pathway? ... 49
2 Notification of Use of ADM ... 51
3 System-Level and Design/Post-Implementation Safeguards ... 53
4 The Power of (and Power within) Government ... 54
5 The Power Imbalances Due to ADM Scale, Speed and Depth ... 56
6 Why Risk-Based Frameworks Misread the Welfare Domain ... 57
7 Harms-Based Regulation and Enforcement ... 59
8 Why Front End Preventive Measures Are Problematic for Government Human Services Mass Decision-Making ... 60
VI Conclusion ... 62

I Introduction

Technological innovation in government administration has surged in recent times. Digitisation of government records and data, automation of decision-making in government administration in place of human decisions, and government use of both basic and increasingly sophisticated forms of artificial intelligence ('AI') such as ChatGPT and chatbot interfaces for dealing with the

public, are increasingly deployed by all levels of Australian government.[1] A key driver of government adoption of automated decision-making and artificial intelligence is efficiency and restraint on fiscal outlays.

This article concentrates mainly on federal government use of *automated* decision-making ('ADM') to effectively *make* human services decisions affecting citizens. It largely applies the definition of ADM devised by Rashida Richardson for 'prescriptive' (ie regulatory) purposes as 'any systems, software, or processes that use computation to … replace government decisions, judgments, and/or policy implementation that impact opportunities, access, liberties, rights, and/or safety', modified to exclude automation that 'aid[s]' but does not effectively 'replace' human decision-making.[2] AI, by contrast, is defined as any 'engineered system that generates predictive outputs such as content, forecasts, recommendations or decisions for a given set of human-defined objectives or parameters without explicit programming'.[3]

Human services account for a high proportion of government outlays (around a third in Australia),[4] so social security has been an early target for digital transformation initiatives in many countries.[5] Several of these initiatives

[1] For a recent New South Wales analysis, see, eg, Luis Lozano-Paredes, 'Mapping AI's Role in NSW Governance: A Socio-Technical Analysis of GenAI Integration' (2025) 7 *Frontiers in Political Science* 1595345:1–17. See generally Yee-Fui Ng and Matthew Groves, *Automation in Governance: Theory, Practice and Problems* (Hart Publishing, 2025); Ryan Calo and Danielle Citron, 'The Automated Administrative State: A Crisis of Legitimacy' (2021) 70(4) *Emory Law Journal* 797; Alexandra Sinclair, 'Automated Decision-Making and Government Opacity' [2025] (January) *Public Law* 8.

[2] Rashida Richardson, 'Defining and Demystifying Automated Decision Systems' (2022) 81(3) *Maryland Law Review* 785, 795–6. A wider formulation is proposed as the basis for defining the mandates of advisory or investigative bodies: at 795. See also Anna Katharina Boos, 'Conceptualizing Automated Decision-Making in Organizational Contexts' (2024) 37(3) *Philosophy and Technology* 92:1–30. Boos discusses the role of 'intention' as part of a 'decision': at 6–8, 11–13; and refines the definition of ADM 'as the practice of using algorithms to solve decision problems, where these algorithms can play a suggesting, offloading, or superseding role relative to humans, and decisions are defined as action triggering choices': at 23 (emphasis omitted). The intersection of AI with ADM is discussed in Senate Select Committee on Adopting Artificial Intelligence, Parliament of Australia, *Final Report* (Report, November 2024) ch 5 ('*Senate Committee on AI*').

[3] *Senate Committee on AI* (n 2) 4 [1.21]. The definition goes on to add that 'AI systems are designed to operate with varying levels of automation, and include many relatively commonplace systems' and lists examples.

[4] Commonwealth, *Budget Strategy and Outlook* (Budget Paper No 1, 12 May 2026) 215.

[5] Alexandra James and Andrew Whelan, '"Ethical" Artificial Intelligence in the Welfare State: Discourse and Discrepancy in Australian Social Services' (2022) 42(1) *Critical Social Policy* 22,

gained notoriety internationally for the harm inflicted on vulnerable populations (termed 'dystopia[n]' by some)[6] and the multiple failures of systems of political and legal accountability.[7] Countries differ in the weight accorded to three different arenas for delivering administrative accountability, namely preventive protections (such as the United States ('US') open hearings on making new administrative rules);[8] review of and remaking of all aspects of challenged decisions (termed 'merits review') by administrative tribunals; and redress through the courts.[9] Australia relies most heavily on merits review,[10] but all three have foundered as a response to similar ADM in welfare challenges internationally.[11]

My aim in what follows is to assess whether recent reforms to Australia's federal accountability and regulatory machinery are equipped to cope with

29. Although the literature concentrates on Western countries, the same issues arise in the global South: see generally Sebastian Smart, 'Algorithmic Discrimination in Latin American Welfare States' (Carr Center Discussion Paper No 2024-05, Carr-Ryan Center for Human Rights Policy, Harvard University, 26 August 2024) <https://www.hks.harvard.edu/centers/carr-ryan/publications/algorithmic-discrimination-latin-american-welfare-states>, archived at <https://perma.cc/97QC-CWVE>.

6 Philip Alston, *Report of the Special Rapporteur on Extreme Poverty and Human Rights*, UN Doc A/74/493 (11 October 2019) 5 [6], 5 [8]; Gijsbert Vonk, 'Welfare State Dystopia as a Challenge to the Fundamental Right of Social Security' (2025) 27(2) *European Journal of Social Security* 82, 84–5.

7 Maarten Bouwmeester, 'System Failure in the Digital Welfare State: Exploring Parliamentary and Judicial Control in the Dutch Childcare Benefits Scandal' (2023) 44(2) *Recht der Werkelijkheid* [Journal of Empirical Research on Law in Action] 13, 13–14; Lucas Michael Haitsma and Maarten Bouwmeester, 'Learning from Control Deficits in the Childcare Benefits Scandal: A Plea for Multi-Level Analysis in Law and Policy Research' (2023) 44(3) *Recht der Werkelijkheid* [Journal of Empirical Research on Law in Action] 57, 65; Maarten Bouwmeester, 'Checks and Balances under Pressure in the Welfare State: Disentangling the Rule of Law Risks of Automation and Welfare Conditionality' (2025) 27(2) *European Journal of Social Security* 98, 99 ('Checks and Balances').

8 See *Administrative Procedure Act*, 5 USC 553, 556–7 (2024).

9 See generally Michael Asimow, 'Five Models of Administrative Adjudication' (2015) 63(1) *American Journal of Comparative Law* 3; Michael Asimow, 'A Comparative Approach to Administrative Adjudication' in Peter Cane et al (eds), *The Oxford Handbook of Comparative Administrative Law* (Oxford University Press, 2021) 577 ('Comparative Administrative Adjudication').

10 Asimow, 'Comparative Administrative Adjudication' (n 9) 587.

11 See Bouwmeester, 'Checks and Balances' (n 7) 99. These failures are caused by limited rule of law, proportionality and democratic accountability in the ADM development stage; post-implementation scrutiny being trumped by a lack of transparency; 'the overwhelming political pressures in favour of welfare conditionality' reforms; the often limited interest to carefully scrutinise government algorithms; complex rules and redress procedures; and a lack of legal aid at the *ex post* stage: at 106–7.

future government ADM challenges, and, if not, what additional measures should be considered. I conclude that adequate pathways for rectification of individual grievances due to ADM have been restored, but to the neglect of measures to prevent or respond to systemic harm and the more powerful *structural* sources of ADM harm in human services. I then make a case for a broader set of measures — ones better attuned to the nature of bureaucracy, the technology at hand and the power asymmetries associated with ADM deployment within government. Drawing attention to research on the limits of orthodox legal remedies operating in isolation from political and managerial controls (including infrastructural linkages), I also touch on possible ways of mitigating against the release of poorly designed programs, including new regulatory bodies and proactive forms of oversight of ADM in government.

II Robodebt and Reflections on Its Causes and Prevention

A *Robodebt's Legal and Administrative Failings*

Australia's Robodebt proved to be 'a massive failure of public administration',[12] an abuse of political leadership and, most worryingly of all, an almost total breakdown of the institutional checks and balances of tribunals, ombudsman offices and other accountability machinery. The scheme raised 794,000 legally unfounded and false debts against 526,000 citizens, with $1.751 billion ultimately required to be repaid or written off,[13] rising to $2.4 billion following a further 2025 settlement costing $548.5 million.[14]

The illegality in Robodebt involved using average fortnightly income in place of actual fortnight by fortnight income, reversal of the onus of proof, and failure to apply the correct standard of proof such that there was no probative

[12] *Prygodicz v Commonwealth [No 2]* (2021) 173 ALD 277, 280 [5] (Murphy J) ('*Prygodicz*'), quoted in *Royal Commission into the Robodebt Scheme* (Report, July 2023) 25 ('*Robodebt Royal Commission*').

[13] *Robodebt Royal Commission* (n 12) 402, 410. The most accessible comprehensive account is Rick Morton, *Mean Streak: A Moral Vacuum and a Multi-Billion Dollar Government Shake Down* (Fourth Estate, 2024).

[14] Samantha Dick, 'Robodebt Victims to Get Further $475 Million Compensation after Class Action Settled', *ABC News* (online, 4 September 2025) <https://www.abc.net.au/news/2025-09-04/robodebt-victims-get-compensation-from-class-action/105734030>, archived at <https://perma.cc/4DVH-38CZ>. See below n 67 and accompanying text.

evidence to ground debt decisions,[15] making those decisions legally 'irrational'.[16] The mathematical (and factual) falsity was assuming an average can reliably speak to its constituent *fortnightly* parts, or that clients had continuous and stable fortnightly incomes (an assumption contradicted by data already held by the relevant government departments).[17] For present purposes, however, it was the *automation* of the raising of advice of overpayments and lack of human decision-makers that constituted Robodebt as a form of ADM, restricting citizens to dealing with what government initially termed its 'Online Compliance Intervention'.[18] It was the failure of so many accountability avenues that rendered it an emblematic example of why reform was needed.

B *Research Findings on Robodebt*

Robodebt has spawned a vast literature. As now briefly sketched, much of the commentary is rather discipline-specific, with writers sharing many conclusions about the core harms and concerns but more divided about the respective contributions of legal, administrative/bureaucratic, and socio-technical/cultural considerations. Only a few dominant unifying themes emerge across disciplines.[19]

[15] Terry Carney, 'The New Digital Future for Welfare: Debts without Legal Proofs or Moral Authority?' [2018] *University of New South Wales Law Journal Forum* 1:1–16, 3–8 ('New Digital Future'). See also Matthew Groves, 'Systemic Failure and Systemic Unfairness' (2023) 30(3) *Australian Journal of Administrative Law* 209, 227 ('Systemic Failure').

[16] Order of Davies J in *Amato v Commonwealth* (Federal Court of Australia, VID611/2019, 27 November 2019) 5–6 [8]–[9] ('*Amato*').

[17] For summaries of the design failures, see Peter Whiteford, 'Why Robodebt's Use of "Income Averaging" Lacked Basic Common Sense', *Austaxpolicy* (Blog Post, 24 March 2023) <https://www.austaxpolicy.com/why-robodebts-use-of-income-averaging-lacked-basic-common-sense/>, archived at <https://perma.cc/9DY5-VWRM>. For more detailed analysis, see Peter Whiteford, 'Debt by Design: The Anatomy of a Social Policy Fiasco' (2021) 80(2) *Australian Journal of Public Administration* 340, 345–8; Katina Michael, 'Algorithmic Bias: Australia's Robodebt and Its Human Rights Aftermath' (2024) 5(3) *IEEE Transactions on Technology and Society* 254, 257–8.

[18] See Andrew Whelan, '"Ask for More Time": Big Data Chronopolitics in the Australian Welfare Bureaucracy' (2020) 46(6) *Critical Sociology* 867, 871–8.

[19] The concept of vulnerability is one such theme: see generally Terry Carney, 'Vulnerability: False Hope for Vulnerable Social Security Clients?' (2018) 41(3) *University of New South Wales Law Journal* 783; Terry Carney, 'Artificial Intelligence in Welfare: Striking the Vulnerability Balance?' (2020) 46(2) *Monash University Law Review* 23 ('AI in Welfare'); Yee-Fui Ng and Stephen Gray, 'Disadvantage and the Automated Decision' (2022) 43(2) *Adelaide Law Review* 641; Maria

Legal commentary, for example, has highlighted failures of accountability machinery,[20] technical limitations to transparency [21] and the scope of what constitutes a 'decision',[22] the role of judicial review and merits review,[23] and the outsourcing of government functions.[24] Reforms canvassed include adoption of the English 'duty to consult',[25] realisation of an entitlement to a 'well-calibrated' decision,[26] human oversight of adverse decisions,[27] and legal aid for strategic litigation.[28]

Governance and administration scholars, for their part, have written about the loss of integrity and trust in government,[29] pointing to various failings such

O'Sullivan, 'The Vulnerability of Social Security Recipients in the Age of Automation: The Role of Public Law' [2024] (April) *Public Law* 203. For a broader discussion about ADM and vulnerability, see generally Antonia Stanojević, 'Algorithmic Governance and Social Vulnerability: A Value Analysis of Equality and Trust' in Jurgen Goossens, Esther Keymolen and Antonia Stanojević (eds), *Public Governance and Emerging Technologies: Values, Trust, and Regulatory Compliance* (Springer, 2025) 335.

20 See, eg, Peter Hanks, 'Administrative Law and Welfare Rights: A 40-Year Story from *Green v Daniels* to "Robot Debt Recovery"' (2017) 89 *AIAL Forum* 1, 9–10; Carney, 'New Digital Future' (n 15) 3, 8–9; Terry Carney, 'Robodebt Illegality: The Seven Veils of Failed Guarantees of the Rule of Law?' (2019) 44(1) *Alternative Law Journal* 4, 4–8 ('Robodebt Illegality').

21 Yee-Fui Ng et al, 'Revitalising Public Law in a Technological Era: Rights, Transparency and Administrative Justice' (2020) 43(3) *University of New South Wales Law Journal* 1041, 1052.

22 See generally Yee-Fui Ng and Maria O'Sullivan, 'Deliberation and Automation: When Is a Decision a Decision?' (2019) 26(1) *Australian Journal of Administrative Law* 21 ('Deliberation and Automation').

23 See Joel Townsend, 'Better Decisions? Robodebt and Failings of Merits Review' in Janina Boughey and Katie Miller (eds), *The Automated State: Implications, Challenges and Opportunities for Public Law* (Federation Press, 2021) 52, 63–9 ('Better Decisions'); Jack Maxwell, 'Judicial Review and the Digital Welfare State in the UK and Australia' (2021) 28(2) *Journal of Social Security Law* 94, 105–8.

24 Lucinda O'Brien et al, 'Government Debt Collection after Robodebt' (2024) 52(3) *Australian Business Law Review* 130, 130–1, 148–57.

25 Groves, 'Systemic Failure' (n 15) 215–19.

26 Benedict Sheehy and Yee-Fui Ng, 'The Challenges of AI Decision-Making in Government and Administrative Law: A Proposal for Regulatory Design' (2024) 57(3) *Indiana Law Review* 665, 688–9, quoting Aziz Z Huq, 'A Right to a Human Decision' (2020) 106(3) *Virginia Law Review* 611, 619.

27 Sheehy and Ng (n 26) 687.

28 Joel Townsend, 'Challenging a Repeat Player: Robodebt and the Role of Legal Assistance' (2023) 30(3) *Australian Journal of Administrative Law* 181, 188–91 ('Repeat Player').

29 Valerie Braithwaite, 'Beyond the Bubble That Is Robodebt: How Governments That Lose Integrity Threaten Democracy' (2020) 55(3) *Australian Journal of Social Issues* 242, 248–56.

as an '"overly responsive" public service culture' (termed 'useful idiots'),[30] themes of bureaucratic organisational 'evil' and Kafkaesque corruption and dystopia,[31] or neoliberal governance as surveillance.[32] Socio-technical systems theorists have discussed Robodebt's faulty development and exclusion of key interests, such as those of users,[33] and the way exclusion and oppression was built into the scheme.[34]

More recently, an interesting body of interdisciplinary, theoretical and reflective literature has begun to emerge.

[30] Paul WF Henman, 'Robodebt Cultures and Useful Idiots: Why Robodebt Was Not a Techno-Failure' (2025) 60(1) *Australian Journal of Social Issues* 4, 11–13, quoting Andrew Podger, *Report to the Royal Commission into the Robodebt Scheme* (Report, 20 February 2023) 3 <https://robodebt.royalcommission.gov.au/publications/andrew-podger-ao-report-robodebt-royal-commission>, archived at <https://perma.cc/K34Y-W7TF> ('*Report to Royal Commission*').

[31] See Adam Graycar and Adam B Masters, 'Bureaucratic Bastardry: Robodebt/Debt Recovery, AI and the Stigmatisation of Citizens by Machines and Systems' (2022) 16(5–6) *International Journal of Public Policy* 333, 340–1; Samuel Naylor, 'Robodebt, Kafka and Institutional Absurdism in Australia' (2023) 48(4) *Alternative Law Journal* 299.

[32] Monique Mann, 'Technological Politics of Automated Welfare Surveillance: Social (and Data) Justice through Critical Qualitative Inquiry' (2020) 1(1) *Global Perspectives* 12991:1–12, 4, 6; Mona Nikidehaghani, Jane Andrew and Corinne Cortese, 'Algorithmic Accountability: Robodebt and the Making of Welfare Cheats' (2023) 36(2) *Accounting, Auditing and Accountability Journal* 677, 684–5.

[33] Roger Clarke, Katina Michael and Roba Abbas, 'Robodebt: A Socio-Technical Case Study of Public Sector Information Systems Failure' (2024) 28 *Australasian Journal of Information Systems*:1–42, 25–6. See also Michael D'Rosario and Carlene D'Rosario, 'Beyond Robodebt: The Future of Robotic Process Automation' (2020) 11(2) *International Journal of Strategic Decision Sciences* 1, 17–19. The unintended consequences of flawed implementation are outlined in Tapani Rinta-Kahila et al, 'Managing Unintended Consequences of Algorithmic Decision-Making: The Case of Robodebt' (2024) 14(1) *Journal of Information Technology Teaching Cases* 165, 166–7.

[34] Sora Park and Justine Humphry, 'Exclusion by Design: Intersections of Social, Digital and Data Exclusion' (2019) 22(7) *Information, Communication and Society* 934, 940–4.

C *Framings Pertinent to Risks of ADM in Human Services*

Four framing ideas from the literature are salient to the present inquiry.

1 *Infrastructural Relationships of Law, Bureaucracy and Politics*

Social theorist Robert van Krieken highlighted what he termed a 'systemic, wilful ignorance in the relationship between law, bureaucracy and politics'.[35] Yee-Fui Ng's important legal monograph responded with a proposed synthesis of all three fields by way of a new tripartite combination of legal, political and managerial controls.[36] Likewise, Jennifer Raso's plea was for what she termed an 'infrastructural approach',[37] having earlier called for 'new justice models centred on the perspectives of system users *within and outside of administrative institutions*'.[38] As Raso elaborates on the infrastructural approach:

> In an era of digital government, responsible AI and administrative law communities must reconceptualize AI systems and the decision-making processes to which they contribute infrastructurally. By 'infrastructurally,' I mean as distributed systems that generate results based on the inputs and interactions of many actors: tools, databases, workers, members of the public, and institutions.[39]

As shown later in this paper, the wide-angle lens of an infrastructural approach to accountability has considerable appeal. However, three other theoretical framings also assist in understanding how Robodebt exposed failures of both individual and systemic justice.

[35] Robert van Krieken, 'The Organization of Ignorance: The Australian "Robodebt" Affair, Bureaucracy, Law and Politics' (2024) 50(7–8) *Critical Sociology* 1379, 1379.

[36] Yee-Fui Ng, *Combatting the Code: Regulating Automated Government Decision-Making in Comparative Context* (Cambridge University Press, 2025) ch 9.

[37] Jennifer Raso, 'Responsible AI: Binaries That Bind' (2024) 69(4) *McGill Law Journal* 417, 436–9 ('Responsible AI'). For an equally jaundiced view of the ability to control the commercial market and associated power of internationally pervasive conglomerate AI companies and platforms through traditional legal principles, see Monika Zalnieriute, 'Against Procedural Fetishism in the Automated State' in Zofia Bednarz and Monika Zalnieriute (eds), *Money, Power and AI: Automated Banks and Automated States* (Cambridge University Press, 2024) 221.

[38] Jennifer Raso, 'Implementing Digitalisation in an Administrative Justice Context' in Marc Hertogh et al (eds), *The Oxford Handbook of Administrative Justice* (Oxford University Press, 2022) 521, 534 (emphasis added) ('Implementing Digitalisation').

[39] Raso, 'Responsible AI' (n 37) 436.

2 *Administrative Burden*

The second interesting idea is the concept of administrative burden. The administrative burden experienced in accessing government largesse is advanced as an explanation for depressing levels of access to, and distortion in the distribution of, public goods such as social security and administrative justice.[40] Administrative burden was originally conceptualised as the difficulties associated with learning about, complying with, and the psychological costs of engaging with, a state program.[41] To this, redemptive and correction of error burdens have more recently been added.[42] These burdens may be unintended or a deliberate choice (such as welfare activation or mandatory internal reviews before external merits review can be accessed).[43] But such burdens also 'have *distributive consequences* and are likely to fall harder on those with fewer resources in the form of human and administrative capital, administrative literacy, and bureaucratic self-efficacy'.[44] Burdens are accentuated when services are contracted out to private sector providers (under neoliberal principles of what is termed 'new public management')[45] or when welfare 'activation' imposes extensive requirements as a condition of continued qualification for human services (termed 'virtual welfare').[46] These access burdens and consequences have

[40] For a recent literature review, see Aske Halling and Martin Baekgaard, 'Administrative Burden in Citizen–State Interactions: A Systematic Literature Review' (2024) 34(2) *Journal of Public Administration Research and Theory* 180, 181. A nuanced refinement, drawing on the work of Michael Lipsky and others, is given in Victoria Adelmant and Jennifer Raso, 'Data Entry and Decision Chains: Distributed Responsibility and Bureaucratic Disempowerment in the UK's Universal Credit Programme' (2025) 45(2) *Oxford Journal of Legal Studies* 415, 418–21.

[41] Halling and Baekgaard (n 40) 181.

[42] Pierre-Marc Daigneault, 'Reconceptualizing Administrative Burden around Onerous Experiences' (2024) 7(4) *Perspectives on Public Management and Governance* 124, 125.

[43] Halling and Baekgaard (n 40) 181.

[44] Ibid (emphasis added) (citations omitted). For a discussion of the way digitally 'linked' infrastructures in the United Kingdom's ('UK') Universal Credit scheme diffuses and shifts administrative burdens onto outlier public services (such as the tax office) or private actors (such as employers reporting their employees earnings), see Adelmant and Raso (n 40) 416–17, 428–31.

[45] Carney, 'AI in Welfare' (n 19) 37–40.

[46] Terry Carney, 'Automation and Conditionality: Towards "Virtual" Social Security?' (2024) 31(1) *Journal of Social Security Law* 32, 40–1 ('Automation and Conditionality'); John Morrison, 'Digital Conditionality: Universal Credit and Its Implications for Social Citizenship' (2025) 47(2–3) *Journal of Social Welfare and Family Law* 320, 323–4.

already been demonstrated in Australian social security,[47] though not yet for accessing administrative justice.

3 *Algorithmic Justice*

The third idea offering potential insights into the infrastructural or systemic risks of ADM in government and how these might be rectified is the concept of algorithmic justice. Initially developed as a banner for public policy advocacy purposes, algorithmic justice was just a neat rhetorical way of bundling for popular consumption the traditional concerns about ADM such as bias, lack of agency or choice, inadequate transparency and the absence of accountability.[48] But this short-changes its wider potential.

A recent analysis contends that algorithmic justice is much more capacious in at least three aspects.[49] The first highlighted aspect is its focus on *broader AI harms* and an accompanying theory of justice going beyond traditional equal distributional access to social goods.[50] The second lies in its concentration on *systemic* as distinct from merely *individual* impacts, picking up not only economic but also sociocultural and political dimensions of ADM. The final appealing feature is its interest in the algorithmic *processes* rather than just various biases or distortions in data sets.[51]

Algorithmic justice conceived in this fashion resonates with the wider-angle lens of whole systems or infrastructural approaches to ADM regulation.

[47] On cashless welfare, see Shelley Bielefeld, 'Administrative Burden and the Cashless Debit Card: Stripping Time, Autonomy and Dignity from Social Security Recipients' (2021) 80(4) *Australian Journal of Public Administration* 891, 895–906. On welfare activation, see Michael McGann and Sarah Ball, 'Beyond Administrative Burden: Activation and Administrative Harm' (2025) 60(2) *Australian Journal of Social Issues* 441, 448. On the administrative burden of responsibilisation, see generally Jeremiah Thomas Brown, 'Citizens Putting In the Work: How Administrative Work Is Shifting in the Digital Welfare State' (2025) 47(2–3) *Journal of Social Welfare and Family Law* 280.

[48] An example is the non-profit Algorithmic Justice League, founded in 2016 in Cambridge, Massachusetts: 'About', *Algorithmic Justice League* (Web Page, 2025) <https://www.ajl.org/about>, archived at <https://perma.cc/KE2Q-U23N>.

[49] Olivera Marjanovic, Dubravka Cecez-Kecmanovic and Richard Vidgen, 'Theorising Algorithmic Justice' (2022) 31(3) *European Journal of Information Systems* 269.

[50] Ibid 273–6.

[51] Ibid 271–3, 278–81. For instance, focusing solely on a set of data points may result in stripped down 'datafied humans' being constructed as the subjects of ADM, replacing human-to-human engagement: at 281. For a critique based on a failure to consider justice for non-human actors, see Joshua C Gellers, 'AI, Design, and More-than-Human Justice' in David J Gunkel (ed), *Handbook on the Ethics of Artificial Intelligence* (Edward Elgar Publishing, 2024) 128.

4 *Differential Harms Specific to Welfare*

The final notable idea is that of the differential harms of ADM. A 'harms' approach, rather than the more popular 'risk' approach, to regulation has recently been advocated as better able to capture the ubiquity, intangibility (which makes them more easily overlooked) and aggregation (incremental cumulation) of adverse consequences of ADM.[52]

Importantly, a harms lens also captures the adverse lived experiences arising from the interaction of ADM with the 'societal power dynamics'[53] of the community at large or the *specific* context in which ADM is deployed. However, the potential harms specific to or *emblematic* of ADM use in welfare administration have been neglected until recently.

The special *personal* vulnerabilities of welfare clients are the most obvious reasons why ADM may be more problematic in this setting than for other areas of public administration,[54] but it is more complicated than that. One complication is the entanglement of ADM with digitalisation and e-governance: the so-called 'digital divide' of lack of access to smartphones or other hardware, or lack of competence in navigating e-governance software interfaces, exists

[52] Sylvia Lu, 'Regulating AI Harms' (2026) 78(1) *Florida Law Review* 323, 334–40. Taxonomies of the range of different harms associated with ADM/AI are not new, but most of the previous literature has concentrated on the front-end issue of promoting less harmful and technically better designed systems: see generally Gavin Abercrombie et al, 'A Collaborative, Human-Centred Taxonomy of AI, Algorithmic, and Automation Harms' (Working Paper, 9 November 2024) <https://arxiv.org/pdf/2407.01294>, archived at <https://perma.cc/Q7ZC-4G67>. Gavin Abercrombie et al describe the construction of regulatory frameworks in passing as the 'governance debate that should rightfully follow': at 2.

[53] Renee Shelby et al, 'Sociotechnical Harms of Algorithmic Systems: Scoping a Taxonomy for Harm Reduction' in Association of Computing Machinery (ed), *AIES '23: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (2023) 723, 723, citing Ben Green and Salomé Viljoen, 'Algorithmic Realism: Expanding the Boundaries of Algorithmic Thought' in Association for Computing Machinery (ed), *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (2020) 19.

[54] In a study of perceived impacts of possible design enhancements for existing online clients of the UK Universal Credit scheme, Colin Watson et al use the term 'harms' to cover the way e-governance exacerbates social exclusion of vulnerable welfare clients: Colin Watson et al, 'Hostile Systems: A Taxonomy of Harms Articulated by Citizens Living with Socio-Economic Deprivation' (Conference Paper No 528, CHI Conference on Human Factors in Computing Systems, 11–16 May 2024) 1. They add:

> We use the term 'harms' … [to] include everything disadvantageous to citizens such as increased time, more effort, greater financial cost, loss of reputation, extra discrimination, additional mental stress, diminished knowledge, curtailed capacity, added neglect, and loss of rights.

irrespective of ADM use.[55] But that harm is amplified by the lack of any (or any easy) access to a *human* administrator or external sources of support when navigating access to welfare or ensuring ongoing compliance.[56]

Relevantly to the present paper, such harms have been conceptualised as new forms of administrative burdens, as in the ethnographic study of welfare by Ida Heggertveit and Hanne Höglund Rydén in Norway and the study by Colin Watson, Adam Parnaby and Ahmed Kharrufa in the United Kingdom ('UK').[57] Other harms differentially experienced by people reliant on welfare include heightened perceptions of stigma associated with ADM-generated debts, loss or reduction of payments that immediately imperil household budgets[58] and the magnification of the speed and scale of harms due to ADM mass decision-making in welfare.[59]

III A Post-Robodebt Thought Experiment

The renovation of accountability machinery recommended by the Royal Commission into the Robodebt Scheme ('Royal Commission') has now largely been

[55] See Carney, 'Automation and Conditionality' (n 46) 33. The differential distribution of 'social capital' and the increasing demands on social capital as the welfare state increases the conditionality of welfare and outsources to clients the task of ensuring compliance, is an alternative conceptualisation: see generally Stephen Baron, John Field and Tom Schuller (eds), *Social Capital: Critical Perspectives* (Oxford University Press, 2000); Peter Billings, 'Income Management in Australia: Protecting the Vulnerable and Promoting Human Capital through Welfare Conditionality' (2011) 18(4) *Journal of Social Security Law* 167.

[56] Carney, 'Automation and Conditionality' (n 46) 33.

[57] Ida Heggertveit and Hanne Höglund Rydén, 'Narratives of Reality: Administrative Burdens and Workarounds in Digital Self-Services' (Research Paper No 13, Americas Conference on Information Systems, 2024) 3 <https://aisel.aisnet.org/amcis2024/soc_inclusion/social_inclusion/13>, archived at <https://perma.cc/45PH-DTLL>; Colin Watson, Adam W Parnaby and Ahmed Kharrufa, 'Precarious Experiences: Citizens' Frustrations, Anxieties and Burdens of an Online Welfare Benefit System' (Working Paper, 14 May 2024) 3 <https://arxiv.org/pdf/2405.08515>, archived at <https://perma.cc/37RD-SWKT>. For an analysis of ADM in Finland's immigration system, see also Hanna Maria Malik and Nea Lepinkäinen, 'Between Algorithmic and Analogue Harms: The Case of Automation in Finnish Immigration Services' (2022) 5(3) *Justice, Power and Resistance* 270.

[58] *Robodebt Royal Commission* (n 12) 330–2.

[59] Albert Sanchez-Graells, 'Resh(AI)ping Good Administration: Addressing the Mass Effects of Public Sector Digitalisation' (2024) 13(1) *Laws* 9:1–15, 7–9. The taxonomies of harm proposed by Watson et al and Sylvia Lu are discussed later: see below Part V(C)(7).

completed.[60] So how well would it cope with a fresh ADM challenge? To answer this question I will conduct a thought experiment to stress test legal and associated measures adopted by the Australian government. These measures were designed to avoid or mitigate harms from similar ADM public policy disasters — particularly in mass decision-making areas such as human services and social security.[61]

A thought experiment, as a scientific methodology, is 'a way for inquirers to focus the power of the imagination. … thought experiments are like laboratory experiments, except they are carried out in the imagination'.[62] The method has ancient lineage across most academic fields, though only labelled as such in 1811 in describing the work of Immanuel Kant.[63] Although implicit in the reasoning of legal philosophers such as John Rawls (his 'original position'), explicit adoption of the method is less common in law.[64] Michael Stuart explains the method which I now apply: '[u]sually there are several phases in a thought experiment: one in which we set up some imaginary scenario, another in which we "see" what happens in that scenario, and, finally, one in which we draw some conclusions'.[65]

A *Background*

It was a stretch even to characterise Robodebt as *any* form of ADM. Even as automated decision-making, it was little more than full automation of a pre-existing process (matching Australian Taxation Office and Centrelink data) by

[60] See Department of the Prime Minister and Cabinet, *Robodebt Royal Commission Implementation Update* (Web Document, March 2026) <https://www.pmc.gov.au/resources/government-response-royal-commission-robodebt-scheme>, archived at <https://perma.cc/A7LU-UBNY>.

[61] See, eg, Commonwealth of Australia, *Parliamentary Debates*, House of Representatives, 10 October 2024, 7060 (Mark Dreyfus, Attorney-General).

[62] Michael T Stuart, 'Thought Experiments' in Vlad Petre Glăveanu (ed), *The Palgrave Encyclopedia of the Possible* (Palgrave Macmillan, 2022) 1644, 1644. See also Michael T Stuart, 'Sharpening the Tools of Imagination' (2022) 200(6) *Synthese* 451:1–22, 4–5. For discussion of the synergies and differences between thought experiments and 'models', see generally Rawad El Skaf and Michael T Stuart, 'Scientific Models and Thought Experiments' in Tarja Knuuttila, Natalia Carrillo and Rami Koskinen (eds), *The Routledge Handbook of Philosophy of Scientific Modeling* (Routledge, 2025) 325.

[63] Stuart, 'Thought Experiments' (n 62) 1645.

[64] Gabriel Doménech-Pascual, 'Thought Experiments in Law' [2021] *Law and Method*:1–21, 3–4, 6.

[65] Stuart, 'Thought Experiments' (n 62) 1644.

withdrawing human investigation and oversight.[66] Significantly, Robodebt was very belatedly brought undone after surviving for more than three years, but only by the courts on a point of law.[67] Merits review was gamed to avoid accountability and ombudsman machinery failed to scrutinise legality,[68] Ministers deflected pressure from social media campaigns such as '#NotMyDebt',[69] and parliamentary committees lacked the guidance from expert submissions or the staffing capacity to unearth the fatal legal flaw.[70] How much longer might Robodebt have continued to evade being brought to book had it not been for judicial review (in itself intolerably slow)?[71] In any future scenario lacking the ingredient of illegality (for example, where there is a mathematical or other blatant factual error), would this once again mean that there would be no *realistic* avenue for ensuring accountability on the merits or for breach of ethical or governance standards?

One way of stress testing against a future government ADM failure is to postulate a future program with all of Robodebt's features except for its illegality. A second is to consider how an ADM program would fare if it was instead a sophisticated machine learning or large language model with similarly politically appealing deliverables. How would or should a reformed system of accountability cope with such a challenge? Against the legal architecture that predated the suite of Royal Commission reforms, it is doubtful that either scheme would have ended short of a change of government.

66 *Robodebt Royal Commission* (n 12) xxiv–xxv.

67 Robodebt was found to be unlawful in *Amato* (n 16) 1 [1]–[2]. This was followed by a successful class action to obtain compensation for the large numbers of people adversely impacted by the scheme: see *Prygodicz* (n 12) 279–82 [1]–[14] (Murphy J). On 23 June 2026, the Federal Court approved the settlement of an appeal against the inadequacy of the original compensation, adding a further $475 million: Order of Beach J in *Prygodicz v Commonwealth* (Federal Court of Australia, VID1252/2019, 23 June 2026) 4 [14].

68 Carney, 'Robodebt Illegality' (n 20) 5–6.

69 See ibid 8; Kate Galloway, 'Big Data: A Case Study of Disruption and Government Power' (2017) 42(2) *Alternative Law Journal* 89, 93–4.

70 Carney, 'Robodebt Illegality' (n 20) 6.

71 Robodebt took years to be brought to heel despite the introduction of legal aid commissions: see, eg, *Victoria Legal Aid Act 1978* (Vic); and reforms such as the *Administrative Decisions (Judicial Review) Act 1977* (Cth). But as Peter Hanks noted, an equivalent ruling by the High Court in 1977 in *Green v Daniels* (a case about the Commonwealth denying claims for unemployment payments over the summer break until the commencement of the next school year) was achieved in a matter of months, *well prior* to enactment of the much touted 'new administrative law' reform: Hanks (n 20) 2–6, discussing *Green v Daniels* (1977) 13 ALR 1. See below n 80 and accompanying text.

B *Stress Testing a Flawed Post-Robodebt ADM Initiative*

How would a newly problematic ADM welfare scheme fare in light of the post-Robodebt reforms? The answer turns on whether affected citizens know ADM is involved, whether they have access to an avenue of review for their individual grievances and whether correction of individual errors results in systemic rectification for all other affected social security clients — especially given the vested interests of public servants and ministers in retaining the fiscal savings that commonly motivate such government ADM initiatives.

1 *How Would Citizens Discover the Use of ADM, as a Potential Flaw?*

By February 2017 some (successful) Administrative Appeals Tribunal ('AAT') applicants, and by March 2018 every interested Australian, knew that a crude version of ADM was in play regarding their grievances about wrong or false debts. However, the details of the Robodebt methodology remained opaque due to the way its actual workings were kept secret, as the Royal Commission noted.[72]

As reforms currently stand, the working of any such future ADM scheme would also remain secret. Transparency guarantees remain inadequate. While ADM can be defined in differing ways,[73] government ADM was intended to be dealt with separately from transparency statement obligations originally proposed to be imposed under Australia's AI regulatory framework.[74] If, like Robodebt, crucial details of the workings of a future newly problematic ADM scheme were confined to cabinet documents, they would remain exempt from freedom of information ('FOI') disclosure since the government has rejected the Royal Commission's recommendation to repeal the automatic exemption

[72] *Robodebt Royal Commission* (n 12) 370–1; Evidence to Royal Commission into the Robodebt Scheme, 24 February 2023, 3832 (Peter Hanks) <https://robodebt.royalcommission.gov.au/system/files/2023-02/transcript-hearing-day-36-24-february-2023.pdf>, archived at <https://perma.cc/227E-JLA2> ('Hanks' Evidence to Royal Commission'). Peter Hanks KC's evidence is summarised in Rick Morton, 'February 24: Robodebt Royal Commission', *The Saturday Paper* (Blog Post, 24 February 2023) <https://www.thesaturdaypaper.com.au/news/2023/02/24/february-24-robo-debt-royal-commission>, archived at <https://perma.cc/6T3B-CHUQ>; Morton (n 13) 388–9.

[73] See above n 2 and accompanying text.

[74] See, eg, Department of Industry, Science and Resources (Cth), *Safe and Responsible AI in Australia: Proposals Paper for Introducing Mandatory Guardrails for AI in High-Risk Settings* (Proposals Paper, September 2024) 8 <https://consult.industry.gov.au/ai-mandatory-guardrails>, archived at <https://perma.cc/NJ58-WWJL>. For discussion of the then proposed AI framework, see Henry Fraser et al, 'Should Australia Follow Europe's Approach to AI Standards and Regulation?' (2024) 5(1) *ANU Journal of Law and Technology* 96, 97.

applied to cabinet documents.[75] Outside that, the 2024 amendment to Privacy Principle 1 of the *Privacy Act 1988* (Cth) might now require reporting its *use*, but not its workings.[76] The obligation imposed on government departments to publicise their use of ADM under the transparency requirements sponsored by the Digital Transformation Agency[77] lacks operational standardisation or a common portal, and has been found to be patchily honoured to date.[78] Otherwise, while it is true that an FOI request would confirm the use of ADM, this relies on someone having sufficient knowledge of possible use to warrant making that request for disclosure.

2 *How Could Individuals Review Flawed ADM?*

This is an important question because merits review by the Administrative Review Tribunal ('ART') is, in both theory and design, intended to provide accessible justice and speedy correction of error.[79]

Australia has placed great reliance on its once internationally lauded system of 'new' administrative law based around merits review by an external

[75] *Freedom of Information Act* 1982 (Cth) s 34; Department of Prime Minister and Cabinet, *Government Response: Royal Commission into the Robodebt Scheme* (Report, November 2023) <https://www.pmc.gov.au/resources/government-response-royal-commission-robodebt-scheme>, archived at <https://perma.cc/QAP8-2G3G>. For a critique, see Yee-Fui Ng and Maria O'Sullivan, 'Robodebt, Transparency and Freedom of Information: Should the Cabinet Confidentiality Exemption Be Retained?' (2023) 30(3) *Australian Journal of Administrative Law* 192.

[76] *Privacy and Other Legislation Amendment Act 1988* (Cth) sch 1 item 88 ('*Privacy Amendment Act*'), inserting *Privacy Act 1988* (Cth) sch 1 cls 1.7–1.9 ('*Privacy Act*'). ADM use is to be disclosed in the privacy policy of an organisation where a computer program uses the personal information of individuals to make decisions, or to do something 'substantially and directly related to making' decisions, that could 'reasonably be expected to significantly affect the rights or interests of an individual': *Privacy Amendment Act* (n 76) sch 1 item 88 cl 1.7. The provisions commence on 10 December 2026: *Privacy Amendment Act* (n 76) s 2.

[77] Digital Transformation Agency, *Policy for the Responsible Use of AI in Government: Version 2.0* (Policy, 2025) 1 <https://www.digital.gov.au/ai/ai-in-government-policy>, archived at <https://perma.cc/WT7D-27BT> ('*Responsible AI Policy*'); Digital Transformation Agency, *Standard for AI Transparency Statements: Version 2.0* (Policy, 2025) 4 <https://www.digital.gov.au/ai/ai-in-government-policy/standard-ai-transparency-statements>, archived at <https://perma.cc/NP32-D2FT>.

[78] Kimberlee Weatherall, José-Miguel Bello y Villarino and Alexandra Sinclair, *AI Transparency in Practice: An Evaluation of Commonwealth Entities' Compliance with Their Obligations Regarding AI Transparency Statements* (Report, ARC Centre of Excellence for Automated Decision-Making and Society, 10 February 2026) 9, 14.

[79] *Administrative Review Tribunal Act 2024* (Cth) s 9 ('*ART Act*').

tribunal[80] — free of either cost or formality in the case of social security.[81] Serious limitations of that model were, however, exposed by Robodebt. First, by the Tribunal's then inability to avoid being 'gamed' by government: for example, by failing to make a reviewable 'decision' and by failing to appeal adverse first tier rulings of the then AAT in order to avoid public knowledge of their existence.[82] Secondly, it exposed both the sheer pace and scale of harms associated with the technological transformation of government decision-making from the adoption of ADM,[83] and the rarity of merits review of such errors.

While the recent reforms have removed the prospect of gaming,[84] even realisation of the right to individual access to justice remains questionable under any future problematic ADM scheme. The successful Robodebt class action settlement[85] put a number on how many *primary* decisions were faulty (794,000 debts impacting 526,000 citizens),[86] enabling a world-first calculation of how many of those mistakes reached external merits review. In this instance, a disturbingly low figure of just one in every 2,000 (0.05%) ever reached an external review, access to which requires prior completion of an internal review, generally by an authorised review officer; this constitutes 1% of the social security caseload of first tier AAT review.[87] This exposes the potential flimsiness of

[80] See generally Brian Jinks, 'The "New Administrative Law": Some Assumptions and Questions' (1982) 41(3) *Australian Journal of Public Administration* 209. The main aims of this 'new' administrative law are correction and redress of individual mistakes, speedy rectification of systemic errors and the honouring of procedural fairness and the rule of law.

[81] Terry Carney, 'Welfare Appeals and the ARC Report: To SSAT or Not to SSAT' (1996) 4(1) *Australian Journal of Administrative Law* 21, 31–2; 'Centrelink Payments', *Administrative Review Tribunal* (Web Page) <https://www.art.gov.au/applying-review/centrelink>, archived at <https://perma.cc/5PN2-2R7C>.

[82] Carney, 'Robodebt Illegality' (n 20) 5–7; Townsend, 'Better Decisions' (n 23) 67. See below Part IV(A)(2).

[83] Margaret Allars, 'Automated Decision-Making and Review of Administrative Decisions' (2024) 58(3) *Georgia Law Review* 1145, 1173, 1182 n 182, citing *Robodebt Royal Commission* (n 12) 488.

[84] See below in Part IV(A)(3).

[85] *Prygodicz* (n 12).

[86] See above n 13 and accompanying text.

[87] *Social Security (Administration) Act 1999* (Cth) s 142(1) ('*SSA Act*'). These figures were calculated from the class action settlement and the compilation of all Robodebt cases compiled by Royal Commission staff: see Exhibit No 3-3493A to Royal Commission into the Robodebt Scheme (24 January 2023). See also Terry Carney, 'Australian Social Services Scandals: Towards Dystopia?' (2026) 48 *Sydney Law Review* 22190:1–20, 10 n 54 ('Social Services Scandals'); Terry Carney, Exhibit No 3-3489 to Royal Commission into the Robodebt Scheme (19 January 2023)

existing ex post facto individualised administrative justice as an answer to the sheer speed and scale of the mass impacts achieved by ADM compared to human administration.[88] Even otherwise commendable accessible justice features of bodies such as the ART, then, are no guarantee of access to justice.[89]

3 *How Would an Individual Review Translate into Systemic Action?*

A normative role in correcting ongoing systemic harm is a key goal of Australian merits review,[90] but Robodebt exposed a lack of machinery to ensure its realisation until rectified by clothing the ART with new powers. Those new powers entrust the senior leadership of the ART with the identification and referral of systemic issues to departmental heads, Ministers and/or the Administrative Review Council ('ARC'), for consideration and mandatory public reporting of the outcomes of referrals.[91]

While a considerable advance over the former AAT or the situation in the UK (where feedback loops for correcting systemic errors are said to be 'virtually non-existent'),[92] there are uncertainties around full and speedy enlivenment of these referral powers. The first hurdle is quick identification and referral of

8 ('Exhibit No 3-3489'). Robodebt cases were not internally tagged, so identification of AAT matters for the Royal Commission entailed manual inspection of all relevant debt files to locate suspicious 'runs' of supposedly stable fortnightly earnings at odds with the fluctuating earnings of the vast majority of clients (the only way of ascertaining the likelihood of Australian Taxation Office aggregate earnings over many months being applied as 'averaged' rather than 'true' fortnightly incomes). This took vast departmental resources and time merely to establish the true number of AAT Robodebt matters. It was not feasible for internal reviews, so the number of unsuccessful requests for internal review, or the number not proceeding to AAT review after obtaining an adverse review by an authorised review officer, remain unknown.

88 Sanchez-Graells (n 59) 9.

89 A key explanation for this lies in the previously discussed concept of 'administrative burden', or as Pierre-Marc Daigneault reconceptualises it, the 'onerous experiences' of accessing government programs (eg applying for social security or seeking merits review) or remaining engaged (eg satisfying compliance conditions or other activation requirements): Daigneault (n 42) 132. See McGann and Ball (n 47) 442.

90 The role of the AAT in strengthening government decision-making was reiterated in Explanatory Memorandum, Tribunals Amalgamation Bill 2014 (Cth) 3 [15]: 'A strong, impartial and effective Tribunal … would strengthen government decision-making'. From the outset of the 'new administrative law' in the 1970s it was recognised that this objective is not easy to realise: see Jinks (n 80) 217–18.

91 Justice Emilios Kyrou, 'Mechanisms in the ART Bill to Thwart Robodebt-Type Maladministration' (Speech, Australian Academy of Law Conference, 18 March 2024) ('ART Bill Mechanisms'). See below n 137 and accompanying text.

92 Gráinne McKeever, 'Vulnerability in the Social Security System' [2024] (April) *Public Law* 194, 202.

systemic issues by the ART. That hurdle was commendably passed with flying colours in the 'apportionment case', discussed in more detail in Part V, even though no inquiry eventuated.[93] The larger challenge is ensuring prompt consideration and ministerial or departmental action to rectify such systemic issues. The legislative framework sets a three-month time limit on departments or agencies the subject of a referral reporting to the ARC on the action they have taken in response to such ART presidential referrals.[94] However, public embarrassment apart, there is neither a guarantee that this timeframe will be met nor that substantive rectification will transpire.

4 *Would Exposure of Flawed ADM Necessarily Be Remedied?*

Robodebt was the brainchild of public service endeavours to meet savings targets: an endeavour given excessive rein by a degraded and unduly politicised public service culture that lent itself to precipitate and launch a stubborn defence of a program reflecting populist political messaging of the then government.[95]

Measures to restore the integrity of public service advice undoubtedly reduce the risk of defective programs reaching maturity.[96] But if and when a defective future ADM scheme does see the light of day, its rectification calls for extremely speedy and effective accountability machinery. This is because ADM generates such high volumes of decisions (20,000 new ADM debts a week as against 20,000 manually calculated debts each year in the case of Robodebt).[97] In the aftermath of the Royal Commission, the bureaucracy did act reasonably promptly to suspend the recovery of overpayments caught by a 2024 ruling of an authoritative panel of the then AAT, which found that the respondent's

[93] See below n 234 and accompanying text. The systemic issue posed by the ruling in *Re Secretary, Department of Social Services and FTXB* (2024) 184 ALD 29 ('*Re FTXB*') was the first referral for inquiry and report by the ARC: Justice Emilios Kyrou, 'Notice of Systemic Issue No 1 of 2024' (16 December 2024) 1–2 [1]–[3]. The ARC declined to undertake an inquiry at its meeting on 31 March 2025: Administrative Review Council, 'Meeting of Administrative Review Council 2024' (Statement of Meeting Outcomes, 31 March 2025) 1 [1] <https://www.ag.gov.au/legal-system/publications/arc-meeting-31-march-2025-statement-meeting-outcomes>, archived at <https://perma.cc/S29W-YTMJ>. However, the illegality of the targeted compliance framework ('TCF') system from April 2022 onwards was not detected by or referred from the ART: see below n 99.

[94] *ART Act* (n 79) s 294B.

[95] Andrew Podger and Donald F Kettl, 'How Much Damage Can a Politicized Public Service Do? Lessons from Australia' (2024) 84(1) *Public Administration Review* 160, 169–70.

[96] See Podger, *Report to Royal Commission* (n 30) 4; ibid 163.

[97] *Robodebt Royal Commission* (n 12) xxiv, xxvii.

earned income apparently was not correctly allocated to the relevant social security payment fortnights.[98] The bureaucracy also, if very belatedly, self-imposed a similar pause once the relevant department appreciated that automated cancellations and other sanctions for failure to comply with the requirements for receipt of JobSeeker payments did not align with the legislative provisions of the so-called targeted compliance framework ('TCF').[99]

In the UK on the other hand, the government refused to make any changes to its generic scheme of income support for people of working age (called Universal Credit) in light of similarly adverse apportionment findings that likewise resulted in dramatic fluctuations in entitlements for people with intermittent earnings.[100] A 'too big to fail' rationale was successfully advanced, and it was concluded that it was too costly to rebuild Universal Credit to avoid penalising

[98] *Re FTXB* (n 93) 77–9 [184]–[187] (Kyrou J, Senior Members Kennedy and Trotter); Iain Anderson, Commonwealth Ombudsman, *Lessons in Lawfulness* (Report, August 2023) 3, 10–11 <https://www.ombudsman.gov.au/__data/assets/pdf_file/0040/299947/Income-Apportionment-Lawfulness.pdf>, archived at <https://perma.cc/JN6R-DJTD>; Ray Griggs, 'Public Statement Regarding Income Apportionment' (Media Release, Department of Social Services (Cth), 23 October 2024) <https://www.dss.gov.au/news/public-statement-regarding-income-apportionment>, archived at <https://perma.cc/9G9D-EEAV>.

[99] 'Decision to Pause Cancellations of People's Social Security Participation Payments for Not Meeting a "Reconnection Requirement" within 4 Weeks', *Department of Employment and Workplace Relations* (Web Page, 3 October 2025) <https://www.dewr.gov.au/assuring-integrity-targeted-compliance-framework/decision-pause-cancellations-peoples-social-security-participation-payments-not-meeting-reconnection>, archived at <https://perma.cc/H2B6-LZZV>. Some paused aspects are due to reactivate from 26 October 2026: 'Assuring the Integrity of the Targeted Compliance Framework', *Department of Employment and Workplace Relations* (Web Page, 4 August 2026) <https://www.dewr.gov.au/assuring-integrity-targeted-compliance-framework>, archived at <https://perma.cc/5EPF-BDKN>. The deficiencies of administration of the TCF, its background and the steps taken to rectify harms inflicted are covered in some detail in Carney, 'Social Services Scandals' (n 87) 5–7, 13–15. The points-based TCF, showing as traffic light (green, orange and red) zones on a smartphone app as the principal mode of access, was introduced in 2018 based on recommendations of an independent review: at 6. See Employment Services Expert Advisory Panel, Department of Jobs and Small Business (Cth), *I Want To Work: Employment Services 2020 Report* (Report, 2018) <https://apo.org.au/node/210776>, archived at <https://perma.cc/7FB2-7HZK>. See also Simone Casey, 'Robo-Compliance in Australian Employment Services' (2025) 60(2) *Australian Journal of Social Issues* 428, 435–7.

[100] Rita Griffiths, 'Universal Credit and Automated Decision Making: A Case of the Digital Tail Wagging the Policy Dog?' (2024) 23(1) *Social Policy and Society* 1, 8, 10.

recipients when departmental records of when remuneration was 'received' did not square with the dates during which that income was 'earned'.[101]

In short, the outcome from stress testing the post-Robodebt reforms on these four aspects is worrying. So where do we turn for greater optimism about the future? The next Part probes both how adequately the reformed accountability and regulatory machinery protects against violations to individual rights, and how adequately it realises the potential to address systemic harms from the deployment of ADM which are revealed in the course of individual merits reviews. While fairly favourable in its assessment of individual rights protections under the new merits review body, the ART, the analysis will expose fault lines for the rectification of systemic issues. I show that — despite the significant strengthening of pathways for referral and redress of recurrent or systemic issues identified during such reviews — the very low incidence of review as a proportion of wrong or doubtful primary decisions renders those reforms a necessary, but quite insufficient, remedy for correction of systemic error.

IV Individual Safeguarding Pathways

The first point to consider about the regulation of ADM in government programs is that reform of pre-existing arrangements may not suffice. Rather than being 'just another technology' responsive to being held accountable by variants of traditional accountability measures that proved adequate for past technologies, ADM may represent a watershed moment where old legal or regulatory paradigms are found wanting, and new ones need to be developed.

A *Reforming Merits Review — Correcting Individual Error*

The ART provides external merits review of administrative decisions. It involves a full reconsideration of the decision (achieved by standing 'in the shoes of the original decision-maker')[102] to determine and apply the 'correct or

[101] Ibid. The system involved trade-offs between timing and stability of income support: see Jane Millar and Peter Whiteford, 'Timing It Right or Timing It Wrong: How Should Income-Tested Benefits Deal with Changes in Circumstances?' (2020) 28(1) *Journal of Poverty and Social Justice* 3, 7–9.

[102] *Frugtniet v Australian Securities and Investments Commission* (2019) 266 CLR 250, 271 [51] (Bell, Gageler, Gordon and Edelman JJ), citing *Shi v Migration Agents Registration Authority* (2008) 235 CLR 286, 299 [40] (Kirby J), 315 [100] (Hayne and Heydon JJ), 324–5 [134] (Kiefel J) ('*Shi*'). This is the phrase commonly used by tribunals and courts to capture the ART's

preferable' decision.[103] The review reassesses findings of fact and law, and the exercise of any discretionary choices, but defers to and presumptively applies departmental or government policy unless there are strong reasons in the individual case for departing from its application.[104]

Once legally mandated internal review by an authorised review officer has been exhausted, social security decisions can be reviewed by the first level of the ART.[105] In contrast to review of most other administrative decisions, a second tier of review is also available as of right from the first ART tier (there is no requirement to obtain permission or show that the case has special features).[106] At the discretion of the senior leadership of the ART (the President or their delegate), an application can now be referred by the ART itself to a newly created panel of senior members, called the Guidance and Appeals Panel ('GAP').[107] Referral is open where a matter raises 'an issue of significance to administrative decision-making' and referral is 'in the interests of justice'; or, again at the President's discretion, a matter may be referred on application of a party if it either raises such an issue or contains 'an error of fact or law materially affecting the decision'.[108] To promote consistency, ART members other than judges must 'have regard' to GAP rulings when deciding cases.[109]

In theory, such a system of merits review is a robust accountability protection to secure just and fair decisions for citizens. Yet, the Robodebt strike rate of only one in every 2,000 primary decisions making it through to merits

obligation to reconsider and remake the primary decision under review as if it were the original decision-maker: see *Miller v Minister for Immigration, Citizenship and Multicultural Affairs* (2024) 278 CLR 628, 633 [14] (Gageler CJ, Gordon, Edelman, Jagot and Beech-Jones JJ); *Nathanson v Minister for Home Affairs* (2022) 276 CLR 80, 94 [7] (Kiefel CJ, Keane and Gleeson JJ).

103 *Drake v Minister for Immigration and Ethnic Affairs* (1979) 24 ALR 577, 589 (Bowen CJ and Deane J). The history and philosophy of merits review conducted by the precursor to the ART (the AAT) was eloquently articulated by Kirby J in *Shi* (n 102) 295–8 [25]–[35].

104 The principles regarding deferral to or departure from a lawful policy were laid down by the Full Federal Court in *Drake* (n 110) 589–91 (Bowen CJ and Deane J), and applied by Brennan J on reconsideration of the matter by the AAT in *Re Drake and Minister for Immigration and Ethnic Affairs [No 2]* (1979) 2 ALD 634, 644–5. See also David Thomas, 'Contemporary Challenges in Merits Review: The AAT in a Changing Australia' (2019) 96 *AIAL Forum* 1, 5–7.

105 *SSA Act* (n 87) ss 129, 135, 142.

106 *ART Act* (n 79) s 131D(1).

107 Ibid ss 39(4), 40, 122.

108 Ibid ss 122(1), 124(1)–(2), 128(1)–(2).

109 Ibid s 110.

review gives rise to serious pause for thought about its fitness for purpose.[110] Three possible culprits arguably share the blame for the low rate of review of robodebts: undue administrative burdens, narrowness of what qualifies as a 'reviewable decision' and deliberate government 'gaming' of review.

1 *Administrative Burden*

The first likely culprit was the previously identified concept of administrative burden.[111]

During the life of Robodebt, there were two prior levels of *internal* review, both of which were operationally mandatory requirements — a *departmentally* imposed reconsideration by the original decision-maker ('ODM') and the *legally* mandated review by an authorised review officer.[112] Internal review preconditions are a known barrier to accessing external tribunal review and constitute an administrative burden or onerous experience, particularly when review times blow out well beyond the already generous 49 days the department sets for completion.[113] Similar barriers have been shown to depress access to merits review in the UK.[114] ODM reconsideration was never a legislative requirement and was quietly replaced some time in 2022 by provision for an 'explanation by a [subject-matter expert]'.[115] While a subject-matter expert is likely

[110] See above n 87 and accompanying text.

[111] See above Part II(C)(3).

[112] Terry Carney, 'Automation in Social Security: Implications for Merits Review?' (2020) 55(3) *Australian Journal of Social Issues* 260, 267.

[113] See, eg, 'Authorised Review Officers (ARO) Review Delay Issue', *Commonwealth Ombudsman* (Web Page) <https://www.ombudsman.gov.au/complaints/case-studies/authorised-review-officers-aro-review-delay-issue>, archived at <https://perma.cc/9AZZ-K8UV>; 'Explanations and Formal Reviews of a Centrelink Decision' (Web Page, 17 July 2026) <https://www.servicesaustralia.gov.au/explanations-and-formal-reviews-centrelink-decision>, archived at <https://perma.cc/R6FU-NSXE>. See above n 89.

[114] Gráinne McKeever discusses the mandatory review precondition and, in the case of unemployment, the lack of or delayed response when requests are keyed in the online 'journal': McKeever (n 92) 198–200. See generally Archie Drake et al, 'Legal Contestation of Artificial Intelligence-Related Decision-Making in the United Kingdom: Reflections for Policy' (2022) 36(2) *International Review of Law, Computers and Technology* 251, 261–3.

[115] Department of Social Services (Cth), *Guides to Social Policy Law: Social Security Guide*, version 1.321 (online at 14 October 2024) 6 Reviews, Debts and Payment Recovery, '6.1 Changes in Income and Assets: Reviews' [6.1.7] <https://webarchive.nla.gov.au/awa/20241025074405/https://guides.dss.gov.au/social-security-guide/6/1/7>. The current review pathway is outlined in Department of Social Services, *Guides to Social Policy Law: Social Security Guide*, version 1.340 (online at 1 July 2026) 6 Review of Decisions, Circumstances and

to be more senior and experienced than the ODM, the risk remains of the step continuing to be perceived to be yet another mandatory rather than optional requirement, joining the other already psychologically exhausting steps necessary before external merits review can be accessed, and thus contributing to the associated administrative burden.

The culture of merits review tribunals, including public perception of their inaccessibility or 'court-like' processes, can also generate administrative burden in the form of barriers 'at the door' of the tribunal, deterring otherwise eligible applicants. Gráinne McKeever reviewed evidence of a worrying drift into the judicialising of tribunal review of social security in Britain,[116] a concern shared by Lady Brenda Hale, the former head of its peak court, the Supreme Court.[117] It is too early to assess whether this disease will infect the ART, but it is critical that it operate on a genuine understanding of the fundamental differences between review by courts (judicial, adversarial, supervisory policing of legality and fair process) and *administrative* review by tribunals (inquisitorial; often very high volume, necessitating short hearings and rapid decision-making; and review on the merits as well as legality). The influential Callinan review unfortunately did not fully grasp those fundamentals.[118] The resultant higher proportion of legally qualified members among appointments to the ART[119] may heighten the risk of a similar drift away from the deeper, holistic understanding of government administration and of the place of administrative review (such as how to assess the 'preferable' as well as the 'correct' merits outcome) championed by Lady Hale and others.[120]

At the time of writing, no action has been taken to alleviate either of these sources of administrative burden.

Debts, '6.10 Review of Social Security Decisions and the Appeals System' <https://guides.dss.gov.au/social-security-guide/6/10>, archived at <https://perma.cc/KH6U-T3BS>.

116 See McKeever (n 92) 200–1.

117 Lady Hale, 'Courts and Tribunals' (Speech, Annual Upper Tribunal (UTAAC/UTIAC) Lecture, 14 October 2021) <https://www.judiciary.uk/annual-upper-tribunal-utaac-utiac-lecture-2021-2/>, archived at <https://perma.cc/P4CV-3H6M>.

118 IDF Callinan, *Review: Section 4 of the Tribunals Amalgamation Act 2015 (Cth)* (Report, 23 July 2019).

119 See *ART Act* (n 79) ss 207(3), 208(3).

120 See Hale (n 117) 3–6.

2 *The 'Decision' Gateway*

During the life of Robodebt, administrative burden was further compounded by a deliberate departmental practice of refusing to make (and delaying making if pressed to do so by potential debtors) the legally defined final decision.[121] Such a decision is necessary to trigger either advice about or the ability to access external review.[122] This practice consigned applicants to continuing to deal with the administration alone, or abandon their concern.

The Attorney-General's Department ('AGD') 2024 consultation paper on an ADM framework commendably recognised the need to revisit what qualifies as a decision.[123] It rightly identified the constraint on access to both internal and external review consequent on the *Australian Broadcasting Tribunal v Bond* restriction of the meaning of a decision as generally only covering the final end product of a decision-making process.[124] In light of Robodebt, it is hard to deny the need for some broadening of what constitutes a decision for social security purposes in particular.[125] However, determining how much wider the catchment should be is more difficult.

This difficulty is because only a small proportion of precursor steps to making welfare decisions (eg actions of algorithms and ADM elements, and digital processing of information) will generate the potentially harmful *consequences* for social security recipients that warrant access to individual merits review pathways or coverage by an ADM protective framework. For example, the use

[121] See Carney, 'Automation and Conditionality' (n 46) 42, citing Evidence to Royal Commission into the Robodebt Scheme, 8 March 2023, 4684–5 (Louise Macleod) <https://robodebt.royalcommission.gov.au/system/files/2023-03/transcript-hearing-day-44-8-march-2023.pdf>, archived at <https://perma.cc/M3ML-7TV7>.

[122] Notification of review rights is triggered once a decision is made, provided it meets the test of being a 'decision': *ART Act* (n 79) ss 4 (definition of 'decision'), 266(1), (4)–(5).

[123] Attorney-General's Department (Cth), *Use of Automated Decision-Making by Government* (Consultation Paper, November 2024) 8, 14–15 ('*AGD Consultation Paper*').

[124] (1990) 170 CLR 321, 337 (Mason CJ, Brennan J agreeing at 365, Deane J agreeing at 369) ('*Bond*'); *AGD Consultation Paper* (n 123) 14. To constitute a reviewable 'decision' under the High Court's ruling in *Bond*, the process must result in a 'final or operative and determinative' outcome: *Bond* (n 124) 337. See also Ng and O'Sullivan, 'Deliberation and Automation' (n 22) 26–9.

[125] While an issue across all portfolios, there is also a much deeper question about the wisdom of maintaining '"decisions" as sites of [public law] challenge' in light of the way digitisation and ADM displaces and disperses decision-making within extensive 'chains' of government and private sector actors, often disempowering both citizens and the departmental bureaucrats who are notionally the point of contact for communicating and attempting to respond to queries about ultimate outcomes that they know little about, and who may rely on other actors for rectification: Adelmant and Raso (n 40) 438–40, 445–6.

of the Job Seeker Classification Instrument to assign people to particular streams of employment services, differing in the intensity and nature of the programs of support (the least intensive being facilitation of self-help),[126] arguably should be caught by an expanded definition of 'decision'. This is because that gatekeeping decision impacts their capacity to comply with JobSeeker payment conditions.[127] On the other hand, profiling tools used to triage investigations into overpayments or possible fraud, I argue, ought not be so classified, at least so far as eligibility for *individual* (merits review) remedies are concerned.[128] This is because the algorithm does not impose an additional or collateral harm (or burden) beyond the identification of the possible fraud and is thus on all fours with the choices made by a human investigator in deciding which cases to prioritise. However, such an algorithm should be fully within the ambit of *regulatory* frameworks to ensure that unjustified systemic biases or other breaches of ideal ADM standards are eliminated.

Because the definition of a decision currently serves that variety of purposes, deciding where to draw the line on any expansion, and then finding language to embody that boundary line, is challenging. Action to broaden the definition of a decision remains a work in progress at the time of writing.[129]

3 *Government Gaming of Review Avenues*

Finally, there is the issue of gaming, including what is termed 'administrative non-acquiescence' with review rulings.[130] Of around 210 first tier AAT

[126] Angelika Papadopoulos, 'Integrity versus Ideology in Automated Assessment: The Jobseeker Snapshot' (2025) 60(2) *Australian Journal of Social Issues* 418, 418.

[127] See, eg, Department of Education, Skills and Employment (Cth), *New Employment Services Trial (NEST) Enhanced Services Assessments Guideline: Job Seeker Classification Instrument (JSCI) and Employment Services Assessment* (*ESAt*) (Guideline, 3 June 2020) 6–7 ('*NEST Assessments Guideline*').

[128] For Australia's social security fraud systems, see generally Scarlet Wilcock, *Policing Welfare Fraud: The Government of Welfare Fraud and Non-Compliance* (Routledge, 2024). Although automated predictions of suspicion are currently not 'decisions' under Australian or European Union legal jurisprudence and ADM/AI frameworks, they can still pose a threat to rights and privacy of citizens: see Francesca Palmiotto, 'When Is a Decision Automated? A Taxonomy for a Fundamental Rights Analysis' (2024) 25(2) *German Law Journal* 210, 224–9. See above n 124.

[129] See, eg, *AGD Consultation Paper* (n 123) 15; *Privacy Amendment Act* (n 76) sch 1 item 99, amending *Privacy Act 1988* (Cth) sch 1 cl 1. See above n 76 and accompanying text.

[130] See generally Michael J Froehlich, 'Administrative Nonacquiescence in Judicial Decisions' (1985) 53(1–2) *George Washington Law Review* 147; Carolyn A Kubitschek, 'Social Security Administration Nonacquiescence: The Need for Legislative Curbs on Agency Discretion' (1989)

decisions ruling Robodebt to be unlawful, unmathematical and unethical, not one was appealed by government to the next level, where both the second tier AAT hearing and ruling by the then General Division would have become public (unlike the in camera hearings and non-public reasons for decision of first tier rulings).[131]

That failure to appeal was palpably by design — to avoid an equivalent ruling being required to be applied across the board rather than the case at hand, thus invalidating the whole scheme, as found by the Royal Commission.[132] Together with the fudging of what constituted a decision, this was a form of litigational gaming at odds with obligations of the government as a model litigant.[133] Closing that serious loophole was understandably identified as a high priority reform,[134] along with ensuring accessible and transparent administrative review on the individual merits of each citizen's case, while hewing close to the traditional tenets of administrative justice.[135] This necessarily included enactment of the 2024 protections against gaming by enabling ART publication of anonymised reasons made in key rulings of its otherwise in camera first and second tier reviews, and adherence by government to model litigant ethics.[136] It also included enactment in 2024 of the previously discussed provisions for addressing systemic issues through referral to departmental heads, ministers and/or the ARC, once identified by the ART.[137] However, transparency now

50(2) *University of Pittsburgh Law Review* 399; Note, 'Government Nonacquiescence Case in Point: Social Security Litigation' (1986) 2(1) *Touro Law Review* 197.

131 See *Robodebt Royal Commission* (n 12) 558; Townsend, 'Repeat Player' (n 28) 182–6.

132 See *Robodebt Royal Commission* (n 12) 558–9, 563; Townsend, 'Repeat Player' (n 28) 186.

133 For discussion of model litigant obligations, and the lack of comprehensive protections against the unequal power of government in the context of administrative review, see Matthew Groves, 'Equality of Arms in Administrative Review' (2023) 46(3) *Melbourne University Law Review* 726, 771–6. There is also an argument for expanding model litigant principles as one of the guides for more innovative regulatory oversight of ADM/AI: Sheehy and Ng (n 26) 683–4, 687.

134 See *Robodebt Royal Commission* (n 12) 564.

135 For a recent but orthodox Australian review, see Amy Elton, 'Towards a Normative Standard of Administrative Justice: Themes and Principled Tensions' (2021) 95(12) *Australian Law Journal* 964.

136 *ART Act* (n 79) ss 70(1), 113(2), (4). See also Carney, Exhibit No 3-3489 (n 87) 11–14.

137 The responsibilities of the President of the ART include informing 'relevant Ministers, relevant Commonwealth entities and the Council of any systemic issues related to the making of reviewable decisions that have been identified in the caseload of the Tribunal': *ART Act* (n 79) s 193(i). Referrals are required to be detailed in the annual report of the ART: at s 242(2)(i). A similar obligation applies to the ARC. Referrals received and ministerial or departmental reports to the ARC of action taken within three months of receipt of any referral made by the

relies on ART *utilisation* of these powers; unlike the second tier General Division of the AAT, there is no longer the check of public hearings of social security matters at the second tier of the ART.[138]

Refreshing the normative role of review and revitalising managerial checks and balances is a separate issue.

B *Refreshing the Normative Role of Merits Review*

Equally or more confronting than correctives for the factors responsible for depressing individual merits review to levels verging on being a 'threatened species', is ensuring that these comparatively rare events do not extinguish prospects of merits review serving its other key function. That is, its normative (or preventive) function of prompting departments and agencies to take remedial action to avoid or reduce the incidence of decisions needing correction.[139]

1 *The Robodebt Reforms Addressing Systemic Issues*

Enactment of amendments to impose on agency heads and their staff a duty to assist the work of the Commonwealth Ombudsman and the Inspector-General of Taxation and Taxation Ombudsman, and expansion of the powers of both bodies to access agency records, certainly boosts confidence in the discharge of their roles.[140] However, the Royal Commission unearthed a culture of 'keeping in with government' that saw powerful existing powers of the Commonwealth Ombudsman unused, including withholding from its reports doubts about the legality of the scheme. So, expanded powers may not suffice.[141]

The reactivation in 2024 of the ARC and its added responsibilities as one of the potential recipients of referrals from the ART of systemic issues perhaps

ART must be included in its annual report: at ss 264(2), 294B. See generally Kyrou, 'ART Bill Mechanisms' (n 91).

138 *SSA Act* (n 87) ss 168(1)–(2). At the end of July 2025, 55 first tier decisions, and 49 second tier social security decisions had been published. Social security matters comprise just over 13% (4,358 of 32,828) of all ART cases finalised: Administrative Review Tribunal, 'ART Caseload Report for the Period 14 October 2024 to 30 June 2025' (Report, 2025) 1 <https://www.art.gov.au/help-and-resources/statistics>, archived at <https://perma.cc/644G-92LX>.

139 See above nn 90–4 and accompanying text above.

140 See *Oversight Legislation Amendment (Robodebt Royal Commission Response and Other Measures) Act 2025* (Cth) sch 1. The Act implements recommendations of the Royal Commission by amending the *Ombudsman Act 1976* (Cth) and *Inspector-General of Taxation Act 2003* (Cth).

141 *Robodebt Royal Commission* (n 12) 586–96.

inspires the most confidence.[142] The 2026 finalisation by the ARC of a statement of position regarding implementation by agencies of administrative review decisions of courts and tribunals[143] should, for example, help overcome the weak normative impact of ART rulings on primary decision-making. However, there is no legislative protection against the future reinstatement of executive directions precluding legal aid commissions from undertaking public interest litigation, nor any assurance of adequate funding to enable bringing appropriate test case litigation that *identifies* potential systemic issues.[144]

Renovations such as these to existing accountability machinery, then, only go part of the way to redressing harm from any future defective scheme.

2 *Managerial Reforms as an Answer to Systemic Issues*

Orthodox managerial analysis tends to emphasise correction of systemic error as the responsibility of executive government.[145] Measures for doing so include feedback loops and escalation of issues between frontline decision-making and senior policy levels of departments, and responding to systemic issues thrown up by internal (ie 'in-house') monitoring channels such as random sample audits.[146]

In-house processes are where the greatest volume and range of issues can be detected. So, the measures recommended by Andrew Podger such as restoration of higher level professional expertise specific to social security policy, and enhanced compliance with Australian public service standards of independence and fidelity to the APS Code of Conduct, may improve responsiveness of

[142] See *ART Act* (n 79) pt 9. The ARC's functions include inquiring into systemic issues: at s 249(1)(c). The ART President's functions include referral of systemic issues for the consideration of ministers, departments and agencies, and the ARC: at s 193(i). Referrals and progress is to be detailed in annual reports of the ART and ARC: at ss 242(2)(i), 264(2), 294B.

[143] Administrative Review Council, *Statement of Position: Terms of Reference* (Web Document, 8 August 2025) <https://www.ag.gov.au/legal-system/publications/arc-statement-position-terms-reference>, archived at <https://perma.cc/HXG4-XKGB>.

[144] In a US system according more weight to rule-making protections, Amy Widman identifies lack of legal aid as a key barrier to accessing administrative justice: Amy Widman, 'The False Premise of State Administrative Adjudication' (2024) 61(1) *Harvard Journal on Legislation* 137, 175–8.

[145] See, eg, Elton (n 135) 970.

[146] Daniel E Ho, David Marcus and Gerald K Ray, *Quality Assurance Systems in Agency Adjudication: Emerging Practices and Insights* (Report, 30 November 2021) 9–21 <www.acus.gov/sites/default/files/documents/ACUS_QA_Report_Final_Nov30.pdf>, archived at <https://perma.cc/HED6-4CWG>.

executive government to systemic issues.[147] While consistent with Jerry Mashaw's 1983 model of bureaucratic rationality,[148] reliance on managerial measures is, however, grounded in heroic optimism about the capacity and commitment of the bureaucracy to institutional self-correction. This places undue confidence in routinised 'proactive' correctives for the likely fairly negligible systemic corrective flowing from exercise of the reactive avenue of external ART review of individual cases.[149] Absent strong and regular accountability through, say, arms-length external audit office performance audits,[150] inappropriate executive government fiscal and policy goals of a degraded public service culture are therefore likely to continue to prevail over competing values of fairness, legality and sound administration.[151]

147 Podger, *Report to Royal Commission* (n 30). For earlier recommendations, see Andrew Podger, 'Public Sector Informant: Public Service Must Learn from Robodebt Fiasco', *The Canberra Times* (online, 6 July 2021) <https://www.canberratimes.com.au/story/7321036/public-service-must-learn-from-robodebt-fiasco/?cs=14264>, archived at <https://perma.cc/PA38-6QWR>. The APS Code of Conduct is contained in *Public Service Act 1999* (Cth) s 13.

148 Most recently, see Jerry L Mashaw, 'Models of Administrative Justice' in Marc Hertogh et al (eds), *The Oxford Handbook of Administrative Justice* (Oxford University Press, 2022) 375, 376–8. Mashaw's trinity of bureaucratic rationality, moral judgement and professional treatment has been supplemented by Michael Adler's additions of managerialism, consumerist logics and market logics: Raso, 'Implementing Digitalisation' (n 38) 524, citing Michael Adler, 'A Socio-Legal Approach to Administrative Justice' (2003) 25(4) *Law and Policy* 323, 332–3.

149 See Charles F Sabel and William H Simon, 'The Management Side of Due Process in the Service-Based Welfare State' in Nicholas R Parrillo (ed), *Administrative Law from the Inside Out: Essays on Themes in the Work of Jerry L Mashaw* (Cambridge University Press, 2017) 63, 68–9, 76.

150 There is a large literature on internalisation and operationalisation of such governance standards through impact assessments or concepts of public or social value, but serious reservations are held about their lack of remedial purchase: see, eg, Andrew D Selbst, 'An Institutional View of Algorithmic Impact Assessments' (2021) 35(1) *Harvard Journal of Law and Technology* 117, 122–5; Colin van Noordt and Gianluca Misuraca, 'Evaluating the Impact of Artificial Intelligence Technologies in Public Services: Towards an Assessment Framework' in *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance* (Association for Computing Machinery, September 2020) 8, 12–15.

151 This pessimism was borne out by the August 2025 Ombudsman's report into the failure by the Department of Employment and Workplace Relations to align its automated suspensions or cancellations of unemployment payments (the TCF) with legislative requirements: Commonwealth Ombudsman, *Automation in the Targeted Compliance Framework: When the Law Is Changed but the System Isn't* (Report, August 2025) 3–4 <https://www.dewr.gov.au/assuring-integrity-targeted-compliance-framework/ombudsman-draft-report-and-independent-assurance-review-statement-assurance>, archived at <https://perma.cc/KXT6-RF6V> ('*Automation in the TCF*'). The litany of reasons canvassed reprise many of those excoriated in the *Robodebt Royal Commission* report, including internal failure to coordinate policy and legal

So where does that leave internal departmental processes of responding to external ART review decisions? The Australian Public Service has historically been antagonistic to external review of its decisions and dismissive of such rulings being treated as revealing systemic lessons; that sentiment has been especially strong within successive departments responsible for social security.[152] AAT decisions were commonly characterised as lacking consistency and rulings were often confined to their individual facts.[153] This culture was deeply entrenched and thus highly resistant to change, highlighting the significance of the ART's new referral avenues that enable consideration of systemic normative implications of decisions.

Those reforms commendably went much further than either the fast track internal tribunal referral pathways to more authoritative merits review rulings thoughtfully canvassed by Maria O'Sullivan or the narrower ideas floated in 2023 in the AAT issues paper.[154] The normative effectiveness or otherwise of these avenues, however, remains to be seen.[155]

sections of the department (with policy dominant over a weak legal branch) or to liaise adequately with Services Australia, inadequate internal and external consultation or documentation of policy development, failure to seek or record external advice about possible legal or other issues, and inadequate quality assurance mechanisms to detect and promptly rectify emerging concerns: *Robodebt Royal Commission* (n 12) 20–31; Commonwealth Ombudsman, *Fairness in the Targeted Compliance Framework: When Decisions Are Made beyond Your Control* (Report, December 2025) 32–47 <https://www.dewr.gov.au/assuring-integrity-targeted-compliance-framework/ombudsman-draft-report-and-independent-assurance-review-statement-assurance>, archived at <https://perma.cc/643Z-6N29>.

152 See Hanks, 'Administrative Law and Welfare Rights' (n 20) 5.

153 Commonwealth Ombudsman, *Learning from Merits Review: Best Practice Principles for Agency Engagement with Merits Review* (Report, December 2024) 27–8 [3.6].

154 O'Sullivan recommended structural procedural review, aggregation of claims and new merits review tribunal case management approaches to systemic concerns: Maria O'Sullivan, 'Automated Decision-Making and Human Rights: The Right to an Effective Remedy' in Janina Boughey and Katie Miller (eds), *The Automated State: Implications, Challenges and Opportunities for Public Law* (Federation Press, 2021) 70, 82–8 ('ADM and Human Rights'). The AGD canvassed the ARC and the resolution of significant legal issues in Attorney-General's Department (Cth), *Administrative Review Reform* (Issues Paper, April 2023) 21–3, 80–3 <https://consultations.ag.gov.au/legal-system/administrative-review-reform-issues-paper/>, archived at <https://perma.cc/S23G-RK3M>.

155 The discussion of the way decision and data chains in the UK's Universal Credit scheme serve to frustrate these normative mechanisms gives much pause for thought: Adelmant and Raso (n 40) 441–6.

V Addressing Systemic Risks of ADM in Government

This Part turns attention to proposals for the implementation of the Royal Commission's remaining unimplemented recommendation (enactment of a regulatory framework to govern the use of ADM in government), and how these mesh with the reforms already enacted to merits review and systemic policy referral avenues.

I open by discussing the regulatory implications of the way automation transforms the speed, scale and potential for harm within mass decision-making portfolios (evidenced by human decision-makers' 20,000 debts a year compared with 20,000 a week under Robodebt).[156] My conclusion is that the existing and currently proposed machinery is unable to react swiftly enough to counter the speed and scale of such potential harms. The rarity of access to individual merits review or equivalent judicial or regulatory pathways, I conclude, renders it no match for redressing harms of this order of magnitude. In addition, I show that the political and fiscal drivers within government make it unlikely that a future large-scale ADM savings measure would be foreshadowed and co-designed or piloted in advance, rendering impractical what are otherwise theoretically attractive protections.

A *The Royal Commission and Regulation of ADM in Government*

Adoption and enforcement of a framework governing use of ADM by government were two key recommendations of the Royal Commission.[157] Although quite common internationally,[158] this was one of the last areas of reform tackled by the Labor government during its first term, and then not revisited until July 2026.

Two initiatives were taken. First, to progress the Royal Commission recommendation for an enforceable framework, a consultation paper — *Use of Automated Decision-Making by Government* ('*AGD Consultation Paper*') — on the use of ADM across government was issued by the Attorney-General's Department in November 2024. Over the period to 15 January 2025, officials spoke with 46 stakeholder organisations and ultimately received 'over 40' formal

[156] See above n 97 and accompanying text.

[157] *Robodebt Royal Commission* (n 12) 488. Implementation of both proposals was endorsed in the report of the Senate Select Committee on Artificial Intelligence in Recommendation 12: *Senate Committee on AI* (n 2) 141 [5.130].

[158] See, eg, Smart (n 5) 18–24. Smart reviews South American government initiatives against international normative guidance pronouncements.

submissions and 20 responses to its short online questionnaire.[159] Secondly, in February 2025, Services Australia posted on its web page a 'transparency' statement on automation and AI use in the welfare portfolio,[160] followed in May 2025 by publication of its 21-page *Automation and Artificial Intelligence Strategy 2025–27* ('*Automation and AI Strategy*').[161] These two Services Australia initiatives enshrine commendable goals and principles, though both underestimate the challenge of identifying when ADM (or AI) is *actually* deployed by government, as shown by recent research in New South Wales.[162] However, it is the adequacy of an ADM framework, and its monitoring and enforcement, that is critical to any durable reform attuned to social and economic conditions.[163]

I argue that the *AGD Consultation Paper* provided an inadequate foundation for durable reform. While it sought feedback on options regarding the form and content of an ADM framework, the paper neglected the crucial question of monitoring and enforcement.[164] It also took a narrow conservative view of the scope of the framework, devoting undue attention to traditional *liberal*

[159] 'Automated Decision-Making Reform', *Attorney-General's Department* (Web Page, 22 August 2025) <https://consultations.ag.gov.au/integrity/adm/>, archived at <https://perma.cc/UG62-3TTR>.

[160] 'Automation and Artificial Intelligence Transparency Statement', *Services Australia* (Web Page, 10 October 2025) <https://www.servicesaustralia.gov.au/automation-and-artificial-intelligence-transparency-statement?context=22>, archived at <https://perma.cc/X4WR-YVSL> ('Services Australia Transparency Statement').

[161] Services Australia, *Automation and Artificial Intelligence Strategy 2025–27* (Web Document, 2025). <https://www.servicesaustralia.gov.au/sites/default/files/2025-05/automation-and-ai-strategy-2025-27.pdf>, archived at <https://perma.cc/AY4V-6E9A> ('Automation and AI Strategy').

[162] See José-Miguel Bello y Villarino et al, 'Are We Regulating the Right Digital Systems? Testing Emerging Artificial Intelligence Frameworks against Real-World Public Sector Systems' (2025) 48(4) *University of New South Wales Law Journal* 1165, 1194.

[163] The pace and direction of the progression of further reforms was significantly impacted by the elevation of productivity as a priority of the government's second term of office following the May 2025 election: see Productivity Commission, *Harnessing Data and Digital Technology* (Interim Report, August 2025) iv. The Productivity Commission was tasked with inquiring into '[h]arnessing data and digital technology' as one of five pillars of productivity reform, recommending pausing steps to implement previously proposed 'mandatory guardrails for high-risk AI': at iv, 16–21.

[164] José-Miguel Bello y Villarino et al, Submission to Attorney-General's Department, *Automated Decision-Making Reform* (15 January 2025) 5–7, 10–11 <https://consultations.ag.gov.au/integrity/adm/consultation/view_respondent?uuId=50936674>, archived at <https://perma.cc/QQW3-Q59H> ('AGD Submission').

remedies such as ensuring transparency on the use of ADM[165] and rights of redress for aggrieved individuals.[166] I criticise the neglect of reforms targeting *structural* issues such as co-design prior to the introduction of ADM, adequate accommodation of the diversity of circumstances of users, consideration of the inequitable access to technology such as smartphones (the 'digital divide'), dealing with systemic errors or differentiating the framework across different public administration settings. I worry that it also overlooked the limitations of law and orthodox reforms in addressing *power imbalances*. The particular vulnerabilities of citizens engaging in *high volume* human service sectors such as social security were also insufficiently recognised in my opinion, likely resulting in an overly anodyne and superficial set of protections, as now discussed.

In formulating its approach, the Royal Commission discussed contemporary international ADM frameworks such as the *OECD AI Principles*,[167] endorsing that 'landscape of regulation and principles' as the general basis on which reform should be founded.[168] However the Commission also recognised that human services, such as social security, have *special* requirements, writing that it is 'a context in which human discretion and judgment is a vital component', so human oversight and 'in some cases partial automation with the final decision made by a human arbiter' would be the ideal outcome.[169] The first of

[165] Ibid 17–20. The various issues and international approaches to ensuring transparency are reviewed in Rita Matulionyte, 'Regulating Transparency of AI: A Survey of Best Practices' (Working Paper, Macquarie University, 29 August 2023).

[166] Bello y Villarino et al, 'AGD Submission' (n 164) 20–1.

[167] *Recommendation of the Council on Artificial Intelligence*, OECD Doc OECD/LEGAL/0449 (3 May 2024, adopted 22 May 2019) 7–11 ('*OECD AI Principles*').

[168] *Robodebt Royal Commission* (n 12) 481. The *OECD AI Principles* were summarised in *Robodebt Royal Commission* (n 12) 480 as requiring that ADM/AI should be designed in a way that

> respects the rule of law, human rights, democratic values and diversity, and include appropriate safeguards; for example, enabling human intervention where necessary. Transparency and responsible disclosure should be features … people should be informed when they are the subject of automated decision making and told how they can challenge its outcomes. AI and automated decision-making systems should function in a robust, secure and safe way, with potential risks continually assessed and managed. The principles call for accountability for organisations and individuals developing, deploying or operating AI systems for their proper functioning.

[169] *Robodebt Royal Commission* (n 12) 487, quoting Australasian Society for Computers and the Law, Allens Hub for Technology, Law and Innovation and Gradient Institute, Submission No ANON-24KG-9BRC-8 to Royal Commission into the Robodebt Scheme (3 February 2023) 3. For elaboration of the importance for vulnerable citizens of retaining human rather than ADM administration in government services because of the ability to inject 'empathy': see generally

its two recommendations, Recommendation 17.1, embodies that exceptionalist thinking, listing three uncompromising bullet points relevant for human services:

> 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.[170]

The *AGD Consultation Paper* seemingly failed to appreciate that the uncompromising language of the three specific safeguards was intentional.

B *Transparency of ADM Use*

A key consultation question in the *AGD Consultation Paper* was: '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?'[171] In social security or human services domains, I contend that this is the wrong question.[172] Transparency regarding government use of ADM should not be qualified.

1 *The Case for Unqualified Transparency of ADM Algorithms*

The case for keeping the purity of Royal Commission Recommendation 17.1 in terms of *unqualified* transparency is demonstrated by considering the

Malou Beck and Sofia Ranchordás, 'Empathy as a Public Value: Overcoming Administrative Vulnerability and Rehumanizing (Digital) Government' in Jurgen Goossens, Esther Keymolen and Antonia Stanojević (eds), *Public Governance and Emerging Technologies: Values, Trust and Regulatory Compliance* (Springer, 2025) 291.

170 *Robodebt Royal Commission* (n 12) 488.

171 *AGD Consultation Paper* (n 123) 31. Options are discussed: at 15–20.

172 For discussion of four ways the use of ADM can be rendered opaque and thus beyond judicial review, see Sinclair (n 1) 9–15.

previously mentioned Job Seeker Classification Instrument ('JSCI').[173] Not only is the basis of the JSCI algorithm currently secret, but unlike counterparts in most other countries, there is neither data about its accuracy (its false positives and false negatives) nor any routine human component of decision-making or override of outcomes.[174]

This decades-old logistic regression-based algorithm, derived in-house from undisclosed data points,[175] has massive implications for the welfare and life chances of an unemployed person (and others of working age who are subject to welfare activation obligations). This is because the JSCI determines what stream of privatised employment services a person is allocated to,[176] ranging from the comparatively negligible resourcing of the 'self-service' stream, to the extensive funding of services in the stream for people with complex needs.[177] The algorithmically determined decision about stream allocation, while *intrinsic* to a person's ongoing qualification to their social security payment since compliance with stream obligations is mandatory, is currently not one that qualifies as a sufficiently mature or final decision for the purposes of accessing merits review.

Deployment of the JSCI for purposes for which it was not originally designed, such as its incorporation at one time as part of NDIS decision-

173 See *NEST Assessments Guide* (n 127).

174 See Papadopoulos (n 126) 421; Patrick Gallagher and Ray Griffin, '(In) Accuracy in Algorithmic Profiling of the Unemployed: An Exploratory Review of Reporting Standards' (2025) 24(3) *Social Policy and Society* 401, 407–8.

175 See Papadopoulos (n 126) 421; Gallagher and Griffin (n 174) 407.

176 The network of employment services providers, previously variously called the Job Network and Jobactive, is currently called Workforce Australia: 'Workforce Australia Employment Services', *Department of Employment and Workplace Relations* (Web Page, 16 September 2 024) <https://www.dewr.gov.au/workforce-australia-employment-services>, archived at <https://perma.cc/KF48-E3RH>.

177 *NEST Assessments Guideline* (n 127) 5–7. For a review of the accuracy or otherwise of the predictions, see Gallagher and Griffin (n 174) 406–11. See generally Sam Desiere, Kristine Langenbucher and Ludo Struyven, 'Statistical Profiling in Public Employment Services: An International Comparison' (Working Paper No 224, Employment, Labour and Social Affairs Committee, Organisation for Economic Co-Operation and Development, 20 February 2019) <https://www.oecd.org/en/publications/statistical-profiling-in-public-employment-services_b5e5f16e-en.html>, archived at <https://perma.cc/N4UF-CD47>; Sam Desiere and Ludo Struyven, 'Using Artificial Intelligence to Classify Jobseekers: The Accuracy–Equity Trade-Off' (2021) 50(2) *Journal of Social Policy* 367.

making,[178] further strengthens the case for unqualified transparency.[179] As I envisage it, operationalisation of the entitlement to transparency about the use of ADM should include *explicit* mention of its use on public websites,[180] along with a right of access to the code on request.

There has been some progress on parts of that agenda. The December 2025 departmental update on the TCF debacle committed to public servants again always being 'involved' in TCF sanctioning decisions,[181] and the May 2026 government announcement of major reform to employment services included replacing the JSCI.[182] This was followed in July 2026 by a commitment to developing 'a framework to better regulate the use of automated decision-making within federal agencies, led by the Attorney-General, recognising the importance of ensuring fair, accurate and transparent government decision-making'.[183] There is, however, too little detail for real optimism about how well this will be achieved.

178 See Georgia van Toorn and Terry Carney, 'Decoding the Algorithmic Operations of Australia's National Disability Insurance Scheme' (2025) 60(1) *Australian Journal of Social Issues* 21, 27.

179 Due to its biased outcomes, a similar case for full transparency is made regarding the predictive algorithm used by criminal courts in many US states as a prediction of recidivism for the purposes of sentencing: Christoph Engel, Lorenz Linhardt and Marcel Schubert, 'Code is Law: How COMPAS Affects the Way the Judiciary Handles the Risk of Recidivism' (2024) 33(2) *Artificial Intelligence and Law* 383, 384–6.

180 The 2025 'Services Australia's Transparency Statement' (n 160) falls well short of this, going no further than identifying the types of *functions* where ADM can be used.

181 'An Update on the Targeted Compliance Framework', *Department of Employment and Workplace Relations* (Web Page, 17 December 2025) <https://www.dewr.gov.au/assuring-integrity-targeted-compliance-framework/announcements/update-targeted-compliance-framework>, archived at <https://perma.cc/JU4W-3DSL>.

182 Amanda Rishworth, 'National Press Club Address' (Speech, National Press Club, 27 May 2026); Department of Employment and Workforce Relations, *Shaping the Future of Employment Services: Discussion Paper* (Paper, 27 May 2026) 23.

183 Andrew Charlton et al, 'AI Consumer Safety Priorities' (Press Release, 20 July 2026) <https://www.minister.industry.gov.au/charlton/media/ai-consumer-safety-priorities>, archived at <https://perma.cc/E7QS-ASUD>; Holly Tregenza, 'Federal Government to Tighten Rules around Use of AI in Departmental Decision-Making', *ABC News* (online, 20 July 2026) <https://www.abc.net.au/news/2026-07-19/government-tighten-rules-ai-artificial-intelligence/106933212>, archived at <https://perma.cc/8S65-LJE6>. This followed the announcement of a more centralised response to AI, coordinated by a new office within the Department of Prime Minister and Cabinet: Anthony Albanese, Tim Ayres and Andrew Charlton, *AI in Australia's Interests* (Media Release, 15 July 2026) <https://www.pm.gov.au/media/ai-australias-interests>, archived at <https://perma.cc/M3WP-JQT5>.

2 *Third-Party Proprietary Rights*

With regard to proprietary rights of third-party providers of ADM systems or software components of welfare administration, in my opinion the same unqualified transparency regarding all aspects bearing on the rights of clients of welfare should apply, and likewise when administration of such services are outsourced to non-government entities (as is the case for Australia's employment services).[184]

Unfettered transparency of such algorithms for citizens engaging with welfare services irrespective of setting should be seen as the necessary cost of government introducing ADM.[185] If business reality means that no commercial technology provider will sell its ADM service to government on that basis, then government should build its own transparent system in-house, even if that is more expensive. If a previously acquired (existing) system is compromised in terms of an ability to render it transparent, government should, within a reasonable time, negotiate transparency with the third-party provider or replace that system with one that is compliant. I envisage this being determined by an independent oversight body (existing or newly created) entrusted with overall responsibility for ensuring compliance with the government ADM framework. To follow any other rule requires vulnerable welfare clients to trade away their entitlement to full transparency — that surely is why the Royal Commission framed Recommendation 17.1 as an unqualified right.

C *Review Pathways and Decision-Level Safeguards*

The *AGD Consultation Paper* asked: 'What decision-level safeguards should there be for persons affected by decisions made using ADM (for example,

[184] For an accessible review of outsourcing of employment services, see generally Mark Considine, *The Careless State: Reforming Australia's Social Services* (Melbourne University Press, 2022) ch 2. Current authority places algorithms owned by private providers beyond the reach of current accountability measures: see, eg, *O'Brien v Secretary, Department of Communities and Justice* [2022] NSWCATAD 100, [86]–[92].

[185] Services Australia adopted a neutral position on this issue. Under the subheading '[b]uild versus buy', it listed as a benefit of in-house development 'the ability to better explain automation and AI, offering full transparency of the model', but efficiency was then listed among the benefits of purchasing technology from third parties: *Automation and AI Strategy* (n 161) 19. 'The criteria of decision-making [about building or buying] will be based on factors including whether it is fit for purpose, its accuracy, maintainability, ongoing monitoring, workforce requirements, tuning requirements and explainability.'

review rights)?'[186] It is surely axiomatic that citizens impacted by questionable ADM decisions should have access to review. But the choice and breadth of that pathway is more fraught.

1 *A New ART Pathway or a Purpose-Designed Pathway?*

One possible avenue for airing concerns about ADM use or design would be enabling it to be raised as an adjunct at ART hearings reviewing the correctness and 'preferability' of individual welfare decisions, or perhaps making it a separate ground for review.

The first could trigger exercise of the ART's new discretionary powers of referral of systemic issues to departmental heads and/or the ARC for further consideration.[187] But both ideas may strain the limits of administrative review,[188] posing not only challenges in defining the breadth of the expanded rights but also potential case overload due to vexatious applications.[189] And in light of the soberingly low proportion (one in 2,000) of robodebts making it through to merits review by the then AAT,[190] it would be foolish to place undue store in the corrective power of such avenues.

Individual pathways for raising concerns *directly* with the broad oversight body recommended by the Royal Commission have greater appeal. In Recommendation 17.2, it recommended:

[186] *AGD Consultation Paper* (n 123) 31. The paper discusses options for post-decision safeguards including merits review: at 29–31.

[187] See above n 137 and accompanying text.

[188] For recent reviews of public law protections for ADM/AI, see generally Paul Miller, 'A New "Machinery of Government"? The Automation of Administrative Decision-Making' in Monika Zalnieriute and Zofia Bednarz (eds), *Money, Power and AI: Automated Banks and Automated States* (Cambridge University Press, 2024) 116; Will Bateman, 'Algorithmic Decision-Making and Legality: Public Law Dimensions' (2020) 94(7) *Australian Law Journal* 520.

[189] But see José-Miguel Bello y Villarino, 'A Tale of Two Automated States: Why a One-Size-Fits-All Approach to Administrative Law Reform to Accommodate AI Will Fail' in Zofia Bednarz and Monika Zalnieriute (eds), *Money, Power and AI: Automated Banks and Automated States* (Cambridge University Press, 2024) 136, 141–2 ('Two Automated States'). José-Miguel Bello y Villarino argues that there is limited friction with traditional administrative law principles from reforms empowering

> [t]he affected party, alone or in conjunction with others affected by the same or similar decisions from that system, need to be able to — at least — (i) explore why their decision can be distinguished from similar cases deserving a different administrative response; (ii) be able to raise new distinguishing factors (attributes) not considered by the system; and (iii) challenge the whole decision system on the basis of the process of pre-certification of the system and its subsequent monitoring as the system learns.

[190] See above n 87 and accompanying text.

> 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.[191]

Bodies suggested as candidates for this role included an enhanced Office of the Australian Information Commissioner or the creation of an 'AI Safety Commissioner'.[192]

The only action at the time of writing is on transparency of use of automation: the February 2025 Services Australia transparency statement on automation and AI use in the welfare portfolio,[193] the May 2025 release of its 21-page *Automation and AI Strategy* and the July 2026 foreshadowing of a whole government framework.[194] But arguably the most effective machinery suggested so far is the proposed new dual function oversight body sketched by Benedict Sheehy and Yee-Fui Ng.[195] Their proposal has two main strands: (1) provision for individual case reviews on application or in high profile matters automatically; and (2) a 'continual monitoring function — implementing something like [the] "right to a well-calibrated machine decision" ensuring that the AI is properly serviced, updated and functioning as intended'.[196]

Regulatory oversight and enforcement was, however, overlooked entirely by the *AGD Consultation Paper*.[197] The AI Safety Institute, established in 2026 as part of the *National AI Plan* released on 2 December 2025, is designed to oversee and liaise across government on AI adoption, but its enforcement powers are unclear.[198] The reversal of the first-term Albanese government's plans for

191 *Robodebt Royal Commission* (n 12) 488.

192 Ibid 485.

193 'Services Australia Transparency Statement' (n 160). The statement, to be updated annually, commits Services Australia to nine principles for safe and responsible use, including offline trialling of automation and AI in controlled environments; only using such technologies outside that environment 'when we have identified and implemented controls to protect against any potential negative impacts'; continuing to 'monitor and evaluate, and immediately pausing any systems if they stop meeting assurance and governance requirements'; and having a 'human in the loop' to check system outputs 'where appropriate'.

194 Tregenza (n 183).

195 Sheehy and Ng (n 26) 687–9.

196 Ibid 688–9, quoting Huq (n 26) 619 (emphasis omitted).

197 See Bello y Villarino et al, 'AGD Submission' (n 164) 10–11.

198 Department of Industry, Science and Resources (Cth), *National AI Plan* (Policy, 2025) 31 <https://www.industry.gov.au/publications/national-ai-plan>, archived at

freestanding legislation and regulatory guardrails for AI raises doubt about the robust enforcement of the *National AI Plan,* or any future action from the AGD consultation, reputedly bogged down in interdepartmental committees and differing views about monitoring and enforcement,[199] despite the more promising July 2026 announcement.

While it seems clear that 'light touch' regulation will prevail for the private sector, there is perhaps a glimmer of hope that this might be elevated a notch or two to something approaching 'medium-light touch' oversight of government use of ADM. I say this because the December 2025 update of the *Policy for the Responsible Use of AI in Government* did inject

> requirements for agencies to develop a strategic approach to adopting AI, establish an approach to operationalise the responsible use of AI, ensure designated accountability for AI use cases, and undertake risk-based use case-level actions.[200]

The policy is also replete with 'mandatory' language about the obligations of agencies, including those 'risk-based' assessments of ADM use,[201] echoing ideas floated in the *AGD Consultation Paper*. But while ever hopeful, I remain dubious about the adequacy of levels of monitoring and enforcement of standards for ADM use in government.

Regrettably, no further action by government was evident on any of these fronts at the time of writing.

2 *Notification of Use of ADM*

A requirement for citizens to be notified that ADM has made or has contributed to a decision (as distinct from the transparency of, say, an algorithm) is provided for under arts 13(2)(f) and 15(1)(h) of the European Union's ('EU')

<https://perma.cc/HM2G-Q4YK>. As Justin Hurwitz wryly observes, however, '[e]veryone seems to want' standards for AI but they cannot agree on what it is that they are backing: Justin (Gus) Hurwitz, 'The AI Standards Alignment Problem' (2025) 8(1) *Journal of Law and Innovation* 1, 1.

199 See 'Introducing Mandatory Guardrails for AI in High-Risk Settings: Proposals Paper', *Department of Industry, Science and Resources* (Web Page) <https://consult.industry.gov.au/ai-mandatory-guardrails>, archived at <https://perma.cc/JZU8-HJTU>.

200 *Responsible AI Policy* (n 77) 4.

201 Ibid 5, 14–16.

General Data Protection Regulation.[202] Such an entitlement is fundamental to the ability to take advantage of post-decision protections (such as a right to reinsert a human into the decision-making or launch an application for review).[203]

Notification will, however, be meaningful only if its timing and content resonates in the lived lives of affected citizens (which occurs when they are able to understand and decide what, if anything, to do with that knowledge).[204] What matters is its empirical value (or otherwise) in facilitating *contestability* by the recipient of the information.[205] On timing, it follows that notification should occur at the earliest point where the information is sufficiently useful to recipients, given their diverse situations and circumstances, to equip them to consider taking action, such as seeking internal review of an ADM decision. Notification should be repeated at subsequent stages, such as when advising of a right to ART review.

The stated intention of ADM provisions inserted in the *National Disability Insurance Act 2013* (Cth) is promising in that regard,[206] but the *adequacy* of the substantive content of the notification hinges on a variety of factors, including whether it is sufficiently meaningful as to enable correction or airing of concerns about the ADM decision-making, and whether it effectively *communicates* to the scattergram-like diversity of linguistic ability, literacy levels and

202 *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)* [2016] OJ L 119/1 ('*General Data Protection Regulation*'). See also Palmiotto (n 128) 214.

203 See, eg, *General Data Protection Regulation* (n 202) art 22(3).

204 This is an as yet unresearched empirical and socio-legal question. It is hinted at in Joshua Krook et al, 'A Systematic Literature Review of Artificial Intelligence (AI) Transparency Laws in the European Union (EU) and United Kingdom (UK): A Socio-Legal Approach to AI Transparency Governance' (2025) 5(4) *AI and Ethics* 4069, 4071, 4076.

205 Mireia Yurrita, Agathe Balayn and Ujwal Gadiraju, 'Generating Process-Centric Explanations to Enable Contestability in Algorithmic Decision-Making: Challenges and Opportunities' (Working Paper, 2023) 2; Michael Ridley, 'Human-Centered Explainable Artificial Intelligence: An Annual Review of Information Science and Technology (ARIST) Paper' (2025) 76(1) *Journal of the Association for Information Science and Technology* 98, 109.

206 The stated intent of the new ss 59A–59D of the *National Disability Insurance Act 2013* (Cth) is to enable automation of 'specific administrative actions in a safe, consistent and transparent way', while preserving 'the important role of a human delegate' and 'the rights participants have, including for review', by requiring existing decision notifications to 'inform a person that it was taken by the operation of a computer program': Revised Explanatory Memorandum, National Disability Insurance Scheme Amendment (Securing the NDIS for Future Generations) Bill 2026 (Cth) 132.

cultural attributes of those affected.[207] Adequate understanding is not easily achieved. Even in the final year or so of Robodebt, this level of understanding was not present. While the informed public was increasingly aware of strongly held concerns about the scheme and its use of a crude form of ADM, the actual details of Robodebt's methodology were not sufficiently known.[208] Those finer details remained opaque even to Victoria Legal Aid Commission solicitors and the very experienced counsel Peter Hanks KC in the two Federal Court challenges.[209] This was due to the way its actual workings were kept secret, as the Royal Commission noted.[210] So, there are notifications and *notifications*: those that merely pay lip-service, and those that effectively arm people with the knowledge needed to take advantage of avenues for redress.

Of course, even if there had been sufficient detail about Robodebt's workings, there was, however, effectively *nothing* apart from a Federal Court challenge that these citizens could do about it at that time,[211] an issue to which I now turn.

3 *System-Level and Design/Post-Implementation Safeguards*

Influential scholars are increasingly questioning the capacity of existing or proposed legal and public service reforms to check what they identify as enormous *power* imbalances at the heart of Robodebt-type ADM scandals. Thanks to the thoughtful recommendations of the Royal Commission, there are now fresh channels, including arming the ART with powers to refer systemic issues and ensuring transparency through the publication of reasons for key decisions.[212] These measures are complemented by a suite of intrinsically very worthy

[207] See Jutta Treviranus, 'The Three Dimensions of Inclusive Design: A Design Framework for a Digitally Transformed and Complexly Connected Society' (PhD Thesis, University College Dublin, 8 May 2018) 19–21, 233 <https://openresearch.ocadu.ca/id/eprint/2745/>, archived at <https://perma.cc/6KTJ-8LRZ>; Shari Trewin et al, 'Considerations for AI Fairness for People with Disabilities' (2019) 5(3) *AI Matters* 40, 51–2.

[208] *Robodebt Royal Commission* (n 12) 167.

[209] Ibid 370; Hanks' Evidence to Royal Commission (n 72) 3832.

[210] See above n 72.

[211] This was because of the deliberate gaming of administrative review and judicial avenues for considering individual complaints: Carney, 'Robodebt Illegality' (n 20); Townsend, 'Better Decisions' (n 23) 67.

[212] *ART Act* (n 79) ss 113(2), 122. See also Kyrou, 'ART Bill Mechanisms' (n 91); Tom Brennan, 'The Administrative Review Tribunal as a Response to the Robodebt Scandal' (2024) 31(1) *Australian Journal of Administrative Law* 17, 29–31.

changes to bolster public service institutions and processes, including the roles of the Office of Legal Services Coordination, Commonwealth Ombudsman, and Office of the Australian Information Commissioner, in the development and oversight of new ADM proposals.[213] The question is whether these channels are sufficient to counter the overwhelming power of government.

4 *The Power of (and Power within) Government*

Both Monika Zalnieriute and Jacob Priergaard conclude that much more is required.

Zalnieriute targets the impotence of any purely liberal framework of transparency, individual rights, due process and procedural protections.[214] She writes that 'to date, AI scholarship and regulation have embraced [such measures] without questioning their capacity *to tackle power in the Automated State*.'[215] Priergaard turns to internal politico-bureaucratic dynamics of the public service.[216] He argues that legal and administrative accountability reforms will fail unless supplemented by measures addressing entrenched administrative practices at the heart of three interlinked *institutional* drivers of Robodebt.[217] Those three drivers were identified as: the budget process, departmental splits between policy setting and service delivery responsibility, and the social security compliance framework.[218]

New welfare policy proposals in Australia are mostly the result of Cabinet Expenditure Review Committee requirements for portfolios to find savings (at times including meeting an 'efficiency dividend' in the form of an annual percentage reduction in the department's inflation-adjusted or real budget).[219] As Priergaard explained, the added Australian attraction of efficiency gains from ADM initiatives in welfare was that, unable to change policy set in another department, the welfare services delivery portfolio was left with just 'two major levers from which to generate budget savings on the service side of the portfolio

[213] See especially *Robodebt Royal Commission* (n 12) chs 19, 21–2.

[214] Zalnieriute (n 37) 224.

[215] Ibid (emphasis added). Although obviously cognisant of the way state bureaucracy constitutes (or is complicit in amplifying non-state) sources of oppressive power over citizens, Zalnieriute devoted all her attention to the power exercised by non-state actors such as AI platforms and corporations: at 222.

[216] Jacob Priergaard, 'Not My Debt: The Institutional Origins of Robodebt' (2025) 84(1) *Australian Journal of Public Administration* 142, 144.

[217] Ibid.

[218] Ibid. Priergaard gives a detailed and illuminating review of the budget process: at 148–50.

[219] Ibid 149–50, 152.

— service efficiencies and compliance.'[220] Turning to the Royal Commission reforms, Priergaard lamented that:

> The recommendations do not substantively address the institutional pressures of the budget nor the structure of the social services portfolio that led to Robodebt. The new legal oversights to new policy proposals are welcome changes but do not address the power imbalance at the centre of budget making that delivered the atomised austerity of line department budgets.[221]

The message here is that *structural* problems call for structural rather than distant or overlaid remedies incapable of altering underlying structural or power dynamics.[222]

Structural remedies cannot simply be plucked off the shelf, without modification. For instance, when significant savings are estimated over the forward budget years, as in the case of Robodebt ($4 billion or more),[223] the political reality is that there will never be any advance publicity about the measure. So although it is commendable that Services Australia's 2025 transparency principles commit to co-design, consultation with stakeholders, and other principles espoused as a protection against poor ADM,[224] those processes are unlikely to be undertaken *prior to* a future large-scale budget measure being announced. Any protection from an overlay of co-design thus needs to be imposed on the way the proposal is *subsequently* developed — principally in the form of a commitment to initial rollout as a 'trial', open to subsequent modification in light of co-design feedback. Ensuring that the Department of Finance understands and accepts that this means anticipated savings may significantly decline will be no less a political challenge in the future than it was for Robodebt (including ministerial loss of face or promotion prospects when savings are significantly downgraded).

This, I suggest, is an important reality check against undue reliance on new public service policies and procedures for pre-deployment consultation.

[220] Ibid 152.

[221] Ibid 155.

[222] For a similar critique of overlays such as 'personalization' or 'relationality' without addressing the structural features of neoliberal (often outsourced) governance under 'new public management theory', see also John Clarke and Janet Newman, 'Debate: Beyond the New Public Management?' (2025) 45(1) *Public Money and Management* 6.

[223] *Robodebt Royal Commission* (n 12) 410.

[224] Refer to points four, seven and eight in 'Services Australia Transparency Statement' (n 160).

5 *The Power Imbalances Due to ADM Scale, Speed and Depth*

Noam Kolt concluded that neither EU nor recent US regulatory initiatives meet the accountability challenge posed by big scale, sophisticated machine learning AI programs.[225] He argued that they fail to address three 'black swan' risks.[226]

Systemic risk, the last of the black swans, is the relevant one for present purposes. This is namely: 'In addition to harming large numbers of people, AI systems can cause severe damage to social and political institutions.'[227] Kolt made a plea for what he termed 'algorithmic preparedness', captured in five new principles:

> The first principle ('anticipation') concerns the goals of regulating AI technologies: *AI governance should aim to, among other things, anticipate and mitigate large-scale societal harm from AI systems* — a goal that is neglected by current regulatory proposals. The subsequent principles ('diversification,' 'scalability,' and 'experimentation') concern the means for achieving that goal: *AI governance should adopt a portfolio approach comprised of diverse, uncorrelated, and highly scalable regulatory strategies, while continually exploring and evaluating new regulatory strategies*. The final principle ('recalibrating risk') concerns balancing the benefits and costs of AI regulation. It suggests that *cost-benefit analysis of AI governance interventions should place greater weight on worst-case outcomes*.[228]

Australia's tardy progress towards the regulation of AI in Labor's first term of office, before abandoning freestanding regulation in favour of existing protections in its second term,[229] failed Kolt's 'anticipatory' principle. The *AGD Consultation Paper* on design of a separate regulatory framework for government ADM likewise made no mention of Kolt's 'first principle' of 'anticipation' as a high order goal.[230] Nor did it mention Kolt's associated idea of taking a portfolio

[225] Noam Kolt, 'Algorithmic Black Swans' (2024) 101(4) *Washington University Law Review* 1177, 1240.

[226] Ibid 1182–6.

[227] Ibid 1185.

[228] Ibid 1185–6 (emphasis altered).

[229] See above n 163. For the tension between innovation and regulation and recent policy phases, see generally Kai-Ti Kao, 'From Robodebt to Responsible AI: Sociotechnical Imaginaries of AI in Australia' (2024) 10(3) *Communication Research and Practice* 387. See also Fraser (n 74).

[230] See Kolt (n 225) 1185; *AGD Consultation Paper* (n 123).

approach (somewhat reminiscent of John Braithwaite's regulatory pyramid).[231] Neither was there any discussion of committing to undertaking Kolt's 'worst case' cost–benefit calculus. The same can be said of the government's July 2026 foreshadowed development of a whole government framework.

By the same token there has been little appetite to date for the kind of case by case risk-related and proportionate assessments proposed in Europe.[232] But, as now explained, I argue that risk is a problematic criterion for ADM in welfare settings.

6 *Why Risk-Based Frameworks Misread the Welfare Domain*

Risk is, of course, a component of any evaluation of possible harms associated with the miscarriage of administrative decision-making. But there are serious flaws in schemes that calibrate the degree of intensity of regulatory protections in ADM to tiers or categories of risk. The regulatory or political convenience of such regulatory frameworks does not equate with best practice or a sound evidentiary basis. In the welfare domain, I suggest that risk-based distinctions have no appeal at all.[233]

A case in point is the so-called 'apportionment' scandal. At least 100,000 overpayments (and underpayments) were generated because Services Australia's computer systems incorrectly applied the law about how to apportion fortnightly earnings when the fortnight in which monies were earned overlapped two Centrelink payment fortnights (as is not uncommon).[234] Although

[231] See generally Ian Ayres and John Braithwaite, *Responsive Regulation: Transcending the Deregulation Debate* (Oxford University Press, 1992) 35–51; John Braithwaite, 'The Essence of Responsive Regulation' (2011) 44(3) *University of British Columbia Law Review* 475, 480–90. For a recent application, see Orly Lobel, 'The AI Regulatory Pyramid: A Taxonomy and Analysis of the Emerging Toolbox in the Global Race for the Regulation and Governance of Artificial Intelligence' (2025) 57(4) *Loyola of Los Angeles Law Review* 859, 867–901.

[232] See generally European Law Institute, *Model Rules on Impact Assessment of Algorithmic Decision-Making Systems Used by Public Administration* (Report, 2022).

[233] For a similar critique of EU risk-based regulation of AI (which thus fails to catch important ADM systems), see Maarten Bouwmeester, 'Strengthening Safeguards and Controls in the Digital Welfare State: Insights from the Dutch Rechtsstaat in the Aftermath of the Childcare Benefits Scandal' in Yee-Fui Ng and Matthew Groves (eds), *Automation in Governance: Theory, Practice and Problems* (Hart Publishing, 2025) 127, 131–3 ('Strengthening Safeguards').

[234] Ibid 1. The long-running administrative mistake in applying the law as to when earnings are 'first earned, derived or received' pursuant to the *Social Security Act 1991* (Cth) s 1067-H23 ('*Social Security Act*') for the purpose of calculating fortnight by fortnight working age social security payments resulted in differing findings about the correct legal basis for correction of the errors on the part of a ruling by an authoritative ART GAP-like panel of the then AAT:

ultimately resolved by retrospective legislation and an arbitrary compensation scheme,[235] the analysis necessarily resulted in large *ultimate* discrepancies, built up over long periods of seemingly small fortnight-by-fortnight calculation errors.[236] Would the small fortnightly possibility of such small dollar value errors characterise those decisions as low risk? Of course not, because the *ultimate* harm here would have been of exactly the same magnitude of mass ADM error as was the case in Robodebt.

Building rigid risk-based variables into regulatory frameworks also inevitably involves potentially elitist calculus about what citizens consider to be high or low risk. It poses issues such as what level of monetary loss *matters enough* in the lives of social security recipients and invites a form of moral calculus regarding what *kinds* of ADM mistakes or concerns citizens are sufficiently 'aggrieved about', so as to warrant a greater degree of responsiveness, awareness or other protection. Ever since WG Runciman's path-breaking sociological research in 1966,[237] it has, for example, been known that people living at the economic margins may feel *more deeply hurt* than more economically secure citizens by accusations of breach of community values (such as welfare fault) or of incurring *any level* of debt (a subjective harm that may be greater than the objective harm caused by its impact on their credit standing).[238] On both grounds, I argue that it is wise to avoid building risk-based distinctions into an ADM framework so far as it applies to welfare and human services.

Instead, as contended for by Sylvia Lu, I argue that the focus should shift from risk regulation to *harm regulation*.[239]

Re FTXB (n 93) 63–8 [127]–[150] (Kyrou J, Senior Members Kennedy and Trotter); and a subsequent majority ruling of the Full Court of the Federal Court: *Chaplin v Secretary, Department of Social Services* (2025) 311 FCR 44, 82–7 [184]–[206] (Thawley and Hespe JJ, Kennett J dissenting at 94–6 [240]–[253]). The High Court of Australia dismissed an appeal from the Full Court: *Chaplin v Secretary, Department of Social Services* [2026] HCA 22, [69] ('*Chaplin*'). The Court held that when a decision-maker is satisfied of the date income was 'received' but cannot be satisfied of the date it was 'earned' or 'derived', s 1067G-H23 of the *Social Security Act* instructs the decision-maker to consider the income in the fortnightly instalment period it was 'received': *Chaplin* (n 234) [67]. See also Carney, 'Social Services Scandals' (n 87) 7–8; *Lessons in Lawfulness* (n 98) 10–11.

235 See Carney, 'Social Services Scandals' (n 98) 15, discussing *Social Security and Other Legislation Amendment (Technical Changes No 2) Act 2025* (Cth).

236 *Lessons in Lawfulness* (n 98) 11.

237 WG Runciman, *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth-Century England* (Routledge, 1966).

238 See ibid 9. For findings on the harms incurred by low socio-economic clients of the UK Universal Credit scheme, see Watson et al (n 54) 9–12; Morrison (n 46) 321–3.

239 Lu (n 52) 333. The EU risk-based model has an incomplete coverage of harms: at 385–8.

7 *Harms-Based Regulation and Enforcement*

Lu's regulatory trilogy — combining inclusive harm-based 'impact assessments', individual remedies and full disclosure — warrants further consideration as a possible starting point for adaption to fit the human services sector.[240] Certainly, her all-purpose taxonomy is by no means exhaustive of all the arguably relevant welfare harms previously discussed above.[241] It is built from four 'primary harms' redolent of civil justice concerns, namely 'eroding privacy, undermining autonomy, diminishing equality, and impairing safety'.[242] Most pertinent to ADM in human services, perhaps, are two cross-cutting 'aggravating' factors of paucity of accountability and 'algorithmic opacity'.[243]

Though also less than fully comprehensive — overlooking the sheer scale and speed of ADM in welfare, for example — the harms taxonomy proposed by Colin Watson et al is perhaps the most apt and most inclusive analysis to date regarding the harms on which to base regulation of ADM in human services.[244] The Watson taxonomy includes four dimensions: 'expenditure of resources' (for startup, operation and upkeep of online engagement); 'loss of agency' (loss of control, independence and capability); 'accrual of risk' (to reputation, of obsolescence or of abuse); and 'service specific' harms.[245] These dimensions resonate better with the experience of ADM in human services. For example, the first is particularly apt to the more recent TCF fiasco, where the Commonwealth Ombudsman found inadequate IT coordination, underinvestment and lack of ongoing oversight were at the heart of it operating contrary to law.[246]

[240] Ibid 375–94.

[241] See above Part II(C)(4).

[242] Lu (n 52) 323.

[243] Ibid 368–75.

[244] See Watson et al (n 54) 1.

[245] Ibid 13–15.

[246] In the first of two reports, the Ombudsman was scathing of the Department of Employment and Workplace Relations for failing to ensure that its automated compliance system implemented an April 2022 amendment to s 42AF(2) of the *SSA Act* making payment cancellations discretionary rather than mandatory (as they previously were), for the 10-month delay in pausing such penalties once the issue was recognised, and for failing for over three years to introduce a 'digital protection framework' as required by sch 1 item 159A of the *Social Security Legislation Amendment (Streamlined Participation Requirements and Other Measures) Act 2022* (Cth): *Automation in the TCF* (n 151) 13–19, 31–2.

8 *Why Front End Preventive Measures Are Problematic for Government Human Services Mass Decision-Making*

Systemic measures, as distinct from individual grievance resolution or other individualised measures, potentially offer vastly more effective protections, but are problematic for government social security administration. The *AGD Consultation Paper* relevantly asked: 'What pre-implementation safeguards should apply where ADM is intended to be used?'[247]

Much ink has been spilt, but to little gain for present purposes, about the pros and cons of various theoretical protections at the front end of consideration of ADM in government decision-making.[248] This point in the lifecycle arguably *does* warrant the greatest investment or attention, because of the obvious preventative benefits flowing from properly designed ADM that avoids the rapid mass harms that ensued from schemes such as Robodebt. The problem I see with most of the ideas for front end protections in human services administration, however, is that they misread the *enormous power* of government both as the initiator of introducing ADM and in its subsequent ability to dominate, if not largely control, the public narrative in favour of the initiative being sponsored. In theory, full consideration of proposals at, say, public hearings conducted by an independent external body as a precondition to proceeding — akin to obtaining development approval from a town planning or environment protection agency — would probably be an adequate pre-deployment protection. However, this has never been a serious option for a host of political feasibility and other reasons, such as cost and delay.

Potential protections prior to the full deployment of ADM include consultation (co-design) and staged development (piloting and evaluation). The most effective and touted of these front end protections is co-design with end users.[249] Surprisingly, the *AGD Consultation Paper* completely overlooked co-design: neither the word nor the concept was ever mentioned.[250] Services Australia's strategy did little better, devoting just two not entirely convincing

[247] *AGD Consultation Paper* (n 123) 31.

[248] See, eg, ibid 20–1.

[249] See Treviranus (n 207) 76; Georgia van Toorn, 'Inclusion Interrupted: Lessons from the Making of a Digital Assistant by and for People with Disability' (2024) 41(1) *Government Information Quarterly* 101900:1–10, 5–8.

[250] The options canvassed were the pre-implementation consideration of 'fairness and efficiency', 'legal compliance' and 'errors and impacts on individuals', along with 'risk assessments': *AGD Consultation Paper* (n 123) 20–1.

sentences across 21 pages.[251] Was this a sign of the failure to understand what effective design principles entail? Or did it reflect a deeper issue of an inability of law and its orthodox regulatory portfolio of policy options to appreciate systemic inequalities of governmental power (ie the political economy of government bureaucracies and decision-making)? Regardless, as foreshadowed above, co-design is potentially both bane and boon.

I contend that co-design of ADM is mainly bane in government settings because it most often is hijacked by government as a top-down process for public ratification of a predetermined ADM initiative by way of token consultation with end users.[252] True co-design and its associated preventive/protective value arises only if *end users* are granted the power of initiation (isolating a decision-making 'need' and deciding to select ADM to resolve that need), given carriage of the form taken by ADM, and able to influence its subsequent rollout and modification.[253] So, to my mind, co-design offers little. More creative and innovative pre-deployment measures are still required.[254]

All this adds force to my argument that ADM use in welfare administration needs to be made accountable to a regulatory body with powers of

251 *Automation and AI Strategy* (n 161) 4, 7. The introduction includes a commitment to 'human-centric' design 'supported by our customer-centred design approach, ensuring customers and staff are at the heart of everything we do': at 4. '[S]trategic partners' are mentioned as facilitating 'early co-design of initiatives, understanding emerging concerns or issues within the community, and ensuring robust and evidenced based decision making': at 7.

252 Joe Tomlinson, *Justice in the Digital State: Assessing the Next Revolution in Administrative Justice* (Policy Press, 2019) 77–80; Victoria Adelmant and Joe Tomlinson, 'Who Builds Digital Government? Accountability in the Private Sector's "Agile" Reconstruction of the Administrative Justice System' [2023] (April) *Public Law* 196, 197–8.

253 See Tomlinson (n 252) 79, citing Lorne Sossin, 'Designing Administrative Justice' (2017) 34(1) *Windsor Yearbook of Access to Justice* 87, 87–9.

254 See, eg, O'Sullivan, 'ADM and Human Rights' (n 154) 82–8. Extrapolating traditional administrative and judicial review doctrine is feasible where ADM outputs can be compared to human systems, but not in a more generative system which 'designs or evaluates policy decisions': Bello y Villarino, 'Two Automated States' (n 189) 143–8. Additionally, human rights, data protection and algorithmic accountability are unable to address collective as against individual bias: Adamantia Rachovitsa and Niclas Johann, 'The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch *SyRI* Case' (2022) 22(2) *Human Rights Law Review*:1–15, 13. See also Terry Carney, 'The Automated Welfare State: Challenges for Socio-Economic Rights of the Marginalised' in Zofia Bednarz and Monika Zalnieriute (eds), *Money, Power and AI: Automated Banks and Automated States* (Cambridge University Press, 2024) 95, 114–15. Reform achievements and limitations in response to the Dutch child care benefits scandal are canvassed in Bouwmeester, 'Strengthening Safeguards' (n 233) 147–51.

enforcement, bringing a harms-based rather than risk-based lens to its work. As shown, neither part of this agenda has yet been seriously progressed.

VI Conclusion

I began this article by reviewing some of the research on the causes and ways of avoiding a repetition of ADM public policy disasters such as Robodebt. By way of a thought experiment I then stress tested recent legal measures and associated reforms adopted to futureproof against any repetition of such episodes, paying particular attention to the mass decision-making characteristic of automated administration of human services such as social security benefits. That review found that the reformed accountability and regulatory machinery offers greater protection against violations to individual rights, along with some promising, but as yet untested, potential for addressing systemic harms from the deployment of ADM and AI in human services. I argued that these reforms do too little to address the powerful structural sources and reasons for such risks and that more creative and innovative measures directed at structural sources of harm are urgently required.

In the welfare services domain (indeed, all those high volume areas of government decision-making impacting ordinary citizens), I have argued that the Royal Commission Recommendations 17.1 and 17.2 — for the creation of clear pathways for affected citizens to seek review of ADM decisions, plain language explanation of when automation is used and how it works and unrestricted access to algorithms for expert scrutiny — ought be implemented exactly in the form drafted, without any offsetting considerations diluting their impact.[255] Allied reforms enacted to overcome government gaming of administrative review and provide effective avenues for referral of systemic issues for examination and rectification by departments or oversight bodies such as the ARC also found favour with me.

Turning to the political and fiscal drivers of overly hasty or poorly designed government ADM initiatives, I suggest that much work remains to be done in countering the real risk of another Robodebt-type scheme being unleashed on an unprepared public, leaving the magnitude and duration of harms inflicted hostage to how rapidly the currently renovated machinery is able to react. Because complaints-based or citizen-initiated pathways for accessing individual merits review or equivalent judicial or regulatory pathways are so *rarely*

[255] See *Robodebt Royal Commission* (n 12) 488. Outside that welfare/human impacts policy domain, in areas such as commerce and customs, there may be a case for more nuanced framing.

invoked, I argued that relying solely on correction through such avenues is seriously inadequate. And because automation in human services like social security moves at warp speed and on a massive scale — reaching 20,000 fresh debts a week in the case of Robodebt — I argued that to do so essentially pays lip service to access to justice when the reality is that ADM justice is neither accessible nor sufficiently swift.

I concluded this review by reprising themes raised by commentators questioning the limits of orthodox legal remedies, the unwisdom of law operating in isolation from political and managerial controls,[256] the inadequacy of narrowcasting visions that overlook infrastructural linkages,[257] and the insights offered by wider conceptions of algorithmic justice.[258] Based on a more speculative review of the structural drivers of ADM adoption in government and possible constituents of more innovative and capacious reform, I outlined possible measures to prevent (or at least minimise the risk) of the release of poorly designed programs at the outset, including by establishing regulatory bodies and proactive oversight processes that *automatically* detect and review possible harms without reliance on citizen complaints alone.

Sofia Ranchordas and Louisa Scarcella perceptively summed up my central message when they wrote that

> in our current digital society, there are three phenomena that simultaneously connect and disconnect citizens from government and impede millions of individuals from exercising their rights on equal terms: bureaucracy, technology, and power asymmetries.[259]

My central contention is that effective ADM accountability in human services settings calls for concurrent action on all three fronts: by regulating the way ADM is deployed within the bureaucracy, by recognising the multiple sources of potential harm from deploying ADM in human services and by being realistic about the power asymmetries and fiscal drivers of the adoption of ADM in welfare.

[256] See above n 36 and accompanying text. In a similar vein, though with different tentative visions of possible solutions, see Adelmant and Raso (n 40) 438–45.

[257] See above n 37 and accompanying text.

[258] See above nn 49–51 and accompanying text.

[259] Sofia Ranchordas and Louisa Scarcella, 'Automated Government for Vulnerable Citizens: Intermediating Rights' (2021) 30(2) *William and Mary Bill of Rights Journal* 373, 375.