



THE UNIVERSITY OF  
**SYDNEY**

## **The University of Sydney Law School**

Legal Studies Research Paper Series

No. 20/16

March 2020

### ***Automation in Social Security: Implications for Merits Review?***

Terry Carney

This paper can be downloaded without charge from the  
Social Science Research Network Electronic Library  
at: <http://ssrn.com/abstract=3561286>

# 'AUTOMATION IN SOCIAL SECURITY: IMPLICATIONS FOR MERITS REVIEW?'

TERRY CARNEY\*

As the role of technology steadily grows in justice systems around the world, the UK Ministry of Justice [and other agencies] have taken the step of being global pioneers. They are now in the process of putting many court and tribunal processes – as well as court administration systems – on to a digital footing. Tribunals – which hear many more challenges to the decisions of public authorities than the courts do via judicial review – are a major focus of these changes (Tomlinson 2019: 37).

## I INTRODUCTION

Technological transformation of government administration is not new, but neither Australian social security nor merits review before the Administrative Appeals Tribunal ('AAT') lay claim to join the UK as a 'global pioneer'. On one view, information technology, and specifically artificial intelligence ('AI'), is just the most recent technological advance to wash up on Australian shores: the 21<sup>st</sup> century's equivalent of the 1961 IBM golf ball electric typewriter, or the Osborne laptop of 1982. However as explained below, AI is a very capacious term, a taxonomy of which isolates forms which differ significantly from each other, calling for a more nuanced policy evaluation of its impact on justice and other values of merits review of administrative action.

AAT external merits review involves stepping into the shoes of the original decision-maker to determine afresh what is the correct and preferable decision. Merits review of social security decisions is or will be impacted by AI in two main ways. First, in changing the way disputes present themselves for adjudication (changing their 'decisional character' and their 'evidentiary form'). Second, in altering the way tribunals 'hear' and 'decide' those disputes. In social security the first has long been

present, though recently has become more complicated;<sup>1</sup> while the second – as for other Australian tribunals – lies in the future, if perhaps a nearer future than many may appreciate. After all, in the UK the digital tribunals project is ‘starting in the Social Security and Child Support Tribunal (SSCS) and then moving on to the First-tier Tribunal (Immigration and Asylum) Chamber ...’ (Tomlinson 2019: 38).

As already mentioned, AI is a compendious and fluid term. As Sourdin (2018: 1116) explains, ‘AI can include machine learning, natural language processing, expert systems, vision, speech, planning and robotics’. For present purposes it embraces everything from mere ‘digitisation’ (ie computerisation of records) and use of simple expert systems which merely ‘automate’ decision-making, through to deployment of very sophisticated ‘machine-learning’ algorithms and big data-sets able to ‘best’ and thus ultimately replace human decision-makers. In Australia, social security records were fully digitised a decade ago, when paper claims, Centrelink letters or other documents were scanned or created as digital records. Paper records were relegated to proof of identity and meeting the requirements of bodies such as the courts (eg for fraud prosecutions). Or (for a time) to service merits review bodies.<sup>2</sup>

A helpful taxonomy of the functional impacts of AI divides them into three types: the ‘supportive’ (aids to humans), the ‘replacements’ (automation of previously human activity), and the ‘disruptive’ (different *forms* of justice) (Sourdin 2018: 1117). Use of software, such as its Multi-Cal program to make debt calculations once data points are entered, is an example of Centrelink’s early adoption of supportive AI. The controversial Online Compliance Initiative for debt recovery (‘OCI’, commonly known

---

\* Emeritus Professor of Law, The University of Sydney (Eastern Avenue, University of Sydney, NSW 2006, AUSTRALIA; fax: +61 2 9351 0200; email: [terry.carney@sydney.edu.au](mailto:terry.carney@sydney.edu.au)); Visiting Research Professor, University of Technology Sydney

<sup>1</sup> Further, Terry Carney, ‘Automating Australian Social Security: Boon, bane or just bungled?’ a paper presented at ANU Public Law Weekend “Technology, Public Law and Public Administration” Canberra, 1 November 2019.

<sup>2</sup> Thus the 2012-13 *Annual Report of the Department of Human Services* wrote that ‘In 2011-12 the department created 233 837 paper customer files. As a result of digitisation, in 2012-13 we significantly reduced the number of new paper files created to 1086’: p 248.

as ‘robo-debt’) introduced in mid-2016 and which relies on a computer algorithm to automatically generate debts from cross-matching Tax Office data with Centrelink records, exemplifies ‘replacement’ AI. Such *automated* decision-making, drawing on long-standing provisions deeming the outcomes to be a ‘decision’ for review purposes, now operates on a large scale in some Centrelink settings, with over half a million robo-debts raised (if not yet collected) since mid-2016 (Carney, 2019b). Potentially disruptive AI in the form of machine learning properly so-called, however, currently remains largely in the design or pilot stage (Carney 2019a).

This article reviews illustrative aspects of social security (Centrelink) original decision-making and merits review, highlighting the complexity and nuanced character of evaluating greater reliance on AI over human decision-making systems. It argues that technological change is inevitable, and is not necessarily either an unadulterated boon or bane, but calls for careful planning and a comparative assessment of the pros and cons of AI versus human systems. In social security, particular attention should be given to avoiding discrimination against citizens who are technology poor or otherwise vulnerable (for a nice description of the ‘digital divide’, see Toohey et al. 2019: 145-147, Yu 2020), and ensuring that the first of the two tiers of merits review (AAT1) is able to deliver justice while meeting its statutory obligation to be ‘fair, just, economical, informal and quick’<sup>3</sup> (Sourdin et al 2019).

Although vulnerabilities for social security recipients include but are not limited to issues of lack of technological capacity, or stamina in accessing on-line decision-making and review systems (Carney 2018b, Wing 2017), this article focuses selectively on some emblematic issues, including difficulties with Centrelink’s on-line reporting and uploading of documents, difficulties of deciphering reasons for decisions, and the confusing system designs that deter or effectively lock citizens out of accessing

---

<sup>3</sup> *Administrative Appeals Tribunal Act 1975* (Cth) s 2A(b). Other objectives are to be accessible, be proportionate to the matter, and promote public trust and confidence in its decisions: s 2A(a), (c), (d) respectively. Second tier merits review is to the General Division.

independent review avenues. Issues and challenges representative of the implications of supportive, replacement and disruptive AI will also be canvassed.

## II ONLINE CHANGES TO THE CHARACTER AND FORM OF MATTERS FOR REVIEW

Online administration is a deliberately compendious term for the purposes of this article, intended to capture the variety of ways AI can impact what earlier was termed the 'decisional character' or 'evidentiary form' of the matters presenting for merits review.

### *A. Computational 'jargon in' equals computational 'jargon out'?*

First tier merits review by the now Social Services and Child Support Division of the AAT ('AAT1'; previously the SSAT, pre-amalgamation) has long grappled with mainly print-offs of digital outputs as Centrelink's documentary evidence. Under neoliberal governance pressure of tight funding, formerly multi-member AAT panels now are rare, hearing times are severely compressed, and oral decisions are encouraged (Carney and Bigby 2018). Legislative limitations currently preclude use of ADR,<sup>4</sup> and online dispute handling has not been contemplated. While the Division still copes commendably with its workload, personal experience attests that the pressure imposed on tribunal members in interpreting the raw outputs of digitisation of records is not inconsiderable.

Apart from a short authorised review officer ('ARO') statement of reasons for the internal review decision, documentation tendered to AAT1 hearings already comprises reams of raw print outs of electronic file notes and other information. For over a decade the bulk of the information in Centrelink 'files' has been presented in digital form, such as ADEX or Multi-Cal rate and debt calculation records; as the raw unformatted input fields used to generate actual Centrelink correspondence (but not the letter itself); and

---

<sup>4</sup> This would be rectified under measure 30 of the Callinan Report: Callinan, Hon IDF, 'Report: Review: section 4 of the Tribunals Amalgamation Act 2015 (CTH).' Canberra: Attorney-General's Department Australia <https://www.ag.gov.au/Consultations/Documents/statutory-review-tribunals-act-2015/report-statutory-review-aat.pdf> (December 2018, published July 2019), para 1.32.

as ‘file notes’ keyed to screen, often replete with shorthand acronym codes for events and transactions. The AAT now receives all these papers electronically (including the ARO’s decision statement), printing off paper copies for despatch to the applicant and the tribunal member in advance of the hearing. Experience in sitting on such cases over nearly 40 years confirms that these papers make for a challenging read, even when Centrelink has put the print-outs into chronological order and grouped the material (eg by grouping all letters, and all debt calculations). The evidentiary form of this material, then, adds significantly to the decisional burden imposed on AAT members.

For example, establishing whether a person had received sufficiently full information in a letter of advice about a rate reduction being challenged outside the period where automatic arrears are payable, actually hinges on determining a very few simple facts. It involves ascertaining whether the underpayment was advised in writing (if not arrears are unlimited), the date of that advice letter,<sup>5</sup> and the actual wording of the rate advice (to determine if it is sufficiently fulsome at law).<sup>6</sup> Locating this information within the morass of Centrelink verbiage and codes however is particularly challenging; the pertinent information is ‘buried-away’ in unformatted data strings, since no formatted copy of the actual letter is stored or provided (though of course Centrelink could readily produce one by re-running the input information). File notes likewise frequently present not as discrete records for each date ‘segment’ appearing on the page, but with easily overlooked later annotations tucked away within these text blocks (requiring very close scrutiny to establish accurate event timelines). Lurking in the material may be broader textual observations, crucial to understanding earlier or later events, or which corroborate the version of events being advanced by an applicant.

---

<sup>5</sup> *Social Security (Administration) Act 1999* ss 109(1) [unlimited arrears if no adequate ‘notice’], 109(2) [arrears of up to 13 weeks if challenged within that period of receipt of notice, otherwise only from date of correction].

<sup>6</sup> The requirements to constitute *adequate notice* are set out in cases such as *Secretary, Department of Family & Community Services v Rogers* [2000] FCA 1447 Cooper J at paras [33]-[34]; *Austin v Secretary, Department of Family and Community Services* [1999] FCA 938, (1999) 92 FCR 138.

The already substantial degree of difficulty facing review bodies can inadvertently be compounded by otherwise attractive technological innovations such as tribunal members deciding to adopt a ‘justice on a tablet’ approach to handling the hearing papers. Some AAT1 members have elected to dispense with the AAT registry’s paper print-off of file documents at hearings. Instead they rely on a tablet or laptop to view the pre-tagged PDF file documents (using colour highlights or other tags marked onto the electronic-file). The justice and wisdom of this is problematic however. During the course of a hearing it is not uncommon for a comment to be made to a tribunal member that jogs their memory of some relevant matter. If manually ‘marked up’ on the paper file in preparation for the hearing, that material generally can be located manually while conversing with the applicant. This serves the very important justice principle of enabling that information to be introduced into the conversation for confirmation or otherwise. This simply is not able to be achieved by scrolling through what may be several hundred pages ‘on screen’. Nor does the short time of one hour set aside for AAT1 hearings readily allow a later opportunity to put this material to the applicant in person. Adjournment for telephone follow-up after conclusion of the hearing is the only realistic option, and then only if member workloads permit.

AI initiatives such as *fully* electronic hearing papers at AAT1, then, presently would come at a cost to the quality of justice.<sup>7</sup>

### *B. Contextual features matter?*

The evidentiary form of presentation of a merits review issue to AAT1 is indeed shaped by AI (as later explained) but regard must always be had to the other factors at play within the *overall* procedural and decisional context.

The earlier history of social security merits review offers a classic example of the way the overall operational design makes a difference. Prior to amalgamation the

---

<sup>7</sup> A similar hypothesis is advanced by Tomlinson, writing that ‘[s]uccess rates between paper and oral appeals differ significantly, and online appeals could have similar consequences’: Joe Tomlinson. *Justice in the Digital State: Assessing the next revolution in administrative justice*. (Bristol: Policy Press (Open Access), 2019), 50.

operational structure of the predecessor tribunal to AAT1 (the Social Security Appeals Tribunal 'SSAT') had the capacity to discover what then US Secretary of Defence Donald Rumsfeld popularised in 2002 as 'unknown unknowns'. At least during the decades before being readied for amalgamation, the SSAT could readily uncover such not uncommon situations as where Centrelink mistakenly recorded a client's required advice of their income in the wrong place, such as under family tax benefit ('FTB') payments instead of the working age payment to which the information actually related. Previously this was discoverable because the SSAT received *all* file papers (including FTB paperwork), not just those deemed relevant to the decision by Centrelink; and because it also had access to the Centrelink mainframe computer (helped further by members or support staff with knowledge of its multiple acronym codes and processes<sup>8</sup>). That was all lost when mainframe access was cut off and only the so-called section 37 'T document' papers forwarded, consistent with AAT processes in its other jurisdictions. Such an 'unknown' in this example now cannot be detected because the FTB papers would no longer be provided. In the result, the former SSAT's more extensive practical powers of inquisitorial inquiry were diminished in the interests of AAT uniformity and independence from the agency being reviewed. No longer can unknown unknowns be detected on review. Absolute purity of principle has its price.

Shifting from human decisions to those made or assisted by AI also changes the evidentiary *form* of the way issues present. For instance AI-generated primary decisions may lack the quality of 'face-explainability' (ie of being accompanied by *adequate* reasons). Explainability is insisted on under EU law, but it is not guaranteed elsewhere

---

<sup>8</sup> The Callinan Report regrettably demonstrated little understanding of the skill set required of competent AAT members, or of their legitimate support needs from staff at and under member direction; instead opting for what I may term a 'judicial' model where a monopoly of legally qualified members would decide everything on a more adversarial than inquisitorial basis: Callinan, n 4, paras 1.8 [measure 6], 8.20, 10.34.



(Olsen et al. 2019).<sup>9</sup> The most pertinent of contemporary examples of this problem within Centrelink is its recent development of a number of smart-phone app interfaces. These apps deliver information about payments, downloadable letters, advice about future appointments, and enable recipients to upload any required documents or information.<sup>10</sup> The undoubted convenience and other benefits of such technology for the many, nevertheless comes at a cost. It comes at the price of blurring the way ‘decisions’ are made (now made ‘virtually’ in response to data flows) and changes the very *geography* of governance of clients – shifting its location from being transacted at or in contact with a Centrelink office, into the ‘virtual’ space (Sleep and Tranter 2017: 506). They also raise other challenges to traditional protections of rights of citizens (Henman 2019).<sup>11</sup>

A case in point is the ParentsNext (‘PN’) program for sole parents at risk of long-term dependence. This was piloted in 2016 and rolled out nationwide in July 2018 at the same time as the implementation of a revamped Targeted Compliance Framework (‘TCF’) covering all working age social security recipients. PN targets clients who have been in receipt of Parenting Payment for more than six months without receiving income from employment and with a child under 6 years. The reformed TCF commendably now fosters compliance mainly by suspending and then restoring payments with back-pay on compliance, reserving actual rate reductions or non-payment periods for those few ‘wilfully’ doing the wrong thing. However both PN and

---

<sup>9</sup> In part this is due to major macro-level design differences between Australian and European models of accountability for administrative action. By relying heavily on administrative *courts* in many EU countries this elevates the importance of having reasons against which legality can be tested, while Australia’s *de novo* administrative review on the merits (stepping into the shoes of the primary decision-maker) diminishes the significance of reasons for primary decisions: further, Michael Asimow, ‘Five Models of Administrative Adjudication.’ (2015) 63(1) *American Journal of Comparative Law* 3-31.

<sup>10</sup> This, rather than say Centrelink failures to comply with protocols for what should be contained in reasons for the decision under review, remains most pertinent because AAT review is *de novo* and even if no or the wrong lines of reasoning are deployed, the decision generates its own ‘presumed’ grounding in the correct sub-set of reasoning possibilities.

<sup>11</sup> See for example, Paul Henman, ‘Of Algorithms, Apps and Advice: Digital social policy and service delivery.’ (2019) 12(1) *Journal of Asian Public Policy* 71 at 74-75 [‘street-level’ bureaucracy becomes ‘screen-level’ bureaucracy], 76-77 [some fundamental provisions of accountability were transformed under robo-debt].

TCF rely very heavily on use of digital (e.g. smart-phone) technology. Smart phones became the preferred mode of contact for reporting compliance (to be notified 'on the day') and for advising about and communicating a person's compliance status (a 'traffic light' system for alerting being at risk of or in actual breach status).

Associated with this, job-matching agencies assumed full *operational* responsibility, under their contract for service provision, for coding the acceptability<sup>12</sup> or otherwise of reasons for non-compliance. Instead of residing in the hands of skilled agency caseworkers (or a Centrelink delegate acting on provider recommendations) as previously, that function is now undertaken by front-desk clerical staff. This is because the task is now 'constructed' as a checklist exercise against a list of standard excuses. Default rules automatically allocating clients onto digital reporting without adequate assessment of their digital literacy or capacity to comply, further compounds difficulties for vulnerable ParentsNext clients and AAT members alike (in deciding whether subsequent sanctions should be upheld or otherwise). For clients there is the natural anxiety about looming penalties (Senate Community Affairs: 2019: paras [1.52]-[1.59]) and their loss of confidence in review processes due to inability formally to challenge notifications of the 'default points'<sup>13</sup> accumulating towards the trigger points for suspension or loss of payment (generally, Casey 2019). AAT members face the challenge of ascertaining whether document uploads or compliance notices and activities failed for technical reasons such as a drop-out, or were not ever actually initiated in the first place.

---

<sup>12</sup> The ultimate *legal* responsibility for determining acceptability of an excuse for a 'mutual obligation' or a 'work refusal' failure (as participation events are now labelled) remains that of a Centrelink delegate and is incapable to being delegated to an outside body: *Social Security (Administration) Act 1999* ss 42AI [reasonable excuse], 42AJ [excuses must pre-date the activity requirement unless there are reasonable grounds for not doing so].

<sup>13</sup> Unreviewable because the points in isolation do not result in an 'operative' decision but instead are simply steps along the path towards one, and thus have not matured to the stage of becoming a reviewable 'decision' as defined in *Australian Broadcasting Tribunal v Bond* (1990) 170 CLR 321.

### *C. The powers and procedures of the tribunal matter*

Of course the overall array of legislative powers within the review system design also has some part to play in accommodating the additional ‘adjudication complexities’ around AI, as now sketched.

Automated decision-making is already fully legitimised in social security, unlike the NDIS where automated ‘casework’ decisions would be much more problematic (for elaboration, Carney et al. 2019). It is legitimised in social security by ‘deeming’ computerised decisions to be reviewable ‘decisions’, as is the case for those involving human input.<sup>14</sup> AAT review powers when dealing with rate and debt decisions fortuitously however already are very ample ones. It was the challenge of performing very complex rate calculations (a product of a highly targeted system of payments) which led to empowering the AAT either itself to ‘assess’ a rate of payment when one is overturned, or to direct Centrelink to do so.<sup>15</sup> This proved necessary even though the Social Security Act 1991 deliberately rewrote the rules so that it became possible for very patient citizens to calculate their rate of payment by following the ‘steps’ of the relevant rate calculators, which were included in the law for the first time. Anyone familiar with the complexity of the reams of pages of Multi-Cal ‘long form’ rate calculations would understand why referral back to Centrelink is virtually always

---

<sup>14</sup> This has been the case since 2001, though in 1999 general authority was given to make or record a decision by computer, and from 1989 to that date, to ‘record’ it by computer: Will Bateman, ‘Automatic Public Law.’ a paper presented at “Public Law Weekend” – Centre for International and Public Law, ANU, Canberra, 3 November 2018. Equivalent provisions of the *Social Security Act 1991* covering automatic rate adjustment or cancellation decisions existed earlier, such as s 75A:

75A If:

- (a) a person is receiving an age pension on the basis of data in a computer; and
- (b) the pension is automatically terminated or the pension rate is automatically reduced by the operation of a provision of this Act; and
- (c) the automatic termination or reduction is given effect to by the operation of a computer program approved by the Secretary stopping payment or reducing the rate of payment of the pension;

there is taken to be a decision by the Secretary that the automatic termination or rate reduction provision applies to the person’s pension.

<sup>15</sup> Now *Social Security (Administration) Act 1999* s 177(a)(b) [previously s 149(2)(3) of that Act but originally *Social Security Act 1991* s 1253(2)(a)(b)].

chosen by the tribunal. Absent this power, cases involving a rate determination quite simply would have become impossibly hard for the AAT to resolve.

In short, the powers and procedures with which the adjudicative tribunal is furnished are another critical component of its capacity to accommodate complexity such as that introduced by AI.

#### *D. Machine learning is a different order challenge?*

The largest challenge however is surely the one looming on the horizon: that of machine learning.

This is not the place for a definitive analysis of the benefits and potential difficulty posed by machine learning algorithms, the introduction of which into Australian public administration probably is a long way off. Such algorithms use big data sets, dividing them into two parts and using one as a training set to discern underlying patterns that are then adaptively improved on in interaction with the rest of the data over time. So they are characterised by their ability to *replicate* complex human decision-making and then adaptively ‘learn’ how to do even better, at least when operating *within* the limits of their sphere of capability (for a nuanced discussion of the policy balance sheet: Zalnieriute et al 2019). This ‘besting’ of many human systems not only on cost and reliability but also on ‘quality’, builds pressure for their rapid adoption. But adoption certainly ought not be thought inevitable. Predictive analytics like Lex Machina (an off-shoot of LexisNexis) and other such machine learning applications which predict how judges (‘human black boxes’) decide cases, not only risk widening the economic barriers to access to justice (high costs of access to the systems: Sourdin et al 2019: 27) but also seriously overstate their capacity to accurately capture the ‘social and human’ essence of legal decision-making. This both undermines their utility and threatens the legitimacy of the legal system (Pasquale and Cashwell 2018).

The task, then, is to ensure that any machine learning is appropriately deployed and is well designed – as robo-debt demonstrably was not, even though it was only a simple *automation* system (indicating how much higher are the stakes for higher order

AI design). Good AI design covers many things, including compliance with principles of legality, transparency and procedural fairness; and rendering decisions capable of review. Transparency of decision-making is critical to these objectives, yet it is very challenging to achieve (Williams 2018: 4).

The EU (and thus presently the UK) applies the solution of requiring citizens to be notified when the decision lacks any human element and to have the right to request a human decision for a period of time after being advised of this (Veale and Brass 2019: 19). However this may prove to be an empty protection. As has astutely been suggested, the human being 'added to the loop' will often simply defer to and confirm the complex computer-made outcome (Olsen et al. 2019: 4). So other protections provided in the EU's General Data Protection Regulation ('GDPR'), such as third party audits and data protection impact assessments, may be more effective antidotes (Kaminski and Malgieri 2019: 3-4, 11-13). Susan Morse, for her part, instead places store in hastening slowly. She speculates that, when unable to offer full transparency or confidence, governments will adopt the 'precautionary' approach of *undershooting* the potential legal mandate, tightly restricting adoption of machine learning to the 'safe' zones (Morse 2019). This seems a rather heroic assumption in Australia, however, in light of the unchecked overreach by robo-debt in pursuit of maximum revenues (Carney 2019c).

Yet it is also important to remain grounded, and not attribute magical differences to machine learning (however mathematically complicated the weightings and algorithms) as compared to human decisions (Surden 2020 forthcoming: pp 5, 14). As Olsen and colleagues recently so pithily put it, 'we argue that the inner workings of an algorithm is not what is in need of explanation, but rather, the human interaction with the *output* of the algorithm' (Olsen et al. 2019: 10). In an Australian merits review context, this boils down to whether the person understands the basis for the decision and the information taken into account in making it. Or as Joe Goldenfein puts it, AI system explanations:

must be geared towards challenging decisions more than justifying them. It must be situated around exposing how an automated system may have used the wrong data; how the data used may not represent the totality of the data relevant to the question; how the system may have miscalculated or not understood the significance of that data; or how the rules, when applied to that data, might not produce the desired outcomes (Goldenfein 2019: 59).

It surely is hard to disagree.

### III FUTURE ONLINE JUSTICE CHALLENGES

The previous section identified a number of significant implications of AI regarding the current ability of merits review, and tribunal members, to do ‘justice’ to applicants and themselves. However the ‘virtual’ or at least the digital tribunal may not be far away in merits review, given a recent spate of interest in and experimentation with online dispute resolution (‘ODR’; for a recent though review, see Cashman and Ginnivan 2019).

In September 2018, the Victorian Civil and Administrative Tribunal (‘VCAT’) engaged a firm called MODRON<sup>16</sup> to run a small one month pilot of fully online (digital) dispute resolution of small claims,<sup>17</sup> with a view to a possible full roll-out in 2022 (Hendry 2018, Cashman and Ginnivan 2019: 42-43). Two months later, Churchill Fellow Katarina Palmgren’s study report recommended, among other things, that Victoria ‘[e]stablish [an] online court as a division of the Magistrates’ Court of Victoria with the jurisdiction to deal with low value civil claims up to \$10,000’ (Palmgren 2018: 51, rec 2). As Palmgren reports, British Columbia Canada already has a fully integrated online Civil Resolution Tribunal (further, Salter 2017, Cashman and Ginnivan 2019: 43-44).<sup>18</sup> Similar but bolder proposals have been made for the UK (Palmgren 2018: 9)

---

<sup>16</sup> See [https://www.modron.com/solutions\\_courts\\_tribunals](https://www.modron.com/solutions_courts_tribunals).

<sup>17</sup> VCAT’s summary of the pilot reported: ‘65 cases using the technology, with 71 parties participating in online hearings and 21 cases settled beforehand, an indirect result of online dispute resolution case management showing that online dispute resolution is a fast, cost-effective option for people with civil disputes at VCAT’: <https://www.vcat.vic.gov.au/news/sharing-vcats-online-dispute-resolution-experience>.

<sup>18</sup> Hangzau China is another to be mentioned: <https://www.nortonrosefulbright.com/en/knowledge/publications/71e0aa1e/online-dispute-resolution-and-electronic-hearings>.

including the English, Civil Money Claims Online (Cortes 2018: 105, 113-118), along with other initiatives globally (Cashman and Ginnivan 2019: 43-47).

Although ODR presently shrinks or attenuates rather than completely *replaces* the role of human adjudicators (such as by channelling parties into AI-assisted self-managed settlements or conciliations, or by supplementing human adjudication), the shift away from human oversight is both very real (Sourdin 2018) and is ‘trending’ under the press for fiscal savings. Thus in Mexico the Expertus system assists judges with advice ‘on the determination of whether the plaintiff is or is not eligible for granting him/her a pension’ (Sourdin 2018: 1119), and in Ontario, Canada, the Social Assistance Management System (SAMS) has since 2014 been involved in eligibility decision-making (Alston, 2019: 7, paras [21]-[22]).

It needs to be appreciated here that ODR is necessarily *transformative* of justice, such as in shifting from a more formal ‘hearing centric’ to a more informal ‘conversational’ process (Tomlinson 2019: 54). Writing about UK tribunal digitisation initiatives, Tomlinson therefore raises eight sets of issues to be considered in its design.<sup>19</sup> One of these is what digitisation might mean for *other* parts of the administrative system (such as mandatory internal review prior to accessing external review), and the overall normative impact in improving primary decision-making. Tomlinson writes:

The introduction of digital tribunals prompts multiple questions .... For instance, in the context of social security, there is a possibility that – next to an online tribunal procedure – MR [mandatory internal review] looks obsolete. How will the two systems – one paper-based and the other online – work together? ... There is plenty of room for creative improvements here too.... The prospect of digitalisation presents the opportunity to build in better and quicker feedback loops that consume less time, effort and money (Tomlinson 2019: 59).

In social security this ‘wider administrative justice system’ is often quite complex. The decision chain for review in Australia, for instance, already comprises

---

<sup>19</sup> The eight are: gatekeeping between on-line and traditional systems; respect for traditional justice values; design of communication platforms; ensuring fairness; avoiding digital exclusion; accommodating the wider justice landscape; data collection and management; and efficiency: Tomlinson, (n. 7),54-61.

many links or steps. A query made after being informed of the original decision (human or AI) not uncommonly leads first to a 'reconsideration' by the original decision-maker ('ODM'). If unsatisfied, Tomlinson's mandatory internal review ('MR') can be requested (in Australia the 'ARO' review). If still unsatisfied with that outcome, only then can *external* AAT1 review be sought for the first time. OCI 'robo-debt' effectively converted the first of these steps into a 'virtual' reconsideration. This is because the system insisted on exclusive engagement with the online compliance system, with applications for formal ARO or AAT review improperly dissuaded or unlawfully refused until this was completed (National Legal Aid Secretariate 2019: 28-30).

In practice this insistence on OCI contact could be considered to convert it into a 'disruptive' impact of AI. This is because obtaining and uploading all fortnightly payslips from up to seven years previously, proved so time consuming and daunting for citizens (there being only difficult to access telephone contact as an alternative, with little or no Centrelink officer assistance) that many alleged debtors simply 'gave up'. People were either unaware of the potential of or were unwilling to invest further in formal ODM, ARO and AAT1 review pathways. Instead they cut their losses by 'accepting' what in truth may still be false or over-inflated 'debts' (Carney 2019b). To reprise and slightly adapt Hazel Genn's evocative characterisation, effectively such a system becomes one that is 'about *just* a settlement' (by exhaustion or misinformation) rather than one 'about *just* settlement' (Genn 2012: 411). The merits review promised at law is thus rendered but a pale shadow of the justice it offers in theory.

One of the many other choices in designing ODR is between the so-called court-oriented (ie parallel with paper-based adjudication) and an 'integrated' or mainstreamed model. The Dutch Rechtwijzer for family divorce disputes is a parallel design (Smith 2018) while BC Canada adopted an integrated design for its Civil Resolution Tribunal (as favoured by Palmgren). One of the supposed bugbears with parallel systems is working out the gate-keeping between the electronic and the



traditional pathways, with concerns variously expressed about too few (or even too many) entering the online justice arm. Volume of business I argue is a very crude measure, because it conceals the value stance of the evaluator. For example the Dutch opt-in scheme has been said to have drawn ‘too few’ customers, an outcome put down to the available options being too blunt and the time for deciding whether to opt-in too pressured.

But it surely is more concerning if too *many* vulnerable citizens opt into on-line justice systems lacking the face-to-face hearings able to detect signs of distress, lack of comprehension of proceedings, or other cues of inadequate participation. Mandating online justice options as standard parts of the process, as under an integrated model, certainly deals with caseload volume concerns, but it does not necessarily tackle the important *compositional* equity concerns about *compounding* disadvantage for the digitally less competent segment of the caseload, nor the many other justice system values in play (for a thorough review, Cashman and Ginnivan 2019: 47-61, Surden 2020 forthcoming). For instance it may fail to meet the justice precept of ‘even-handedness’ (Sourdin et al 2019: 23-24).

Of course similar compositional equity issues *already* arise with delivery of oral rather than written decisions, and other more traditional procedural variations to hearings. Pressure to adopt such measures likewise often is driven by neoliberal governance policies which starve the machinery of justice of adequate funds. For better or worse, however, such measures at least remain firmly in the hands of tribunal members rather than being made by potentially vulnerable applicants themselves (as in parallel online justice) or being designed into the justice system (as in an integrated online justice model). Members of course are not necessarily good judges of such matters either, being prone to over-estimate their ability to dispense the same quality of analysis and reasoning when electing to deliver *ex tempore* decisions compared to the discipline of writing up (and the reflection that this entails), or to under-estimate the

temptation to allow self-interest to pick the less onerous option (for a summary of the research, Sourdin 2018: 1128-29). So it is not a clear-cut choice between human and AI.

Once again the policy balances prove to be more complicated and nuanced than is obvious on a more superficial consideration, whether considering the rather ‘luddite’ present or the more technologically ‘utopian’ future. Addressing these challenges of ODR futures will test us all, but they cannot be avoided. For otherwise it is surely true that ‘[w]ithout international standards, monitoring and global, cross-jurisdictional regulation of ODR, [risks] the software designer becoming a gatekeeper for access to justice’ (Wing 2017: 19). Yet in taking responsibility for ODR design rather than letting software designers set the values and principles, it is important not to over egg either present arrangements or their futurist technological alternative. Both are policy options. Both have merits and deficiencies. Nearly 70 years ago Alan Turing therefore proposed pitting purely human and purely AI decision-making (and by implication decision review) to a blind face off. A blind face off before a jury of human beings who, without knowing which decision path was which, would assess the quality and acceptability of outcomes of the human and the AI processes.

Without formalising the Turing test (as intriguingly is proposed by Olsen, 2019: 25-26), it might at least become one key benchmark when evaluating AI changes. For as Cashman and Ginnivan (2019: 61) demonstrate, for all their attraction of speed, cheapness, efficiency and user satisfaction, ‘in the design and implementation of such platforms, important objectives in terms of access to justice, open justice and procedural fairness need to be accommodated’.

#### IV CONCLUSION

The principal driver of adoption of AI, whether in government primary decision-making or in merits review processes, is of course economic efficiency and fiscal cost-cutting (eg, Palmgren 2018: 19-20). But as this brief review has demonstrated, many other fundamental values are implicated, such as those of substantive justice,

procedural fairness, distributive justice impacts (such as from the ‘digital divide’), compositional equity from differential treatment between on-line and human decision-making, and preservation of the transparency of open justice (Sourdin et al 2019; Tomlinson 2019).

All the economic or other supposed benefits of AI or other efficiency measures call for rigorous independent policy evaluation, but this is rarely undertaken. Thus Centrelink’s OCI ‘robo-debt’ initiative was optimistically projected to yield massive savings (up to \$3.7 billion over budget out-years). This involved capturing the outstanding 93% of presumed debts when ATO data-matching of earnings differed from Centrelink payment records.<sup>20</sup> Discrepancies which were previously left uninvestigated and unconfirmed because it was uneconomic to do so (further, Carney 2018a: 10) . However OCI’s poor design raised major rule of law and good governance issues. For its part, online dispute resolution such as the ‘digital tribunal’ being fast-tracked in the UK in social security, is likewise portrayed as both more efficient and as potentially widening access to justice (Palmgren 2018: 18-19, Cortes 2018), just as *ex tempore* oral decisions are touted as not only quicker and cheaper than written reasons but are thought to promote greater participation in and respect for the justice of the decision handed down. But there is virtually no research or evidence-base to ground reliable policy-making, whether about AI in primary decision-making or in the form of a digital tribunal (Tomlinson 2019: 8-9).

Leah Wing rightly notes, ‘[t]he ways in which we design ODR systems and manage data within them are central to whether they magnify the risk or the opportunities for access to justice’ (Wing 2017: 17). The same is true of AI, including sophisticated machine learning algorithms, within social security administration (Carney 2019c). In one sense it is trite to say that sound design is the difference between

---

<sup>20</sup> The so-called ‘match’ is of quite *incommensurate* data-types: ATO fortnightly averages of *supposed* income derived from total earnings of all employment, against the accuracy or otherwise of the fortnight by fortnight *actual* earnings as reported by the person as the basis for their legal rate entitlement. Further, Terry Carney, ‘The New Digital Future for Welfare: Debts without legal proofs or moral authority?’ (2018) *UNSW Law Journal Forum* 1-16.

acceptable and unacceptable uses. Robo-debt for instance surely would not have been so problematic had it been designed in accordance with the principles laid down by the Administrative Review Council in 2004 (Administrative Review Council 2004),<sup>21</sup> or had that oversight body not been ‘unlawfully’ wound up, as Callinan concluded (2018/19: para [1.27]).

Social security is, however, rather special in one very important way. Citizens subject to primary decision-making by Centrelink (whether by human hand or by AI) and those aggrieved clients who turn to the AAT for merits review, are disproportionately comprised of vulnerable individuals, whether due to age, location, human capital resources, mental illness, education, or other markers (Carney 2018b). That is why it is argued here that very *particular* attention needs to be paid to avoiding discrimination either in primary decision-making or on review. Attention that of course equally should be extended to other settings where vulnerable client populations are served by review, including consumer claims, adult guardianship and mental health, to name but three.

There are also some wider longer-term issues to be mindful of. This includes that AI ultimately is a *technological* not a human system. To coin a phrase, it sings from a different song sheet to that of the more discretionary forms of justice administered by human beings. It is less nuanced and more didactic in form. So although social security law has already lost much of its discretionary flexibility, a ‘slower burn’ risk remains. This is the risk that over time ‘[c]omputerisation is apt to *change the nature of an administrative process*, translating public administration from a person-based service to a *dehumanised system* where expert systems replace officials and routine cases are handled without human input’ (Harlow and Rawlings 2019: 19-20 [my emphasis]). There is also

---

<sup>21</sup> If followed there would have been no doubt about the legality of all debts raised. It would have resulted in a two stage process, where data matches led to the perfectly reasonable invitation for alleged debtors to upload or otherwise provide fortnightly pay slips or bank records. But when unavailable or not provided for whatever reason, under stage two Centrelink would have continued to bear the legal responsibility of obtaining fortnightly data (using its powers to require employers or other institutions to do so) essential to meeting its onus of proof obligation.

a synergistic risk. This is the name given to the phenomenon where AI decisions and digital ODR tribunals come to constitute a *mutually reinforcing* adaptive spiral, one which slowly ousts any vestiges of more equitable adjudication associated with traditional primary decision-making pathways (further, Re and Solow-Niederman 2019: 5).

Some of these changes are already well advanced. What is canvassed in this article are some of the ways AI is further strengthening some of the already disturbing winds of cultural change blowing through the AAT (Lucy 2017). These include the way ADR and other AI initiatives may further deteriorate rather than improve the quality of merits review decision-making or distort access to justice (Donoghue 2017).

## References

- Administrative Review Council. 2004. "Automated Assistance in Administrative Decision Making." Canberra: Administrative Review Council <https://www.arc.gov.au/Documents/AAADMreportPDF.pdf>.
- Alston, Philip. 2019. *Report of the Special Rapporteur on Extreme Poverty and Human Rights*. New York: Special rapporteur on extreme poverty and human rights. 1-20. <https://www.ohchr.org/EN/newyork/Pages/GA74thSession.aspx>
- Callinan, Ian Hon. 2018/19. "Report: Review: section 4 of the Tribunals Amalgamation Act 2015 (CTH)." Canberra: Attorney-General's Department Australia <https://www.ag.gov.au/Consultations/Documents/statutory-review-tribunals-act-2015/report-statutory-review-aat.pdf>.
- Carney, Terry. 2018a. The New Digital Future for Welfare: Debts without legal proofs or moral authority? *UNSW Law Journal Forum* (March): 1-16.
- Carney, Terry. 2018b. "Vulnerability: False hope for vulnerable social security clients?" *University of New South Wales Law Journal* 41 (3):783-817.
- Carney, Terry. 2019a. "Automating Australian Social Security: Boon, bane or just bungled?" a paper presented at ANU Public Law Weekend "Technology, Public Law and Public Administration" Canberra, 1 November.
- Carney, Terry. 2019b. "Bringing Robo-debts Before the Law: Why it's time to right a legal wrong." *Law Society Journal of NSW* 58 (August):68-70.
- Carney, Terry. 2019c. "Robo-debt Illegality: The seven veils of failed guarantees of the rule of law?" *Alternative Law Journal* 44 (1):4-10.
- Carney, Terry, and Christine Bigby. 2018. "Social Security and Welfare Rights – What role for social work?" In *Social Work in the Shadow of the Law*, edited by Simon Rice, Andrew Day and Linda Briskman, 361-388. Sydney: Federation Press.
- Carney, Terry, Shih-Ning Then, Christine Bigby, Ilan Wiesel, and Jacinta Douglas. 2019. "National Disability Insurance Scheme Decision-making: Or when tailor-made caseplanning met Taylorism and the algorithms?" *Melbourne University Law Review* 42 (3): 780-812 in press
- Casey, Simone. 2019. Social Security Rights and the Targeted Compliance Framework. *Social Security Rights Review* February.
- Cashman, Peter, and Eliza Ginnivan. 2019. "Digital Justice: Online Resolution of Minor Civil Disputes and the Use of Digital Technology in Complex Litigation and Class Actions." *Macquarie Law Journal* 19:39-79 [available at SSRN: <https://ssrn.com/abstract=3415229>].
- Cortes, Pablo. 2018. "Using Technology and ADR Methods to Enhance Access to Justice." *International Journal of Online Dispute Resolution* 5 (1/2):103-121. doi: 10.5553/IJODR/235250022018005102011.

- Donoghue, Jane. 2017. "The Rise of Digital Justice: Courtroom Technology, Public Participation and Access to Justice." *Modern Law Review* 80 (6):995-1025.
- Genn, Hazel. 2012. "What Is Civil Justice For? Reform, ADR, and Access to Justice." *Yale Journal of Law & the Humanities* 24 (1):397-418.
- Goldenfein, Jake. 2019. "Algorithmic Transparency and Decision-Making Accountability: Thoughts for Buying Machine Learning Algorithms." In *Closer to the Machine: Technical, Social, and Legal aspects of AI*, edited by Office of the Victorian Information Commissioner, 41-61, Melbourne: Office of the Victorian Information Commissioner. [available at SSRN: <https://ssrn.com/abstract=3445873>]
- Harlow, Carol, and Richard Rawlings. 2019. "Proceduralism and Automation: Challenges to the Values of Administrative Law." In *The Foundations and Future of Public Law*, edited by Elizabeth Fisher, Jeff King and Alison Young, forthcoming. Oxford: OUP.
- Hendry, Justin. 2018. "Victoria Looks to Settle Legal Disputes Online" *ITNews*, 10 September. <https://www.itnews.com.au/news/victoria-looks-to-settle-legal-disputes-online-512159>.
- Henman, Paul. 2019. "Of Algorithms, Apps and Advice: Digital social policy and service delivery." *Journal of Asian Public Policy* 12 (1):71-89.
- Kaminski, Margot, and Gianclaudio Malgieri. 2019. "Algorithmic Impact Assessments under the GDPR: Producing Multi-layered Explanations." U of Colorado Law Legal Studies Research Paper No. 19-28 [available at SSRN: <https://ssrn.com/abstract=3456224>].
- Lucy, Juliet. 2017. "Merits Review and the 21st Century Tribunal." *Australian Journal of Administrative Law* 24 (2):121-139.
- Morse, Susan C. 2019. "When Robots Make Legal Mistakes." *Oklahoma Law Review*:Forthcoming. [available at SSRN: <https://ssrn.com/abstract=3340103>].
- National Legal Aid Secretariate. 2019. "Rethink Robo-debt: Building a fair and accurate system people can trust." Canberra: Submission 26 to the Senate Community Affairs References Committee Inquiry into Centrelink's Compliance System [https://www.aph.gov.au/Parliamentary\\_Business/Committees/Senate/Community\\_Affairs/Centrelinkcompliance/Submissions](https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/Centrelinkcompliance/Submissions).
- Olsen, Henrik Palmer, Jacob Livingston Slosser, Thomas Troels Hildebrandt, and Cornelius Wiesener. 2019. "What's in the Box? The Legal Requirement of Explainability in Computationally Aided Decision-Making in Public Administration " University of Copenhagen, iCourts Working Paper Series No. 162, 2019 [Available at SSRN <https://ssrn.com/abstract=3402974>]
- Palmgren, Katarina. 2018. "Use of Online Dispute Resolution to Resolve Civil Disputes." Canberra: Winston Churchill Memorial Trust [https://www.churchilltrust.com.au/media/fellows/Palmgren\\_K\\_2017\\_Use\\_of\\_online\\_dispute\\_resolution\\_to\\_resolve\\_civil\\_disputes.pdf](https://www.churchilltrust.com.au/media/fellows/Palmgren_K_2017_Use_of_online_dispute_resolution_to_resolve_civil_disputes.pdf).

- Pasquale, Frank and Glyn Cashwell. 2018. "Prediction, Persuasion, and the Jurisprudence of Behaviourism." *University of Toronto Law Journal* 68: 63-81.
- Re, Richard, and Alicia Solow-Niederman. 2019. "Developing Artificially Intelligent Justice." *Stanford Technology Law Review*: forthcoming [Available at SSRN: <https://ssrn.com/abstract=3390854>].
- Salter, Shannon. 2017. "Online Dispute Resolution and Justice System Integration: British Columbia's Civil Resolution Tribunal." *Windsor Yearbook of Access to Justice* 34:112-129.
- Senate Community Affairs. 2019. Senate Standing Committee on Community Affairs, Report on 'ParentsNext, including its trial and subsequent broader rollout' [https://www.aph.gov.au/Parliamentary\\_Business/Committees/Senate/Community\\_Affairs/ParentsNext/Report/Section?id=committees%2Freportsen%2F024267%2F27170](https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/ParentsNext/Report/Section?id=committees%2Freportsen%2F024267%2F27170)
- Sleep, Lyndal, and Kieran Tranter. 2017. "The Visiocracy of the Social Security Mobile App in Australia." *International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique* 30 (3):495-514.
- Smith, Roger. 2018. "Online Dispute Resolution (ODR) and Access to Justice" *Law Technology and Access to Justice*, 1 May 2018. <https://law-tech-a2j.org/odr/online-dispute-resolution-odr-and-access-to-justice/>.
- Sourdin, Tania. 2018. "Judge v Robot? Artificial Intelligence and Judicial Decision Making." *University of New South Wales Law Journal* 41 (4):1114-1133.
- Sourdin, Tania, Bin Li, and Tony Burke. 2019. "Just, Quick and Cheap?: Civil dispute resolution and technology." *Macquarie Law Journal* 19: 17-38
- Surden, Harry. 2020 forthcoming. "e Ethics of Artificial Intelligence in Law: Basic Questions." In *Oxford Handbook of Ethics of AI*, edited by Markus Dubber, Frank Pasquale and Sunit Das, Oxford: Oxford University Press. [Available at SSRN: <https://ssrn.com/abstract=3441303>]
- Tomlinson, Joe. 2019. *Justice in the Digital State: Assessing the next revolution in administrative justice*. Bristol: Policy Press (Open Access).
- Toohey, Lisa, Monique Moore, Katelane Dart, and Dan Toohey. 2019. "Meeting the Access to Civil Justice Challenge: Digital inclusion, algorithmic justice, and human-centred design." *Macquarie Law Journal* 19:133-156.
- Veale, Michael, and Irina Brass. 2019. "Administration by Algorithm? Public Management meets Public Sector Machine Learning." In *Algorithmic Regulation*, edited by Karen Yeung and Martin Lodge, Ch. 6, in press. Oxford: Oxford University Press.
- Williams, Rebecca A. 2018. "Rethinking Deference for Algorithmic Decision-Making." Oxford Legal Studies Research Paper No. 7/2019, [Available at SSRN: <https://ssrn.com/abstract=3242482>].



- Wing, Leah. 2017. "Artificial Intelligence and Online Dispute Resolution Systems Design: Lack of/Access to Justice Magnified." *International Journal of Online Dispute Resolution* 4 (2):16-20.
- Yu, Peter. 2020. "The Algorithmic Divide and Equality in the Age of Artificial Intelligence." *Florida Law Review* 72:[Available at SSRN: <https://ssrn.com/abstract=3455772>].
- Zalnieriute, Monika, Lyria. Bennett Moses and George Williams (2019). "The Rule of Law and Automation of Government Decision-Making." *Modern Law Review* 82(3): 425-455