

Professor Philip Alston
United Nations Special Rapporteur on Extreme Poverty and Human Rights
United Nations Office of the High Commissioner for Human Rights
Palais Wilson; 52 Rue Des Pâquis
CH-1201 Geneva, Switzerland
sxtremepoverty@ohchr.org
+41 22 917 9220

4 June 2019

Submission to the OHCHR Thematic report to the United Nations General Assembly on digital technology, social protection and human rights

Monika Zalnieriute, Lyria Bennett Moses, Janina Boughey,
Lisa Burton Crawford and Nayan Bhathela
The Allens Hub for Technology, Law & Innovation
Gilbert + Tobin Centre of Public Law
UNSW Faculty of Law
Sydney, Australia

About Us

We are a group of scholars researching intersections between law & technology, constitutional and administrative law, human rights and legal theory, and collaborating under the *Technologies and Rule of Law* research stream at the UNSW Sydney Faculty of Law, which is a joint initiative of the Allens Hub for Technology, Law and Innovation ('the Allens Hub') and Gilbert + Tobin Centre for Public Law ('G+T Centre'). Allens Hub is an independent community of scholars, led by Professor Lyria Bennett Moses and based at UNSW Sydney, it aims to add depth to research on interactions among technology, law and society. More information about the Allens Hub can be found at <http://www.allenshub.unsw.edu.au/>. G+T Centre, led by Professor Rosalind Dixon and Deputy Director, Dr Lisa Burton Crawford, and staffed by a team of leading public law scholars researching in all aspects of constitutional and administrative law and public law theory. G+T Centre plays a prominent, independent role in public debate on issues vital to Australia's constitutional future including Charters of Rights, federal reform, reconciliation and native title, refugees and migration law and the challenges of responding to terrorism. More information about the G+T Centre can be found at

<http://www.gtcentre.unsw.edu.au/>. The joint research initiative on *Technologies and Rule of Law*, led by researcher Dr. Monika Zalnieriute, aims to enrich academic and policy debates and drive considered reform of law and practice through engagement with the legal profession, the judiciary, government, industry, civil society and the broader community. The opinions expressed in this submission are the views of the authors, and do not necessarily reflect or present the views or positions of the Allens Hub, G+T Centre or UNSW Law.

About this Submission

This submission seeks to respond to the questions raised in the call for submissions by the United Nations Special Rapporteur on Extreme Poverty and Human Rights into his thematic report to the United Nations General Assembly on digital technology, social protection and human rights. As scholars working at the intersection of law and technology, we are delighted to participate in this consultation process led by the Special Rapporteur. In this submission, we draw upon some of the research conducted by the Allens Hub and G+T Centre researchers to assist the Special Rapporteur in relation to the issues raised in his thematic report. The opinions expressed in this submission are the views of the authors, and not necessarily reflect or present the views or positions of the Allens Hub or the G+T Centre. We note that this research does not relate to all the questions raised in the call for submissions, and so we only set out answers in relation to those matters where our research may be relevant, namely, Questions 1, 2 and 3. We are grateful for the opportunity to present our views and hope this submission will assist in informing the Special Rapporteur's thematic report. We are especially grateful to the Office of the United Nations High Commissioner for Human Rights for providing this opportunity to make a submission past the previously advertised deadline.

Question 1: Centrelink and Robo-Debt Case in Australia

‘Robo-debt’ is the colloquial name critics gave to the Australian Government’s controversial program of automatically calculating and recovering debts owed because of welfare overpayments. As detailed by Allens Hub and G+T researchers Monika Zalnieriute, Lyria Bennett Moses and George Williams in their article ‘*The Rule of Law and Government Decision-Making*’,¹ robo-debt involved primarily the use of a pre-programmed human-authored logic system, analogous to an ‘expert system’,² that gathered and analysed income, banking, employment and other information from individuals receiving welfare payments. The system itself was not new. Data-matching had previously been used to trigger further manual enquiries about discrepancies between a welfare recipient’s reported income for tax purposes and their reported income for welfare purposes. Previously, staff from Australia’s welfare agency, Centrelink, would select discrepancies based on risk management profiling and obtain further information to ascertain whether the recipient had been overpaid. The robo-debt program essentially reversed the onus of proof, requiring welfare recipients to *disprove* that they had a debt whenever the system identified a discrepancy. In doing so, it aimed to expedite the process of identifying welfare overpayments, and therefore to significantly increase the number of overpayments that were identified and reversed per year.³

The robo-debt program was introduced as part of a 2015–16 Budget measure, ‘*Strengthening the Integrity of Welfare Payments*’, and a December 2015 Mid-Year Economic Fiscal Outlook announcement pursuant to ss 1222A and 1223 of the *Social Security Act 1991* (Cth), which establishes the legislative framework for the recovery of welfare overpayments.⁴ Under the robo-debt system, data on annual income held by the Australian Tax Office (‘ATO’) was automatically cross-matched with income reported to Centrelink. Because welfare entitlements were originally calculated on the Centrelink figure, a higher income declared to the ATO was taken to mean that the individual concerned had been overpaid and thus owed a debt to the government. The system thus combined data matching, automated assessment through the application of human-authored formulae, and automated generation of letters to welfare recipients. In Australia, such data matching is authorised by the *Data Matching Program (Assistance and Tax) Act 1990* (Cth).

¹ M. Zalnieriute, L. Bennett Moses and G. Williams, ‘The Rule of Law and Government Decision-Making’ (2019) 82(3) *Modern Law Review* 425.

² For more on expert systems, see generally A. Tyree, *Expert Systems in Law* (Sydney: Prentice Hall, 1989).

³ Commonwealth Ombudsman, *Centrelink’s Automated Debt Raising and Recovery System: A Report about the Department of Human Services’ Online Compliance Intervention System for Debt Raising and Recovery* (Investigation Report, 2017) 5 at https://www.ombudsman.gov.au/_data/assets/pdf_file/0022/43528/Report-Centrelinks-automated-debt-raising-and-recovery-system-April-2017.pdf (last accessed 15 May 2019).

⁴ For a detailed explanation of the enabling legislation, see P. Hanks, ‘Administrative Law and Welfare Rights: A 40-Year Story from *Green v Daniels* to “Robot Debt Recovery”’ (2017) 89 *AIAL Forum* 1, 7.

To understand how the robo-debt system worked, it is important to know that income in Australia is reported to the ATO as an annual figure, but to Centrelink as a fortnightly figure. The first step was to check the two annualised income figures against each other. Where their ATO annual income was greater than their Centrelink annualised income, individuals were sent an automatically-generated letter giving them an opportunity to confirm their annual income through an online portal. Those who accessed the online portal were given an opportunity to state their *fortnightly* income (with evidence), whereas those who did not access the portal were assumed to earn a fortnightly figure calculated as the annual ATO figure divided by the number of weeks in a year.⁵ However, the letter sent to individuals did not explain that recording variation in income over the year was important to an accurate calculation of welfare entitlements.⁶ Early versions of the letter were also worded in a way that appeared to some recipients to be a debt notice. The fortnightly income (entered into the online system or derived as above) was used to calculate what the welfare entitlement ought to have been and, where relevant, individuals were automatically sent a debt notice. Debts were raised against every person who did not or could not *disprove* that they were overpaid by producing evidence. In some cases, evidence was required for income earned up to seven years earlier. Error rates have been calculated at approximately 20 per cent.⁷ In the 18 months between July 2016 and December 2018, 445 000 debts were raised through this system, 77 500 of which have been reduced, waived or written off.⁸ This may be an underestimate of the number of erroneous notices due to the number of people who may simply have paid the ‘debt’ without challenge.

The robo-debt program has been criticised by legal experts for being unlawful due to its reversal of the onus of proof, and for using a formula to calculate debts which was known to be flawed.⁹ An inquiry by the Commonwealth Parliament Senate Community Affairs References Committee found that the program breached fundamental principles of procedural fairness (or due process),¹⁰ and recommended that it be put on hold until these issues were rectified.¹¹ The Senate Committee inquiry and a separate inquiry by the Commonwealth Ombudsman found that the program had been poorly designed and implemented. In particular:¹²

⁵ Commonwealth Ombudsman (n 3), 1, 4.

⁶ *ibid* at 9.

⁷ Senate Community Affairs References Committee, Commonwealth Parliament, *Design, scope, cost-benefit analysis, contracts awarded and implementation associated with the Better Management of the Social Welfare System initiative* (21 June 2017) 32–33 <aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/SocialWelfareSystem/Report>

⁸ Department of Human Services, Answer to Question on Notice (HS 34 (SQ19-000027)), Senate Community Affairs Legislation Committee, Additional Estimates 2018-19 (21 February 2019).

⁹ Hanks (n 4).

¹⁰ Australia does not have a federal bill of rights, but the administrative law principle of procedural fairness is substantially the same as due process rights in other jurisdictions, see J. Boughey, *Human Rights and Judicial Review in Australia and Canada: The Newest Despotism?* (Hart Publishing, 2017) Chapter 4.

¹¹ Senate Community Affairs References Committee (n 7), ch 6.

¹² Senate Community Affairs References Committee (n 7), ch 6; Commonwealth Ombudsman (n 3).

- The misleading wording of the letters sent to welfare recipients had caused confusion and stress.
- Vulnerable people, in particular, had found it difficult to obtain the evidence required to disprove their debts, and there was a lack of support available to them.
- The imposition of an automatic 10% recovery fee was punitive and unfair.
- The program lacked transparency, and people were not given information about how their debts were calculated, grounds for challenging their debts, or how to challenge debts.
- The online portal through which people were required to provide evidence disproving their debts was not user-friendly, particularly for vulnerable people, and millions of phone calls to Centrelink went unanswered.

As of 31 December 2018, the program had cost the Government approximately \$475 million.¹³ The Government had raised approximately \$1.5 billion of total debts and received approximately one third of that (\$500 million) in repaid debts.¹⁴

Questions 2 & 3: Lessons from Robo-Debt and Relevance of Rule of Law Values

While the robo-debt system has been progressively modified over time, there are important lessons to be learned from its originally flawed design and implementation in relation to government decision-making generally. A range of government bodies, including executive agencies, parliamentary bodies, the judiciary,¹⁵ police forces and national security agencies, are increasingly deploying tools, including expert systems and machine learning, to automate decision-making or improve their services. These developments have the potential to advance human rights, for example by improving transparency, consistency and equity in high volume decision-making and reducing discrimination. However, they can also impinge on human rights, particularly when poorly implemented. Of particular concern are due process rights,

¹³ Comprising \$375 million from July 2015 to June 2018, plus half of the \$200 million budgeted for the 2018-19 financial year): Department of Human Services, Answer to Question on Notice (HS9 (SQ19-000022)), Senate Community Affairs Legislation Committee, Additional Estimates 2018-19 (21 February 2019).

¹⁴ Senate Community Affairs Legislation Committee, Additional Estimates, Canberra, 21 February 2019, 120-1 (Renee Leon).

¹⁵ On the use of automation by the judiciary, see M. Zalnieriute and F. Bell, 'Technology and Judicial Role' forthcoming in G. Appleby and A. Lynch (eds.), *The Judge, the Judiciary and the Court: Individual, Collegial and Institutional Judicial Dynamics in Australia* (UK: Cambridge University Press, 2020).

privacy, and the right to be free from discrimination. The use of automated systems in government decision-making also presents considerable challenges for the traditional public law oversight mechanisms and principles, particularly the principles of transparency, rationality and accountability. Allens Hub and G+T researchers have analysed legal implications of the automation of different areas and types of government decision-making in their recent work, which the Special Rapporteur might find useful.¹⁶

It is important that any automated systems designed for social protection or any other policy areas address the concerns which arose in the robo-debt case. These include poorly-worded correspondence, inaccuracy of the formula applied in a percentage of cases, issuing debt notices to those not owing money,¹⁷ shifting the burden of proof,¹⁸ and leaving individuals to the mercy of debt collectors.¹⁹ The remainder of this submission will analyse the implications that such flaws in the robo-debt system have on the values of the rule of law. The rule of law is a core constitutional value in most established legal orders, including Australia.²⁰ Our focus in doing so is on three core rule of law concepts that, we suggest, have the widest acceptance across political and national systems: transparency and accountability; predictability and consistency; and equality before the law. We suggest that the lessons to be learned from robo-debt for other areas of public policy and government decision-making focus on the alignment of automated systems with rule of law values, which ultimately hinges on the appropriateness of design choices.

Transparency and Accountability

The automation of government decision-making through the adoption of rules-based expert systems such as robo-debt offers potential benefits in enhancing transparency and accountability, particularly in the area of social protection. As Zalnieriute, Bennett Moses and Williams argue, '[w]hereas a human may come up with justifications for a decision *ex post* that do not accurately represent why a decision was made,²¹ a rules-based system can explain precisely how every variable was set and why each conclusion was reached'.²² It is important to note that such feedback is not *necessarily* provided in rules-based expert systems. The

¹⁶ See M. Zalnieriute, L. Burton Crawford, J. Boughey, L. Bennett Moses and S. Logan, 'From the Rule of Law to Statute Drafting: Legal Issues for Algorithms in Government Decision-Making,' forthcoming in W. Barfield (ed.) *Cambridge Handbook on Law and Algorithms* (UK: Cambridge University Press, 2019).

¹⁷ T. Carney, 'The New Digital Future for Welfare: Debts without Legal Proofs or Moral Authority?' (UNSW Law Journal Forum, May 2018) at <http://www.unswlawjournal.unsw.edu.au/wp-content/uploads/2018/03/006-Carney.pdf> (last accessed 16 August 2018).

¹⁸ Hanks (n 4) 9–11.

¹⁹ Note that this aspect of the program has been modified: Commonwealth Ombudsman (n 3) 9–10, 36, 38.

²⁰ See further Lisa Burton Crawford, *The Rule of Law and the Australian Constitution* (Federation Press, 2017).

²¹ R.E. Nisbett and T. DeCamp Wilson, 'Telling More Than We Can Know: Verbal Reports on Mental Processes' (1977) 84 *Psychological Review* 231.

²² Zalnieriute, Bennett Moses and Williams (n 1) 440.

designer decides what the output of the system will be and whether it will include reasons for its conclusions or decisions. In the case of robo-debt, as explained above, individuals were not provided with clear information as to how debts were calculated in general, or in their individual case.

Thus, where decisions are fully or partially automated in social protection systems, the transparency and accountability of outputs hinges on the accountability of those designing the system *for* the transparency and accountability of the decision-making system itself. Indeed, a similar point is true for all rule of law values. They are unlikely to be found in decision-making and decision-support systems by accident. Therefore those designing such automated systems should be required to design them in ways consistent with the rule of law and be able to give an account of how this has been done.²³ As robo-debt illustrates, greater in-built transparency could lead to a better understanding of how these systems operate, as well as their underlying values, thus revealing what is now obscure. More transparency would also allow affected individuals to challenge such decision-making systems, because information about the variables, inputs and outputs would be available.²⁴

Predictability and Consistency

Automation can also improve the predictability and consistency of government decision-making in the area of social protection. As Zalnieriute, Bennett Moses and Williams suggest, '[u]nlike humans, computer systems cannot act with wanton disregard for the rules with which they are programmed'.²⁵ However, automation also poses many challenges for the rule of law principles of predictability and consistency. One such challenge arises when the code which is applied in an automated decision-making process does not correspond with statutory, common law or policy requirements. The inconsistency in such a case is not in the application of the rule in different cases, but between the rule as formulated and the rule as applied in every case. The robo-debt system is an unfortunate example of such an inconsistency, as the formula used by the automated system in question failed to produce the correct result for many people.²⁶ This was exacerbated by the fact that the processes in place to manage errors were insufficient. There was no human checking of the decision to issue a debt notice. The notice itself was also presented to individuals as a *fait accompli*, with some individuals not receiving earlier

²³ *ibid* at 444.

²⁴ S. B. Starr, 'Evidence-Based Sentencing and the Scientific Rationalization of Discrimination' (2014) 66 *Stanford Law Review* 803, 806.

²⁵ Zalnieriute, Bennett Moses and Williams (n 1) 445.

²⁶ There is some dispute about the rate of error and how these should be characterised. Approximately 20 per cent of people who received debt notices succeeded in providing additional information that demonstrated that no debt was owed: Senate Community Affairs References Committee, Parliament of Australia, *Design, Scope, Cost-Benefit Analysis, Contracts Awarded and Implementation Associated with the Better Management of the Social Welfare System Initiative* (2017) at [2.88].

communications due to address errors.²⁷ The online portal in place to deal with challenges to debt notices was also hard to use,²⁸ with human alternatives inadequate to meet the demand.²⁹ The rate of errors also potentially exceeded the capacity of institutions designed to deal with appeals. Hence, what resulted was a far higher likelihood that the law is being misapplied in ways that are unpredictable and inconsistent.

What these failures in the robo-debt system elucidate is that the benefits of automation in improving the predictability and consistency of decision-making in the area of social protection, as well as government decision-making more generally, can only be realised if the automation process is sufficiently transparent, if it is properly evaluated (for accuracy and for consistency with legal requirements), and if appropriate measures are put in place to manage foreseeable errors. Such measures should include human checking of outputs, clear explanation as to the potential for error and the circumstances in which error can arise, and a transparent and sufficiently resourced process for appeals. As Zalnieriute, Bennett Moses and Williams suggest, these are all questions of design.³⁰

Equality before the Law

We suggest that automation in the area of social protection has the potential to enhance the principle of equality before the law by reducing arbitrariness in the application of law, removing bias and eliminating corruption. Systems that give the same answer when presented with the same inputs help to ensure that similarly situated individuals are treated equally. However, automation of decision-making can compromise individual due process rights because it may undermine the ability of that person to influence or challenge a decision affecting them.³¹ This may be, for example, because they are unable to access or determine the correctness of key information used to make that decision. The robo-debt system is an example of this, as the right to review and rectify information was undermined because the letter sent to individuals by the government did not explain the importance of the income variation over the year for an accurate calculation of welfare entitlements.³² Issuing debt notices without providing a genuine opportunity to correct the erroneous data effectively denied welfare recipients due process rights.³³ Further challenges to equality before the law and due process

²⁷ *ibid* at [3.61].

²⁸ *ibid* at [2.110].

²⁹ *ibid* at [3.98], [3.106], [3.107], [3.119].

³⁰ Zalnieriute, Bennett Moses and Williams (n 1) 448.

³¹ *ibid* at 449.

³² D.L. Kehl, P. Guo and S.A. Kessler, 'Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing' (Harvard Law School Student Paper, Responsive Communities Initiative, Berkman Klein Center for Internet & Society, July 2017) 9.

³³ On the similarity between due process 'rights' and the common law duty of administrative decision-makers to provide procedural fairness see Boughey (n 10) Chapter 4.

safeguards can arise in some cases of automated decision-making due to what could be described as a ‘reversal’ of the burden of proof or the lowering of ‘evidence threshold’.³⁴ This was also illustrated in the robo-debt case, as explained above.³⁵ While debts issued under this automated decision-making process can be challenged, it has been argued that the government failed its responsibility to ensure that it had established the existence of the debt before initiating the claim.³⁶

In understanding the benefits and challenges of automating decisions in relation to social protection, it is crucial to consider both the context of the decision and the type of system deployed.³⁷ A system with pre-programmed rules such as robo-debt could ensure that decisions are made based on factors recognised as legally relevant and hence avoid or minimise the risk of corruption or bias by officials. However, procedural rights and opportunities to check and rectify data on which the decision relies are crucial, as is ensuring that the logic of the system accurately reflects the law. Rule-based expert systems must be designed to include such procedural rights to ensure that there is a systematic and easily-accessible method to exercise such rights as and when required. However, it is important to note that pre-programmed systems, such as robo-debt, should be distinguished from systems based on machine learning and data-driven classification and prediction *about humans* that raise deeper issues for equality before the law. By definition, these systems seek to classify individuals based on their similarity to others who have committed offences or proved to be more in need of particular services. We suggest that such systems should never be used by governments as a basis for defining access to social protection services (or other public services) not only because of the virtual inevitability of discrimination on prohibited characteristics, such as racial or ethnic background, but also because it is not an appropriate logic for allocating government entitlements.

³⁴ On the importance of burden of proof and ‘evidence threshold’ in the context of social welfare in the US, see L. Kaplow, ‘Burden of Proof’ (2012) 121 *Yale Law Journal* 738. For Australia, see, e.g., A. Gray, ‘Constitutionally Protecting the Presumption of Innocence’ (2012) 31(1) *University of Tasmania Law Review* 13. In the context of European Court of Human Rights, see M. Ambrus, ‘The European Court of Human Rights and Standards of Proof: An Evidential Approach Toward the Margin of Appreciation’ in L. Gruszczynski and W. Werner (eds), *Deference in International Courts and Tribunals: Standard of Review and Margin of Appreciation* (Oxford: Oxford University Press, 2014). On due process implications of shifting the burden of proof in the US legal context, see C.M.A. McCauliff, ‘Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees’ (1982) 35 *Vanderbilt Law Review* 1293; P. Petrou, ‘Due Process Implications of Shifting the Burden of Proof in Forfeiture Proceedings Arising out of Illegal Drug Transactions’ [1984] *Duke Law Journal* 822.

³⁵ Hanks (n 4).

³⁶ Carney (n 17).

³⁷ On the use of automation by the judiciary, see M. Zalnieriute and F. Bell (n 15).

Conclusions on Lessons for the Rule of Law Values

Generally speaking, automation can improve government decision-making, particularly in relation to social protection. Zalnieriute, Bennett Moses and Williams suggest that ‘[t]he benefits include cost savings and greater speed, as well as a capacity to enhance the rule of law’.³⁸ However, as the robo-debt case study clearly illustrates, the extent to which automated rules-based decision-making can enhance rule of law values depends entirely on the quality with which such systems are designed and implemented. The most important lesson from the robo-debt case study is that in order for automated systems to be compliant with the rule of law values, there must be human oversight of, and accountability for, decisions made by such automated systems. Such oversight and accountability is particularly important in relation to design choices, implementation, procedural fairness and avenues of appeal. Although converting rule of law values into design specifications that can be understood by system designers and enforced through courts,³⁹ regulation, professional standards, contracts or other mechanisms represents a formidable technical and legal challenge, it is crucial for such a challenge to be tackled in order to ensure that automated systems adhere to the rule of law.⁴⁰ The failures of the robo-debt system illustrate how important it is to incorporate rule of law values into automated systems, especially in relation to social protection frameworks.

³⁸ Zalnieriute, Bennett Moses and Williams (n 1) 454.

³⁹ See in particular Zalnieriute and Bell (n 15).

⁴⁰ Zalnieriute, Bennett Moses and Williams (n 1) 455.