Australian Federation of Disability Organisations

Email: office@afdo.org.au
Website: www.afdo.org.au
Phone: 1800 219 969

A.B.N. 25 105 510 898

Contact: Matthew Hall

T: +61 414 678 520
E: matthew.hall@afdo.org.au

17 January 2025

Use of automated decision-making (ADM) by government - public consultation

AFDO's responses to questions asked by the Attorney General's Department

Consultation paper issued November 2024. Responses to the questions the Department posed were lodged online at https://consultations.ag.gov.au/integrity/adm/consultation/ on 15 February 2025.

1. How should the need for transparency about the use of ADM be balanced with the need to protect sensitive information about business processes and systems?

As the CSIRO noted ("Artificial Intelligence: Australia's Ethics Framework" 2019, p.7) transparency is key to effective and appropriate use of ADM. Transparency measures are essential for accountability.

AFDO accepts the inner workings of some ADM technologies may not be easy to explain. However, it is necessary that sufficient information is provided that helps the public understand how any decision is made. AFDO also acknowledges that if proprietary software is used, the licensor of the software is likely to seek to maintain commercial secrecy or confidentiality in relation to human readable code and the "inner workings" of the software operates. As the CSIRO Framework notes

> "black boxes" in which the inner workings … are shrouded in secrecy are not acceptable when public interest is at stake. (CSIRO Framework, p. 7)

However, this position cannot and must not override the important public interest in:

- maintaining confidence in government processes and administrative decision making
- efficient, accurate, consistent, and interpretable decisions made according to law, and
- affording procedural fairness, and ensuring processes are not infected by jurisdictional error, bad faith, conflicts of duty and interest, or actual or perceived bias.

Transparency levels should be as high as possible and proportionate to the severity of adverse human rights impacts. If ADM is used in decision-making processes that carry high risks to human rights (and AFDO submits that it must not – see response to **Question 4** below) that ADM system must be subject to particularly high standards as regards the explainability of processes and outputs.

AFDO submits that ADM systems must only use:

(a) software the rights in which are owned by the Commonwealth, or

(b) open-source software.

If proprietary software is the only commercially available solution that meets the Commonwealth's requirements, the licensor must accept that the overriding public interest of transparency in ADM will necessarily moderate and reduce the extent to which the source code and software processes used by the Commonwealth will be subject to secrecy or confidentiality restrictions, and that they will enter the public domain to some extent, ranging from testing, regulation that requires transparency in the key priorities and fairness measures used in the system, through to measures enabling external review and monitoring. These options are dealt with further below, in response to **Question 2**.

Any software licensor that is not prepared to agree to these requirements, as a minimum will not be able to provide or offer to provide ADM software or services to the Commonwealth.

If the government is concerned this position will mean that ADM licensors will not be prepared to licence necessary software components, the government can rely on:

(a) the statutory licensing scheme concerning the use of copyright material for the Crown "for the services of the Commonwealth" in Division 2 of Part VII of the Copyright Act 1968 (see section 183 in particular)

(b) to the extent to which any software component may be protected by valid Australian patents, the similar scheme for acts done "for the services of the Commonwealth" in the Patents Act 1990 (see sections 163 and 165 in particular), or

(c) new statutory provisions for ADM that expressly provide government with immunity from civil and criminal proceedings, including infringement of copyright and breach of confidence (see, as an example, sections 90, 91 and 92, Freedom of Information Act 1982).

2. What transparency rules would be appropriate to build into the framework?

Like the arrangements under the Canadian government's Directive on Automated Decision-Making (p. 3), the Commonwealth must publicly release the source code and user documentation for ADM systems or software components owned by the Commonwealth, and for all open source ADM systems or software components.

In addition, for ADM systems or software components owned by the Commonwealth, and for all open-source ADM systems or software components, the Commonwealth must:

(a) disclose, the key priorities and fairness measures used in the ADM system, and any assumptions or processes that are dependent on or influenced by the characteristics or behaviour of the person inputting data as part of the decision-making process

(b) prior to any implementation:

  (i) test the ADM system, and

  (ii) publicly disclose testing processes and test results or other data, and

(c) following implementation:

  (i) regularly review, monitor, and assess the ADM system and the decisions the ADM system has made, and

  (ii) publicly disclose the monitoring and assessment processes and results or other relevant data.

AFDO recommends the Commonwealth adopts rules for proprietary software like those in the Canadian Directive (p. 3) and the CSIRO Framework (p. 7):

(a) the licensor must provide to the Commonwealth all released versions of any proprietary software components (and all associated documentation) used for ADM

(b) the licensor must disclose to the Commonwealth, and must allow the Commonwealth to further publicly disclose, the key priorities and fairness measures used in the ADM system, and any assumptions or processes that are dependent on or influenced by the characteristics or behaviour of the person inputting data as part of the decision-making process

(c) the Commonwealth must retain the right to:

(i) publicly disclose, without any confidentiality or secrecy restrictions, user documentation for the ADM system

(ii) access, review, monitor, assess and test (and authorise third parties to review, monitor, assess and test) the ADM system, and

(iii) publicly disclose testing processes and test results or other data. This must, include all released versions of proprietary software components.

Testing and assessment must include consideration of completeness, relevance, privacy, data protection, other human rights, unjustified discriminatory impacts and security breaches (see Council of Europe's Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems p. 8).

3. What pre-implementation safeguards should apply where ADM is intended to be used?

Quality of data and assessment of risks

Prior to implementation the agency must assess the provenance, quality and possible shortcomings of the data being put into and extracted from an ADM system, the possibility of its inappropriate or decontextualised use, and the environments within which the dataset will be or could possibly be used. The agency must determine the risks posed including potential bias or proxies for classifiers such as gender, race, disability, religion, political opinion or social origin. (Council of Europe Recommendation, p7).

Humans in the Loop (HITL)

The concept of 'humans in the loop' (HITL) was developed to ensure that humans maintain a supervisory role over automated technologies (See Rahwan I. 2018. Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology*, 20(1): 5-14).

In designing developing or modifying any ADM system the government must ensure that at every stage of the process there is human oversight. This oversight must provide critical assessment and monitoring of the inputs, algorithms and outputs, exception control, optimisation, and maintenance of automated decision systems. Human oversight must also ensure that errors are addressed, and humans remain accountable.

Further, when dealing with automated decisions that affect large and diverse groups of people, these processes must also adopt the 'society in the loop' design principle. (Rahwan). This means considering the behaviour expected from the technology, and how it aligns with the goals, norms and morals of all relevant stakeholders. This includes people with disability and their representative organisations.

This will ensure ADM systems reflect our laws, adhere to our human rights and social norms, protect our privacy, and take into account the issues, concerns and experiences of all of Australian society, especially those who are marginalised or most vulnerable.

The testing and disclosure requirements outlined in response to Question 2 above are also an essential and mandatory part of pre-implementation safeguards.

4. What system-level safeguards should be required, to ensure that ADM operates appropriately?

When is ADM appropriate to use

ADM is a very useful, highly effective, and entirely appropriate tool for government (and others) too use in circumstances where large volume of decisions have to be made, and those decisions are based on relatively uniform, and uncontested criteria. However, when discretion must (or may) be exercised aby the decision maker, unusual circumstances exist, or exceptions are required, ADM systems (if they are to be used at all) are best used only as a tool to assist human decision makers, and not to make **any** decisions.

We note Article 22 of the European Union's *General Data Protection Regulation* (EU, 2016/679) states individuals have the right

> not to be subject to a decision based solely on automated processing which produces legal effects concerning him or her (*sic*) or similarly significantly affects him or her (*sic*).

Whilst this is not the law of Australia, AFDO submits that the principle is sound and, from an ethical point of view, right. We submit this principle must be a bedrock principle of ADM in Australia, and specifically ADM undertaken by the Commonwealth.

informational self-determination

The design, development and ongoing deployment of ADM systems requires individuals to be informed in advance about the related data processing, including its purposes and possible outcomes (See our response to **Question 7** below). This also needs to give to individuals the ability to control their data, through the ADM process, including through interoperability. Deliberate efforts by individuals or groups to make themselves, their physical environment, or their activities illegible to automation or other forms of machine reading or manipulation, including through obfuscation, must be recognised as a valid exercise of informational self-determination, subject to possible restrictions necessary in a democratic society and provided for by law (Council of Europe Recommendations, p. 7).

Free from automation bias

Automation bias is

> the tendency to disregard or not search for contradictory information in light of a computer-generated solution that is accepted as correct (Parasuraman R, Riley V. 1997. Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2): 230-253).

Relying on automated decisions in situations where they cannot provide a consistently reliable outcome can result in increased errors by commission or omission (that is, by following, or failing to act on, advice from an ADM system in error). These issues are particularly important when using automated decisions in situations requiring discretion.

Good decision making by a person, based on advice or information from an ADM system requires the humans involved to exercise active thinking and to be involved actively in the process. They cannot remain passive, nor allow ADM systems to handle all the tasks of a decision. This is even more important if the ADM system is handling tasks for which the system is not suited (for example, the exercise of a discretion)

Who is responsible

Section 484J of the *Telecommunications Act 1997* authorises the Chair of the ACMA to

arrange for the use, under the **Chair's oversight**, of computer programs to take administrative action that may, or must, be taken by the ACMA under this Part. (our emphasis)

Section 484K (with the heading "Oversight and safeguards for automation of administrative action") goes on to provide as follows:

Chair to ensure administrative action is action that could be validly taken

(1) The Chair of the ACMA must take all reasonable steps to ensure that administrative action taken by the operation of a computer program under an arrangement under subsection 484J(1) is administrative action that the ACMA could validly take under this Part.

(2) Without limiting subsection (1), the Chair of the ACMA must do the things (if any) prescribed by the regulations for the purposes of this subsection.

…

Notice to entities of substituted decisions

(4) If, under subsection 484J(4), the ACMA makes a decision in substitution for a decision the ACMA is treated as having taken under subsection 484J(3), the ACMA must, within 14 days of the making of the substituted decision, give the entity the subject of the substituted decision written notice of the substituted decision.

Note: The notice must also include a statement about review rights (see section 557).

Publication

(5) If the Chair of the ACMA makes an arrangement under subsection 484J(1) in relation to particular provisions of this Part, the Chair must cause a statement to be published on the ACMA's website: (a) to the effect that the Chair has made such an arrangement; and (b) setting out those particular provisions.

Details in annual report

(6) The Chair of the ACMA, when preparing the ACMA's annual report under section 46 of the Public Governance, Performance and Accountability Act 2013 for a period, must include the following information in that report: (a) the total number of substituted decisions made by the ACMA under subsection 484J(4) of this Act in that period; (b) the kinds of substituted decisions so made; (c) the kinds of decisions taken by the operation of the computer program that the ACMA was satisfied were not correct.

These sections were inserted in the Act by the *Telecommunications Amendment (SMS Sender ID Register) Act 2024*, which is described in the consultation paper as recent legislation with "a more complete set of safeguards".

We do not consider these provisions to be sufficient to provide proper and adequate accountability for the ADM system or the decisions made by that system.

Accountability is more than oversight. It is also more than taking all reasonable steps to ensure that an automated decision made by an ADM system is administrative action that the agency could validly take. Accountability requires that a person assumes responsibility for the consequences of decisions made by ADM.

The legal framework for ADM must expressly and clearly provide

(a) the operator (in the Telecommunications Act example, the Chair) assumes the legal effects and bears the consequences of the ADM system and its decisions. This refers to the attribution of the legal effects and consequences to the operator as the decision maker

(b) the operator cannot be free of liability from complying with the ADM's decision or bearing the legal consequences arising from it solely on the grounds that the decision was made by automated means. The operator can also not deny that the decision can be attributed to it on the grounds that the ADM was developed by a third-party provider, or that data was collected from third-party data providers. The decision is not attributed to the programmer, the provider, the distributor or data providers.

(c) the operator is responsible for ensuring that the ADM is fit for its intended purpose and operates as it should.

Testing

Where possible. testing should be performed without using real personal data of individuals, and be guided through a diverse and representative stakeholder process, taking due account of the externalities of the proposed system on populations and their environments, before and after deployment.

Further, testing on the personal data of individuals must be performed with diverse and sufficiently representative sample populations. Relevant demographic groups should be neither over- nor under-represented. The staff involved in these activities must come from sufficiently diverse backgrounds to avoid deliberate or unintentional bias.

Furthermore, the development of algorithmic systems must be discontinued if testing or deployment involves the externalisation of risks or costs to specific individuals, groups, populations and their environments. Relevant legislative frameworks must disincentivise this externalisation.

Special care should be taken in relation to testing in live environments (Council for Europe Recommendations, p.8)

Ongoing review

Throughout the entire lifecycle of an ADM system, from the proposal stage through to the evaluation of effects, the human rights impact of individual systems and their interaction with other technologies must be assessed regularly. This is necessary due to the speed and scale at which these systems function and the fast-evolving technological environment in which they operate. This must also be done based on broad, effective consultations with those affected or likely to be affected.

This must also include adequate oversight by appropriately resourced independent institutions over the number and type of applications to contest decisions by affected individuals. Any results must not only lead to remedial action in specific cases but must also be fed into the systems themselves to avoid repetition of the results, make improvements. Where necessary, the ADM system or its or ongoing deployment must be discontinued, due to the likelihood of negative human rights impacts.

Information on the assessed disputes and resulting follow-up action must be documented regularly and made publicly available. (Council for Europe Recommendations, pp. 8 and 9)

5. What decision-level safeguards should there be for persons affected by decisions made using ADM (for example, review rights)?

When a decision is made by ADM, the decision, and the reasons for the decision, must be communicated to the affected individuals. This must provide a meaningful explanation of how and why the decision was made.

Notification of the decision must be in plain language, easily readable, and in a format accessible by the affected individuals.

Appendix C to the Canadian Directive provides useful examples of the requirements of the notification of the decision. Having regard to those examples, AFDO submit that in communicating the decision to affected individuals, the following safeguards must be followed:

In addition to any applicable legal requirement, the agency must ensure that a meaningful explanation is provided to the affected individuals. The explanation must:

(a) be in a format accessible to the affected individuals

(b) inform the affected individuals in plain, easily readable language of:

  (i) the role of the ADM system in the decision-making process

  (ii) the training and client data, their source, and method of collection, as applicable

  (iii) the criteria used to evaluate client data and the operations applied to process it

  (iv) the output produced by the system and any relevant information needed to interpret it in the context of the administrative decision; and

  (v) a justification of the administrative decision, including the principal factors that led to it, and

(c) inform affected individuals of relevant recourse options.

A general description of these elements must also be made available through an Algorithmic Impact Assessment and discoverable from an easily discoverable and accessible page of the agency's website.

We note the *Telecommunications Amendment (SMS Sender ID Register) Act 2024*, (see response to Question 4 above) does not contain provisions like those recommended. (See, for example section 484G(9) of the principal Act, inserted by this amending Act).

All affected individuals must also be afforded effective means to contest relevant determinations and decisions. This must include an opportunity to be heard, a thorough

review of the decision and the possibility to obtain a non-automated decision. This must include judicial and non-judicial procedures that guarantee an impartial review.

This right may not be waived.

The process must be accessible, affordable, independent, and easily enforceable before, during and after deployment of the ADM decision. This must include ensuring adequately trained staff are available to review the case competently and to take appropriate action effectively, and the provision of easily accessible contact points and hotlines. (Council of Europe Recommendation, p.9)

6. Should individuals be notified of the use of ADM?

Yes.

7. If yes, should notification be required at a specific point in the decision-making process, or should flexibility be provided to agencies about the appropriate time to make a notification?

"Yes" is in response to the first part of this question (notification be required at a specific point in the decision-making process). This is a very poorly worded question.

The existence, process, rationale, reasoning and possible outcome of ADM systems must be disclosed (see Council of Europe Recommendation, p. 8; and Canadian Directive, paragraphs 6.2.1 to 6.2.3):

(a) in general terms, at the collective level, continuously while the ADM system is operational. through all service delivery channels in use by the agency, including its website, and

(b) directly to any individuals concerned in a decision-making process:

(c) prior **to or immediately at the time of commencement of that process, and**

(i) at each time the agency communicates with the individual concerned about the decision-making process.

Any disclosure must be made in a manner and format that is in plain language, easily-readable and accessible.

Any disclosure in general terms must describe:

(a) how the components of the ADM system work

(b) how the ADM system supports the administrative decision

(c) results of any reviews or audits, and

(d) a description of the training data, or a link to the anonymized training data if this data is publicly available.

In addition:

(a) the free, specific, informed and unambiguous consent of participating individuals must be sought in advance, with an accessible means of withdrawing consent, and

(b) any decisions taken or aided by algorithmic systems are identifiable and traceable at the initial interaction, in a clear and accessible manner. (Council of Europe Recommendation, p. 8)

8. Should there be any exemptions to ADM safeguards? If yes, what exemptions should be included and why?

No, for the sound public policy reasons addressed in our response to **Question 1** above.

9. Should the safeguards be different depending on the risks associated with the use of ADM for a particular decision or administrative action?

As submitted above, the safeguard that must be implemented when the decision has legal or quasi-legal consequences, or when the decision involves the exercise of a discretion by the decision maker is that ADM must not be used or, not solely used, to make that decision.

In addition, potentially, pre-implementation safeguards could vary depending on the risks involved. This is the Canadian approach (see Appendix B - Impact Assessment Levels, and Appendix C - Impact Level Requirements to the Canadian Directive). We also refer to the Example Risk Assessment Framework for AI Systems in section 7.2 of the CSIRO Framework.

However, AFDO does not support any variations to the safeguards in relation to notification, consent (including withdrawal of consent), on-going assessment and testing, rights of appeal or contestability, accountability, automation bias, and informational self-determination.