THE USES OF AI IN GOVERNMENT DECISION-MAKING: IDENTIFYING THE LEGAL GAPS IN AUSTRALIA

*Moira Paterson**

INTRODUCTION

The Australian government is increasingly implementing, or considering the implementation of, new technologies based on Big Data Analytics ("BDA"), artificial intelligence ("AI"), and machine learning. This development "corresponds with tight fiscal constraints on governments globally and with rapid growth in the volume, complexity, and subject-matter of decisions made by governments."[1] While it has the potential to improve the efficiency of government service delivery, the use of BDA as a basis for decision-making and actions raises several ethical and legal concerns. The key issues of concern derive from the opaqueness of AI-based decision-making, including its lack of explicability and its potential for bias and lack of accountability.[2]

* Professor of Law, Monash University.

[1] Just. Melissa Perry, Fed. Court of Austl., iDecide: Digital Pathways to Decision, Speech at the 2019 CPD Immigration Law Conference (Mar. 21-23, 2019), http://classic.austlii.edu.au/au/journals/FedJSchol/2019/3.html [https://perma.cc/ERF3-AN38].

[2] For useful overviews of these issues, see FREDERIK ZUIDERVEEN BORGESIUS, COUNCIL OF EUROPE, DISCRIMINATION, ARTIFICIAL INTELLIGENCE, AND ALGORITHMIC DECISION-MAKING (2018), https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73 [https://perma.cc/9QZL-KLAE]; Jenna Burrell, *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms*, BIG DATA & SOC'Y, Jan.-June 2016, at 1.

The issues of opaqueness and lack of explicability arise from the complexity of the processes embedded in algorithms. They are particularly significant in the case of algorithms that utilise machine learning, a process whereby the algorithm self-adjusts based on its own analyses of data previously encountered (known as training data).[3] The use of machine learning makes it difficult to know what is going on inside the program and to understand how it arrives at its conclusions. Machine learning frequently makes use of neural networks,[4] which makes the process of decision-making difficult to explain because it involves the encoding of "[i]mplicit knowledge . . . in numeric parameters—called weights—and distributed all over the system."[5] This problem has been characterised as involving a "mismatch between mathematical optimization in [a] high-dimensionality characteristic of machine learning and the demands of human-scale reasoning and styles of semantic interpretation."[6]

The issue of bias may arise for several different reasons. Barocas and Selbst have identified four sources of unintentional bias in decision-making:

- The definition of "target variable" and "class labels" (i.e., what is being looked for when sorting data and the categories into which it is to be sorted);

- Training data that contains mislabeled examples or biased data collections;

- The methodology used for feature selection (the process used to "make choices about what attributes [are] observe[d] and subsequently fold[ed] into their analyses"); and

---

[3] This process frees the algorithm from predefined preferences.

[4] Margaret Rouse, *artificial neural network (ANN)*, SEARCHENTERPRISEAI, https://searchenterpriseai.techtarget.com/definition/neural-network [https://perma.cc/Y9C6-M2X7] (last updated Aug. 2019) ("[A]n artificial neural network (ANN) is a system of hardware and/or software patterned after the operation of neurons in the human brain.").

[5] Keith Darlington, *Explainable AI Systems: Understanding the Decisions of the Machines*, OPENMINDBBVA (Oct. 11, 2017), https://www.bbvaopenmind.com/en/technology/artificial-intelligence/explainable-ai-systems-understanding-the-decisions-of-the-machines/ [https://perma.cc/GN3E-AGBQ].

[6] Burrell, *supra* note 2, at 2.

- Proxies (where the process of sorting, while not based on protected attributes, may nevertheless use criteria that serve as reliable proxies for membership of a protected class).[7]

The public law frameworks that apply to government decision-making are designed to protect the rule of law and core rights and freedoms. Openness, fairness, accountability and rationality are regarded as fundamental principles of administrative law, in addition to accessibility of judicial and non-judicial grievance procedures, legality and impartiality.[8] In the case of human rights, the three key rights that are implicated are: (1) the right to equality of treatment, (2) the right to privacy, and (3) the right to information, which derives from the right to freedom of expression. They are important from a democratic perspective because "[d]emocracy relies on core commitments to governance by consent of the people whose voices are heard, to governance respectful of the dignity of individuals, to governance constrained by prior commitments to the rule of law, and to governance made accountable by the openness of its processes."[9]

This Article sets out to examine the interaction between uses of automated technologies in Australian government decision-making and the key public law frameworks that govern their use, and to identify and consider what measures are necessary to address the tensions between them.

I. INTERACTIONS WITH KEY PUBLIC LAW FRAMEWORKS

The regulatory frameworks that are relevant to this analysis fall into three groups: (1) administrative law, (2) anti-discrimination law, and (3) information law (including freedom of information and privacy laws.

---

[7] Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671, 688, 715-22 (2016).

[8] MARK ARONSON, BRUCE DYER & MATTHEW GROVES, JUDICIAL REVIEW OF ADMINISTRATIVE ACTION 1 (3d ed. 2004).

[9] Judith Resnik, *Judicial Selection and Democratic Theory: Demand, Supply, and Life Tenure*, 26 CARDOZO L. REV. 579, 591 (2005).

Administrative law

Australia has a comprehensive system of administrative law that includes judicial review by the courts as well as other avenues for review, including review on the merits. The grounds for judicial review derive from the common law but are codified and expanded upon in the *Administrative Decisions (Judicial Review) Act 1977* (Cth), which provides the main avenue of federal judicial review. These rules have been developed in a context where it has been assumed "that the decision-maker using the executive power conferred by Parliament is a human being or an institution composed of humans and that there is a human who will be accountable and responsible for the decision."[10]

Ultra Vires/Illegality

The principle of legality requires that a decision must be made by an authorised decision-maker. In the absence of specific conferral of decision-making power on a computer, "[t]echnology-assisted decision-making assumes a human decision-maker, who will need to be authorised to make the decision, either by the statute conferring the decision-making power or function or through a delegation."[11] This means that:

> if an automated system is being utilised to make part of all of a decision, the use of that system must be authorised. It cannot be assumed that a statutory authority vested in a senior public servant which extends by implication to a properly authorised officer, will also extend to an automated system; nor that authority to delegate to a human decision-maker will permit "delegation" to an automated system.

[10] Marion Oswald, *Algorithm-Assisted Decision-Making in the Public Sector: Framing the Issues Using Administrative Rules Governing Discretionary Power*, PHIL. TRANSACTIONS ROYAL SOC'Y A, Aug. 6, 2018, at 3 (quoting Andrew Le Sueur, *Robot Government: Automated Decision-Making and its Implications for Parliament*, *in* PARLIAMENT: LEGISLATION AND ACCOUNTABILITY 183 (Alexander Horne & Andrew Le Sueur eds., 2016)).

[11] Katie Miller, *The Application of Administrative Law Principles to Technology-Assisted Decision-Making*, 86 AUSTL. INST. ADMIN. L. FORUM 20, 22 (2016).

> Authority to use such systems should be transparent and express.[12]

This matter is dealt with in a number of Commonwealth statutes, via the inclusion of provisions that specifically authorise the person to whom the decision-making authority is allocated to arrange for the use, under their control, of computer programs to make specified decisions under the specific statute.[13] Further, these provisions deem any decision made by the operation of a computer program, as permitted, to be a decision made by that person.[14]

The requirement that a decision-maker has acted within their authority (i.e., that they have not done "something for which they lack legal authority," and in that way exceeded the limits of their power)[15] presents two specific challenges. The first is how to ensure that the rules encoded in a program reflect not just the words of the provision that governs the decision but also the rules of statutory interpretation that govern its interpretation. The second is how to ensure that the system reflects ongoing changes in statutory provisions.

The complexity of the challenges in ensuring that the rules reflect an accurate interpretation of the applicable law has been explained as follows: "laws are interpreted in accordance with statutory presumptions, and meaning is affected by context. These are not necessarily simple questions and the potential for coding error is real."[16]

The second challenge arises from the non-static nature of statutory provisions, which means that automated systems "will need to be kept up to date while maintaining the capacity to apply

---

[12] Just. Melissa Perry, *iDecide: Administrative Decision-Making in the Digital World*, 91 AUSTL. L.J. 29, 31 (2017).

[13] For example, the *Aged Care Act* provides that "[t]he Secretary may arrange for the use, under the Secretary's control, of computer programs for making decisions on the making or varying of determinations under this Division." *Aged Care Act 1997* (Cth) s 23B-4(1) (Austl.).

[14] *See, e.g.*, *Child Support (Assessment) Act 1989* (Cth) s 12A (Austl.).

[15] *See* Jennifer Cobbe, *Administrative Law and the Machines of Government: Judicial Review of Automated Public-Sector Decision-Making*, 39 LEGAL STUDIES 636, 642 (2019) (citing Regina v. Lord Chancellor, ex parte Witham [1998] Q.B. 575 (UK)).

[16] *See* Perry, *supra* note 1.

the law as it stands at previous points in time for decisions caught by transitional arrangements."[17]

Another administrative law requirement is that decision-making must take place within the bounds of the law.[18] The grounds for *extended* ultra vires generally present greater complexity. This is especially so in the case of the duty to take into account all relevant issues and to not take into account any irrelevant issues. This presents challenges because of the large quantities of data involving a multiplicity of variables, both in training models and in the decision-making systems themselves. Therefore, it is important, but also difficult in practice, to be able to identify these variables and demonstrate that they are relevant to any particular decision. A further issue relates to the ability of machine learning systems to infer and predict new information from datasets. This arguably requires demonstration that the system has not inferred, or predicted, and then considered information that is incorrect or irrelevant to the decision.

Another aspect of extended ultra vires that presents challenges is the rule that prohibits fettering the exercise of discretionary powers. As explained by Cobbe, this requires that decision-makers should make a decision "on its merits rather than adopting a one-size-fits-all approach, and they should be prepared to depart from policies or guidelines where appropriate."[19] That may be problematic to the extent that the system applies a statistical model or rule uniformly to all its decisions. This produces outputs that should be "consistent" but do "not facilitat[e] consideration of the particulars of the case at hand" where they deviate from the norm or for some other reason require individual consideration.[20]

A further challenge relates to the need to be able to demonstrate the reasonableness/rationality of a decision and to provide reasons for it. This arises from the risk that problems in

[17] Just. Melissa Perry, Fed. Court of Austl., & Alexander Smith, iDecide: The Legal Implications of Automated Decision-Making, Speech at the Cambridge Centre for Public Law Conference: Process and Substance in Public Law (Sept. 15, 2014), https://www.fedcourt.gov.au/digital-law-library/judges-speeches/justice-perry/perry-j-20140915 [https://perma.cc/966G-D99T].

[18] *See* Cobbe, *supra* note 15.

[19] *Id.* at 646.

[20] *Id.* at 647.

the design of algorithms (or in the content or selection of the data used to train them) may result in decisions that comply with the rules of logic or formal argument but are nevertheless unreasonable. This may arise in part from the ability to use BDA to make "non-intuitive and unverifiable inferences and predictions."[21]

A final issue concerns the manner and extent to which any relevant evidence is taken into account. As summarised by Hogan-Doran, an automated decision may be challenged on a number of grounds relating to the insufficiency of evidence, including the "no evidence rule of procedural fairness."[22] In addition, a decision will lack "probative evidence, to the extent there that it lacks any basis or is unjustifiable on, or is unsupported by, the available evidence."[23]

Procedural Impropriety

A key aspect of the duty of procedural fairness, which is owed to persons whose "rights" or "interests" are likely to be affected by a decision,[24] is the right to be heard. This requires disclosure to them of any adverse information and providing them an opportunity to respond to that material. The duty of procedural fairness is flexible in content and varies according to the context in which the decision is made. Automated decision systems may be ill-suited to incorporate these features. Furthermore, they require care in their implementation to ensure that any procedural protections provided are readily accessible to individuals potentially impacted by adverse decisions. For example, to the extent that individuals are required to interact with a program via computer, it is important that the system is designed to ensure that they are able to exercise their procedural rights without any

[21] *See* Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2019 COLUM. BUS. L. REV. 494, 497.

[22] Dominique Hogan-Doran, *Computer says "no": automation, algorithms and artificial intelligence in Government decision-making*, 13 JUD. REV. 345, 356-58 (2017) (Austl.).

[23] *Id.* at 357.

[24] *See* Kioa v West [1985] 159 CLR 550, 582 (Mason J) (Austl.).

undue difficulty, bearing in mind variations in expected levels of computer literacy.

Amenability to Review

The administrative law issues associated with the use of algorithms can potentially be obviated by requiring some element of human involvement or oversight.[25] However, this will depend on the extent of human agency involved and to which the individual decision-makers concerned feel comfortable in disagreeing with the decision recommended by an *expert system*.

There are two potential avenues for independent review of automated decisions—review on the merits by the Administrative Appeal Tribunal and judicial review by the courts. As noted by Miller, judicial review may be "more difficult and expensive because of the need to engage with, and understand, what the technology is doing."[26] She notes that courts consider two basic issues when reviewing any administrative decision: (1) "What did the statute require?" and (2) "Was that in fact what occurred in this decision?"[27]

Review on the merits[28] and judicial review are generally available only where there is a "decision." For example, the Federal Court's jurisdiction to provide review under the *Administrative Decisions (Judicial Review) Act 1977* (Cth) ("ADJR Act") relates to a decision of an administrative character made under an enactment.[29] The meaning of the term "decision" in this context was considered by the Federal Court in 1999 in *Semunigus v The Minister for Immigration & Multicultural Affairs*,[30] where Justice Finn interpreted it as involving:

---

[25] This is the rational that underlies the restrictions on automated processing in Article 22 of the EU's General Data Protection Regulation. *See* 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) art. 22, 2016 O.J. (L 119), https://eur-lex.europa.eu/eli/reg/2016/679/oj [https://perma.cc/THQ2-WC52] [hereinafter GDPR].

[26] Miller, *supra* note 11, at 27.

[27] *Id.*

[28] The principle avenue for review on the merits is via the *Administrative Appeals Tribunal Act 1975* (Cth) (Austl.).

[29] *Administrative Decision (Judicial Review) Act 1977* (Cth) ss 3, 5 (Austl.).

[30] [1999] FCA 422 (Austl.).

> both reaching a conclusion on a matter as a result of a mental process having been engaged in and translating that conclusion into a decision by an overt act of such character as, in the circumstances, gives finality to the conclusion—as precludes the conclusion being revisited by the decision-maker at his or her option before the decision is to be regarded as final.[31]

In *Pintarich v Deputy Commissioner of Taxation*[32] the Full Court of the Federal Court of Australia considered whether an automatically generated letter, which stated that no general interest charge would be imposed in respect of an amount of outstanding tax, constituted a "decision." The agency officer with authority to decide the terms of repayment had inputted data into a number of fields, which generated a letter that had been sent out from another place.[33] The officer said that he had not seen the letter at the time it was created and that it did not reflect what he intended.[34] It was argued by Mr. Pintarich that the automated correspondence was evidence of a "decision" by the Deputy Commissioner to waive this charge.[35] This argument was rejected by the majority on the basis that there had been no mental process to reach a conclusion.[36] Justice Kerr, in dissent, disagreed and made the point that:

> The hitherto expectation that a 'decision' will usually involve human mental processes of reaching a conclusion prior to an outcome being expressed by an overt act is being challenged by automated 'intelligent' decision making systems that rely on algorithms to process applications and make decisions.[37]

He also commented: "What was once inconceivable, that a complex decision might be made without any requirement of

---

[31] *Id.* para 19. This interpretation was later upheld by the Full Court of the Federal Court. *See* Semunigus v The Minister for Immigration & Multicultural Affairs (2000) 96 FCR 533 paras 55, 101 (Austl.).

[32] [2018] FCAFC 79 (Austl.).

[33] *Id.* para 19.

[34] *Id.*

[35] *Id.* para 133.

[36] *Id.* para 140.

[37] *Id.* para 46 (Kerr, J).

human mental processes is, for better or worse, rapidly becoming unexceptional."[38]

As pointed out by Ng and O'Sullivan,[39] "*Pintarich* throws into uncertainty whether the determinations made utilising the numerous current government automated decision-making schemes would involve making a decision under the *ADJR Act*."[40] They further suggest:

> Where an automated system is exercising discretion and making a decision based on its machine learning, AI or algorithms, then clearly this would not be a decision based on the majority's decision in *Pintarich*, as there is no human mental process of deliberation involved and therefore no decision under the *ADJR Act*.[41]

Anti-Discrimination Law

The right to equality at the federal level in Australia is legislated for via a series of laws which protect specific attributes—the *Age Discrimination Act 2004* (Cth),[42] the *Disability Discrimination Act 1992* (Cth),[43] the *Sex Discrimination Act 1984* (Cth),[44] and the *Racial Discrimination Act 1975* (Cth).[45] The *Racial Discrimination Act* contains a general prohibition against discrimination on the grounds of "race, colour, descent or national or ethnic origin,"[46] whilst the others prohibit direct and

---

[38] *Id.* para 47.

[39] Yee-Fui Ng & Maria O'Sullivan, *Deliberation and Automation – When is a Decision a "Decision"?*, 26 AUSTL. J. ADMIN. L. 21 (2019).

[40] *Id.* at 30.

[41] *Id.*

[42] *Age Discrimination Act 2004* (Cth) pt 3 (Austl.) (prohibiting direct or indirect discrimination of the grounds of age).

[43] *Disability Discrimination Act 1992* (Cth) pt 2 (Austl.) (prohibiting direct and indirect discrimination on the grounds of disability).

[44] *Sex Discrimination Act 1984* (Cth) pt 2 (Austl.) (prohibiting direct and indirect discrimination on the grounds of sex, sexual orientation, gender identity, intersex status, marital or relationship status, pregnancy or potential pregnancy, breastfeeding or family responsibilities).

[45] *Racial Discrimination Act 1975* (Cth) (Austl).

[46] *Id.* pt 2, s 9(1).

indirect[47] discrimination in other specific contexts, including in respect of the administration of Commonwealth laws and programs.

As discussed above, AI-based automated processing can result in bias and discrimination in respect of protected attributes, including discrimination based on proxies. This may be difficult to identify given the multiple ways in which it can arise, and what testing and evidence may be necessary to ensure that it is not resulting in forms of discrimination that are illegal. The use of AI-based decision-making has been identified as presenting two specific challenges in respect of Australian anti-discrimination laws.

First, it makes it possible to discriminate unfairly based on attributes that do not currently receive protection, including on the basis of "a whole range of random factors."[48] A key issue that has been identified elsewhere as a matter of concern is discrimination on the basis of social status or social disadvantage.[49]

A second issue, which has received attention in the context of employers' use of talent analytics, is that it challenges the tests used to define discrimination.[50] Burdon and Harpur contend that that the practices involved in talent analytics "are not automatically covered by anti-discrimination laws because they do not habitually involve decisions regarding a protected attribute," despite having an exclusionary impact.[51] They point out that the test for direct discrimination requires a complainant to establish that the alleged discriminator treated them less favourably

---

[47] *Id.* pt 2, s 9(1A)(a) (prohibiting discrimination through use of compliance "with a term, condition or requirement" that is "not reasonable having regard to the circumstances of the case").

[48] *See* Mark Burdon & Paul Harpur, *Re-Conceptualising Privacy and Discrimination in an Age of Talent Analytics*, 37 U.N.S.W. L.J. 679, 696 (2014) (Austl.); *see also* Oscar H. Gandy Jr., *Engaging Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems*, 12 ETHICS & INFO. TECH. 29, 37 (2010).

[49] *See generally* Maddalena Favaretto, Eva De Clercq & Bernice Simone Elger, *Big Data and discrimination: perils, promises, and solutions. A systematic review*, 6 J. BIG DATA 12 (2019).

[50] *See* Burdon & Harpur, *supra* note 48, at 702 ("In essence, they challenge the clearly defined notions of discrimination based on protected attributes.").

[51] *Id.* at 696.

because the complainant has a protected attribute and that this will be very difficult to establish in practice.[52] They also note that the requirements for indirect discrimination have been interpreted in a highly technical fashion by the Australian High Court in *State of New South Wales v Amery*.[53] In addition, they also make the point that, "[w]hile it may be possible to establish that scoring well on the analytical tool could constitute a requirement, complainants may find it difficult to establish that they experience disadvantage due to the application of a random informational attribute."[54]

Privacy Law

The key privacy issues arising from AI-based government-decision-making relate to the collection, use and disclosure of individuals' personal information. AI-based automated systems generally rely on the availability of extensive personal information, both for their operation and for training purposes.

The key law which protects privacy at the federal level is the *Privacy Act 1988* (Cth). This requires specified entities, including Commonwealth government agencies,[55] to comply with a set of Australian Privacy Principles ("APPs").[56] This contains a number of features which limit its ability to protect personal privacy in the context of AI-based decision-making.

A key feature of the APPs is that they impose limitations on the collection, use and disclosure of "personal information," a term which is defined with reference to a test of identifiability.[57] For information to qualify as "personal information" it must be "about an identified individual, or an individual who is reasonably identifiable."[58] In *Privacy Commissioner v Telstra Corporation*

[52] *Id.* at 697-98.

[53] *Id.* at 698 (discussing State of New South Wales v Amery [2006] 230 CLR 174 (Austl.)).

[54] *Id.* at 699.

[55] *Privacy Act 1988* (Cth) s 6(1) (Austl.).

[56] *Id.* s 15.

[57] *Id.* s 6(1) (defining "personal information" as "information or an opinion about an identified individual, or an individual who is reasonably identifiable: (a) whether the information or opinion is true or not; and (b) whether the information or opinion is recorded in material form or not").

[58] *Id.*

Limited,[59] the Full Court of the Federal Court upheld a decision to the effect that an individual's telecommunications metadata was not information about him but rather related to the way in which the telco provider delivered its service. In reaching its conclusion, the court commented that:

> The words "about an individual" direct attention to the need for the individual to be a subject matter of the information or opinion. This requirement might not be difficult to satisfy. Information and opinions can have multiple subject matters. . . . This will require an evaluative conclusion, depending upon the facts of any individual case . . . .[60]

This approach is problematic because a narrow approach to the question of when information is about an individual may limit privacy protection in the context of BDA.

The limitation of protection to information that is identifiable limits privacy protection because the new BDA environment makes the re-identification of information much easier than in the past.[61] This creates privacy concerns in the context of automated decision-making based on profiling because "de-identified" information can potentially be recombined with identifiable information to produce more detailed profiling.

A similar type of issue arises where data concerning groups (which is not *per se* protected because the APPs apply only to personal data) is added to the profiles of individuals that are members of those groups.[62] Once again, this facilitates more extensive (and therefore more privacy invasive) profiling.

The APPs permit the collection of information other than "sensitive information"[63] without consent, subject to some restrictions, including that it must be collected fairly and legally and for a purpose reasonably necessary for, or directly related to,

---

[59] [2017] FCAFC 4 (Austl.).

[60] *Id.* para 63.

[61] For a useful summary of the issues, see OFF. OF THE VICTORIAN INFO. COMMISSIONER, DE-IDENTIFICATION AND PRIVACY: CONSIDERATIONS FOR THE VICTORIAN PUBLIC SECTOR (June 2018), https://ovic.vic.gov.au/wp-content/uploads/2018/08/De-identification-Background-Paper-Update.pdf (Austl.).

[62] For a useful overview of the issue of group privacy, see Brent Mittelstadt, *From Individual to Group Privacy in Big Data Analytics*, 30 PHIL. & TECH. 475 (2017).

[63] *Privacy Act 1988* (Cth) s 6(1) (Austl.).

one or more the agency's functions or activities.[64] They restrict the use and disclosure of personal information with reference to requirements based on the purpose of collection,[65] subject to exceptions, including where "the use or disclosure of the information is required or authorised by or under an Australian law or a court/tribunal order."[66] The new BDA environment challenges the key assumptions that underlie these requirements—i.e., that data is collected for some specific purpose and it is reasonable to expect that its use will be confined to that purpose. This means that there is pressure for these principles to be overridden by other laws, as permitted by the exceptions for collection, use and disclosure that is "authorised under an Australian law."[67]

Significantly, the APPs in the Act lack several of the protections found in more modern data protection laws. For example, they do not impose any specific limitations on profiling or uses on automated technologies, of the type found in the GDPR[68] and in the Council of Europe's Modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data.[69] These European regimes are reflective of a view that data protection law can and should play a key role in mitigating the harms that can arise from new uses of personal data in BDA, including harms to rights other than privacy. That

---

[64] *See id.* sch 1, APP 3. The collection of sensitive information requires consent but is subject to several exceptions, including where "the collection of the information is required or authorised by or under an Australian law or a court/tribunal order." *Id.* sch 1, APP 3.4(a).

[65] In the case of sensitive information, use or disclosure for the related purpose is permitted only if that purpose is directly related to the primary purpose of collection. In the case of personal information that is not sensitive information, the other purpose must simply be related to the primary purpose (and therefore includes purposes that are indirectly related).

[66] See *id.* sch 1, APP 6.2(b).

[67] *See Health Legislation Amendment (Data-matching and Other Matters) Bill 2019* (Cth) sch 1 (Austl.) (amending the *National Health Act 1953* (Cth) (Austl.) to create authorisations for disclosure under the *Privacy Act 1988* (Cth) (Austl.)).

[68] *See* GDPR, *supra* note 25, arts. 13, 20, 21.

[69] *See* Comm. of Ministers, *Modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data*, 128th Sess., CM/Inf(2018)15-final (2018), https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016807c65bf [https://perma.cc/R7XX-VQL5]

approach makes sense given that the harms would not be possible without collecting and using personal information.

In the case of the GDPR, Article 13 requires that data subjects must be informed about the "existence of any automated decision making, including profiling"[70] and given "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing" and the consequences for them.[71] Article 21 supplements this protection by providing a right to object to automated processing where that processing is justified as being in the public interest or in the exercise of official authority vested in the controller.[72] Exercising this right has the consequence that a data controller can continue with processing only if they are able to demonstrate "compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject."[73] In addition, Article 22 contains a right not to be subject to automated decision making, "including profiling, which produces legal effects" (or similarly significant effects) concerning the data subject.[74] That right is subject to several exceptions, including where the profiling is authorised by a law to which the controller is subject.[75] However, a notable feature of this exception is that it requires that this law must lay down suitable measures to safeguard the data subject's "rights and freedoms and legitimate interests."[76]

Freedom of Information

The Australian *Freedom of Information Act 1982* (Cth) provides the key mechanism for ensuring government transparency, thereby giving effect to the right to receive information. It is relevant in three key respects: (1) it requires the publication by agencies of their "operational information,"[77] (2) it

---

[70] *See* GDPR, *supra* note 25, art. 13(2)(f).

[71] *See id.*

[72] *See id.* art. 21.

[73] *Id.* art. 21(1). If the controller is unable to demonstrate overriding legitimate grounds for the processing, the data subject has the right to require "the erasure of personal data concerning him or her without undue delay." *Id.* art. 17(1).

[74] *Id.* art. 22(1).

[75] *See id.* art. 22(2).

[76] *See id.* art. 22(3).

[77] *See Freedom of Information Act 1982* (Cth) pt II, s 8A (Austl.).

provides rights of access to "documents" in the possession of government agencies,[78] and (3) it provides a mechanism for individuals to access and amend their personal information.[79]

These features have an important role to play in ensuring the transparency of decision-making systems, but to work optimally in this respect they rely on the fact that the information necessary to shed appropriate light on these systems has been created and retained in a way that enables it to be retrieved without activating the provisions which permit non-disclosure on workload grounds. Any records created are subject to the retention requirements set by the Australian Archives under the *Archives Act 1983* (Cth), but currently there is no mechanism (either in that Act or elsewhere) which mandates either the creation of documents or how they are stored and filed.

A further issue is that the access requirements in the Act apply only with respect of documents, rather than information. This is potentially problematic in respect of AI-based decision-making systems as the information necessary to shed light on their operation will not necessarily be contained in any discrete document. While the Act contains a mechanism to require an agency to produce a written document of information that is stored electronically and not in a discrete written form, that mechanism is subject to two constraints. The agency must be able to do so using a "computer or other equipment that is ordinarily available" to it for retrieving or collating stored information.[80] In addition, an agency need not comply if "compliance would substantially and unreasonably divert the resources of the agency from its other operations."[81]

II. THE ISSUES REQUIRING CONSIDERATION

The tensions identified above suggest that there are two primary matters that require consideration—how can it be

[78] *See id.* pts III, IV.

[79] *See id.* pt V. Individuals are able to use the universal access provisions in Part III to access their personal information and the amendment provisions in Part V to request its amendment. There are parallel rights provided under the *Privacy Act 1988* (Cth) sch 1, APP 12 and 13.

[80] *Freedom of Information Act 1982* (Cth) s 17(1)(c)(ii),

[81] *Id.* s 17(2).

ensured that automated decision-making systems are designed and implemented consistently with existing legal requirements and to what extent do the legal requirements require modification, or supplementation by additional guidance materials, so as to address the ethical and policy issues raised by the delegation of human functions to machines.

Measures to Ensure that Implementation is Consistent with Regulatory Requirements

In relation to the first of these, it is important to identify and address the steps involved in the implementation of these new systems. Given the current opacity of algorithms, it is most logical to focus on what takes place before and after they begin to operate in practice.

The design or purchase phase provides an important opportunity to take steps to ensure that systems are implemented consistently with the rule of law (including the laws that frame the exercise of decision-making powers and privacy law which restrict what can be done with identifiable personal information) and do not produce outcomes that are discriminatory or lacking in demonstrable rationality. Guidance about these matters should ideally be included in the rules that regulate procurement activities and guide internal approvals of new technology projects. There may be scope here for the development of standards which can then play a role in defining what is required for good practice.[82]

It is also important to audit the functioning of systems in operation to ensure that they are working as intended and not producing any unfair or discriminatory effects. Given the potential scale of any potential adverse impacts, it is also important to check that the system is producing unpredictable or surprising

[82] In 2017, the International Electrotechnical Commission and the International Organization for Standardization became the first international standards development organizations (SDOs) to set up a joint committee (ISO/IEC JTC 1/SC 42) which will carry out standardization activities for artificial intelligence. *See* Antoinette Price, *First International Standards committee for entire AI ecosystem*, INT'L ELECTROTECHNICAL COMMISSION (Mar. 2018), https://iecetech.org/Technical-Committees/2018-03/First-International-Standards-committee-for-entire-AI-ecosystem [https://perma.cc/H9QR-P6Q9].

outcomes resulting from "the changeable, scalable nature of machine learning algorithms."[83] This ideally should take place first in the context of some form of piloting of the technology, as well as after it has formally been tested.

Identified Weaknesses and Gaps in Existing Legal Frameworks

The above analysis indicates that there are a number of weaknesses and gaps in existing frameworks that need to be addressed by way of legislative amendments.

In the case of administrative law, an important priority will be to amend the ADJR Act to ensure that it applies to decisions made by computers. It is also important that any laws which contain provisions permitting the making of decisions by computers provide clarification as to how those decisions are to be made and reviewed.

The identified weaknesses in anti-discrimination laws concern the tests used and the attributes protected. There is arguably scope for clarification of the current tests for direct and indirect discrimination to ensure that they are sufficient to capture the discrimination harms that can arise in the context of AI-based decision-making. It is also important to consider what additional attributes receive protection (and even whether it is desirable to enact a more general requirement of equal treatment).

In the case of the Australia's *Freedom of Information Act*, a key weakness lies in the fact that it is not underpinned by any requirements for specific documents to be created or for these documents to be created in such a way as to be readily retrievable. There would be logic in expanding the Archives Act to address these issues. There is also a strong case for amending the right of access conferred so that it applies to information (as is the case, for example, in the United Kingdom).[84]

The identified weaknesses in the Privacy Act are the most extensive. The issue, raised by Privacy *Commissioner v Telstra*

---

[83] FRENCH DATA PROTECTION AUTHORITY (CNIL), HOW CAN HUMANS KEEP THE UPPER HAND? THE ETHICAL MATTERS RAISED BY ALGORITHMS AND ARTIFICIAL INTELLIGENCE 50 (Dec. 2017).

[84] Freedom of Information Act 2000, c. 36, § 1 (UK) (conferring a right of access to information).

Corporation Limited,[85] is relatively simple to fix,[86] but there are less obvious solutions available for addressing the current use of identifiability as a touchstone for regulation. The issues that arise in relation to identifiability and group privacy are less easy to resolve, but there may be scope for adding a principle which imposes limits on the extent to which identifiable data can be combined from other sources for profiling purposes.

There is also a case for considering whether it is time to impose limitations on the extent to which other laws can override the collection, use and disclosure limitations. The approach used in Article 22 of the GDPR, which permits an override by another law only where it lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, arguably provides a useful model for consideration.

Finally, there is a strong case for augmenting the APPs with additional provisions along the lines of those found in Articles 13, 21 and 22 in the GDPR. While these are by no means perfect,[87] they go much further in addressing the problems raised by AI-based technologies than the Australian APPs.

CONCLUSION

It is no longer possible to ignore the fact that technology is increasingly altering the reality of how decisions are made by government agencies or the pressures for government agencies to do more with less resources. In these circumstances, it is important to consider what legal measures should be implemented to address the policy issues to which this gives rise. Arguably, the way forward lies both in the adoption of measures which enable such decisions to operate consistently with the safeguards in public law frameworks and in implementing legislative reforms to ensure that the applicable legal frameworks are fit for purpose in this new environment.

---

[85] [2017] FCAFC 4.

[86] For example, the equivalent term "personal data" in the GDPR is defined as "any information relating to an identified or identifiable natural person." *See* GDPR, *supra* note 25, art. 4(1).

[87] *See, e.g.*, Sandra Wachter, Brent Mittelstadt & Luciano Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 INT'L DATA PRIVACY L. 76 (2017).