Consultation on Guidance for Transparency in Automated Decision Making

Monday, 15 June 2026

Hakakire Legal Services | Abuja, Nigeria

Submitted in response to the Consultation on Guidance for Transparency in Automated Decision Making by the Office of the Australian Information Commissioner (OAIC) - 15th June 2026

JUNE 15, 2026.

Key Recommendations at a Glance

- We recommend that the OAIC distinguishes between computer programs that are directly relevant to Automated Decision Making.
- On the "Meaning of Substantially and Directly related to decision making", we rank the nature of output as our first priority followed by degree of reliance. While the ability and likelihood of human override over an ADM decision comes third or fourth.
- On the question of "What factors do you consider increase the likelihood that a decision could affect an individual's rights or interests?", we rank sensitive information as the first priority followed by financial outcomes, intrusive practices and lastly, vulnerable persons.
- We also recommend the nature of the decision; differential treatment; human oversight vs. sole automation; and transparency and redress as additional factors that increase the likelihood that a decision could affect an individual's rights or interests.
- On the question of "What classes of person do you consider to be vulnerable for the purposes of considering a decision's expected effect on a person's rights or interests, and why?", we recommend the following categories of people as vulnerable: Children, Low literacy individuals, Low income citizens, citizens with low or no digital literacy, ethnic minorities and migrants.

Introduction

The Privacy and Other Legislation Amendment Act 2024 introduced an automated decision-making (ADM) Obligation. From 10 December 2026, APP entities that use personal information in ADM with the potential to affect rights or interests will be required to provide information in their privacy policies about the kinds of personal information used and the kinds of decisions made using ADM.

The Office of the Australian Information Commissioner (OAIC) has confirmed that it intends to publish regulatory guidance on this obligation by September 2026 and, in anticipation, has released an issues paper seeking views to help inform the development of such guidance. The OAIC is seeking information and views to inform the development of the Guidance for the ADM Transparency Obligation. This includes views on scope.

Hakakire Legal Services (HLS) welcomes the opportunity to respond to OAIC's Issues Paper on developing Guidance for Transparency in Automated Decision Making (Issues Paper and ADM respectively). HLS is an Abuja-based legal consultancy firm with a practice focus in artificial intelligence governance, digital policy, international trade, and technology law, welcomes the release of these Guidelines and the invitation to contribute to their refinement.

This policy brief sets out our response to the consultation, identifies gaps and areas for strengthening, and offers concrete recommendations.

ADM and the ADM Obligation

Automated decision-making (ADM) is the decision making process of a technology or machine without or with limited human involvement based on factual data or digitally created profiles.

Under the Privacy and Other Legislation Amendment Act 2024 (Privacy Act), a transparency obligation was introduced from 10 December 2026, the types of ADM that require disclosure in APP entities' privacy policies under APP 1.7 are those that meet the following criteria:

1. the entity has arranged for a computer program to make, or do a thing that is substantially and directly related to making, a decision; and

2. the decision could reasonably be expected to significantly affect the rights or interests of an individual; and

3. personal information about the individual is used in the operation of the computer program to make the decision or do the thing that is substantially and directly related to making the decision.

According to AP 1.8, as it relates to these ADM obligations, relevant organisations will need to disclose

1. personal information used in ADM;
2. decisions made solely by ADM; and
3. decisions for which ADM does something substantially and directly related to making the decision.

What the Guidelines Get Right/ Our Commendation on the Guidelines

We commend OAIC for this timely document. While calling for views and opinions, the Issues paper considers the synergy between ADM obligations under the Privacy Act and other requirements including options in other frameworks. For example, anti-discrimination law, administrative law and the General Insurance Code of Practice. These legislative and other obligations are discussed in further detail below.

The issues paper presents 12 targeted questions across six themes (meaning of computer program; substantially and directly related to making a decision; meaning of significantly affecting rights or interests; meaning of making a decision; meaning of arranged for; extent of disclosure). These are relevant questions showing that the OAIC's thinking on ADM transparency is well thought out.

Our Recommendations

1. Meaning of "Computer Program"

The Issues paper states that the meaning of computer program is intended to be interpreted broadly. We note that this meaning is derived from Paragraph 336 of the Explanatory Memorandum states that "The term 'computer program' in APP 1.7(a) is intended to take its ordinary meaning and encompass a broad range of matters, including pre-programmed rule-based processes, artificial intelligence and machine learning processes to make a computer execute a task." We further note that the recognition that while 'computer program' is broad in scope, APP 1.7(a) narrows the types of computer programs, for example, to a thing that is "substantially and directly related to making a decision".

The broad definition of computer programme in this context risks creating uncertainty in the scope of the obligation. Adopting a broad statutory definition means that a broad range of tools and products used in ordinary business operations such as productivity software, workflow management tools, cybersecurity systems, data analytics platforms and customer relationship management systems will fall within this definition. Yet, they are not necessarily connected to the types of consequential decision-making processes that the ADM transparency obligation was intended to address. AI functionalities used in ordinary business operations such as those mentioned above are normally deployed for operational efficiency and user productivity rather than to make decisions that affect the legal rights and obligations of individuals.[1]

Recommendation: We recommend that the OAIC should distinguish between computer programs that are directly relevant to ADM. Such programs should trigger the transparency obligations. In such a scenario, the appropriate question should be whether the system is materially relied upon to make, or meaningfully shape, a decision that could reasonably be expected to significantly affect an individual's rights or interests.

2. Meaning of 'Substantially and Directly Related' to Making a Decision

The Issues Paper provided some relevant factors to enable an entity to assess whether a computer program substantially facilitates and is directly connected to a human decision-maker's decision.

- Degree of reliance on the ADM system output

- Ability and likelihood of human override over an ADM decision

[1] See also BSA Comments on Guidance for Transparency in Automated Decision Making https://www.bsa.org/files/policy-filings/06122026bsaauoaicadm.pdf accessed 15 June 2026.

- Nature of the output (advisory vs determinative)

- Transparency and explainability of outputs

- Integration of ADM into decision-making workflow

We rank these factors according to their degree of priority or importance in determining the extent to which they substantially facilitate and are directly connected to a human decision-maker's decision.

2.1 Degree of Reliance on the ADM System Output

The extent to which an organisation relies on the ADM system output is a critical aspect of what defines how substantially and directly connected to a human decision maker that system is. It is highly unlikely that, where the reliance on the ADM system is minimal, there would be any real effect on the people involved.

Reliance falls within a spectrum - It ranges from strict automation (where systems execute actions without human input) to decision support (where humans retain final authority), with a semi autonomous middle ground, where the algorithm executes a decision automatically unless a human intervenes to stop or modify it. Based on this spectrum, one can determine how much reliance is placed on a system and therefore the extent to which the ADM output is substantially and directly connected to a human decision maker's decision.

We note that properly balancing this reliance is a complex issue and that there are factors such as system explainability, task complexity and AI reliability which must be considered. However, while degree of reliance seems to be critical, we also wish to highlight empirical findings that having humans in the loop as the last line of defense against AI failures assumes that humans are capable of rectifying such AI failures when they occur.[2] Yet, the reality is that humans are not able to achieve appropriate reliance[3] on AI. Here, *appropriate reliance*[4] means that in order to complement the AI system, the human needs to adhere to AI recommendations if and only if

[2] Jakob Schöffer and others, 'AI Reliance and Decision Quality: Fundamentals, Interdependence, and the Effects of Interventions' (2025) JAIR 82 https://www.researchgate.net/publication/388690472_AI_Reliance_and_Decision_Quality_Fundamentals_Interdependence_and_the_Effects_of_Interventions

[3] Reliance is used as an umbrella term for humans' behavior of adhering to or overriding AI recommendations Lai, V., Chen, C., Smith-Renner, A., Liao, Q. V., & Tan, C. (2023). Towards a science of human-AI decision making: An overview of design space in empirical human-subject studies. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, pp. 1369–1385 Vivian Lai, 'Towards a Science of Human-AI Decision Making: A Survey of Empirical Studies' (21 Dec 2021) < https://arxiv.org/abs/2112.11471> accessed 15 June 2026.

[4] Raymond Fok and others, 'In Search of Verifiability: Explanations Rarely Enable Complementary Performance in AI-Advised Decision Making' (12 May 2023) https://arxiv.org/abs/2305.07722 accessed 15 June 2026

these recommendations are correct, and override them otherwise. What is achieved in most cases rather is over-or under-rely on AI recommendations, indicating a deficiency to distinguish correct from wrong AI advice.[5] Even the introduction of additional means of decision support (e.g., explanations) has rarely produced the expected benefits in terms of human-AI complementarity.

Recommendation: Based on the considerations above such as the likelihood that humans may not be able to rectify or achieve appropriate reliance, we are not sure the extent of the relevance and therefore are neutral on this point. We recommend that the degree of reliance on AI systems should be ranked No 3.

2.2 Nature of output (Advisory vs Determinative)

The nature of output, whether advisory or determinative matters to a large extent because outputs translate algorithmic analysis such as pattern recognition or predictive scoring, into actionable results that significantly impact individuals, businesses, or organisational workflows.

In automated decision making, the nature of the output can range from purely advisory suggestions to fully determinative, automatic actions. ADM can produce or deliver a myriad of outputs from a rating to a credit granting decision, from a movie recommendation to the allocation of work assignments, from an estimation of an insurance premium to a decision to remove illegal digital content. The variety of outputs is immense. Thus, the resultant legal consequences and the applicable legal regimes to ADM differ and are potentially multiple and varied.[6]

Purely advisory suggestions as a first category of possible outputs of ADM is where the system provides recommendations, risk scores, or prioritised data, but a human ultimately makes the final decision and executes the action.[7] More determinative outputs cover decisions adopted by the ADM related to the affected person. Known as Superseding (Human-out-the-loop), the decision is fully autonomous. The algorithmic output directly triggers a consequential action with no human intervention or prior review. In such cases, the person interacting with the ADM is directly affected by it. The ADM scores, grants, awards, settles a dispute, handles a complaint,

[5] Jakob Schöffer and others, 'AI Reliance and Decision Quality: Fundamentals, Interdependence, and the Effects of Interventions' (2025) JAIR 82 https://www.researchgate.net/publication/388690472_AI_Reliance_and_Decision_Quality_Fundamentals_Interdependence_and_the_Effects_of_Interventions.accessed 15 June 2025

[6] European Law Institute, 'Guiding Principles for Automated Decision-Making in the EU' (European Law, 2011)x<https://www.europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ELI_Guiding_Principles_for_Automated_Decision-Making_in_the_EU.pdf > accessed 15 June 2025

[7] For instance, a medical diagnostic tool suggesting a likely condition for a patient, which a doctor then reviews before writing a prescription. This often comprises ratings, rankings, predictions, recommendations, or content classification. The ADM lists, prioritises, classifies, rates, filters, ranks or recommends. The resulting output may be 'used' as an input to make subsequent decisions: what to buy, where to stay, which item to choose, whom to deal with. Thus, the output is not a decision which relates to the rights or the status of the person interacting with the ADM, or which has legal effects on them. Ibid.

removes digital content, assigns work, closes an account, dismisses an employee, refuses a request for credit, or decides whether an event notified by a user has to be compensated as per the insurance policy terms and conditions. The ADM produces legal effects concerning the affected person and/or significantly affects their rights, legal or contractual status, or interests.

This means that the ADM decision leads to the decision of whether the affected person is selected or rejected, scored, awarded, demoted, expelled from the platform, dismissed or is somehow directly affected by the decision.[8]

In certain cases, there can be a middle ground, where the system formulates a definitive decision or executes a preliminary action. However, a human monitors the process and holds the authority to override or reverse the outcome. This is known as Offloading (Human-over-the-loop).[9]

The extent to which these outputs affect each individual matters because the outputs determine the degree of effect on the individual - that is - how substantial and directly it is related to the decision. We note that a merely advisory output is likely not to have a substantial and direct effect.

Recommendation: Given that outputs determine the degree of effect the ADM decision has on the individual, we prioritise this categorisation over the degree of reliance because in reliance oftentimes, what is achieved in most cases rather is over- or under-reliance on AI recommendations, indicating a deficiency to distinguish correct from wrong AI advice.[10] Whereas the nature of outputs have a clear categorisation and affects the individual depending on this categorisation. The output therefore can be managed without complexity. We strongly agree that this is a relevant factor and we rank this as No 5.

2.3 Ability and Likelihood of Human Override over an ADM Decision

We rank this classification as third priority because simply including a human is unlikely to prevent systems from producing wrongful or harmful outcomes for individuals. This is frequently due to inadequate implementations or lack of control over the system. For human oversight to

[8] Example: A banking algorithm that automatically declines a credit card transaction due to suspected fraud, or a server that automatically rejects a loan application based on credit scoring. Ibid.
[9] Example: An algorithmic moderation system that automatically flags and hides a social media post, which a human moderator can reinstate upon appeal. Jakob Schöffer and others, 'AI Reliance and Decision Quality: Fundamentals, Interdependence, and the Effects of Interventions' (2025) JAIR 82https://www.researchgate.net/publication/388690472_AI_Reliance_and_Decision_Quality_Fundamentals_Interdependence_and_the_Effects_of_Interventions accessed 15 June 2026.
[10] Jakob Schöffer and others, 'AI Reliance and Decision Quality: Fundamentals, Interdependence, and the Effects of Interventions' (2025) JAIR 82 https://www.researchgate.net/publication/388690472_AI_Reliance_and_Decision_Quality_Fundamentals_Interdependence_and_the_Effects_of_Interventions accessed 15 June 2026

be meaningful and effective,[11] it must be carefully structured, taking into account both the limitations of ADM systems and the complex dynamics between human operators and machine-generated outputs. Given how challenging it is for human oversight to be meaningful and effective, we argue that the ability and likelihood of human override over an ADM decision, as a factor, is not as strong (connected to substantial or directly related) as factors such as reliance on the ADM decision or the nature of output (advisory or determinative).
Generally, individuals might not always be aware that they are subjected to ADM. This can create an imbalance of power between those affected by these systems and those who design, deploy, or control them. As ADM becomes increasingly integrated into processes, tools, and services, it is essential to ensure that these systems are not left to make autonomous, uncontrolled decisions that affect individuals' fundamental rights. Therefore, the involvement of humans as a safeguard against the risks associated with ADM systems (e.g., algorithmic bias and misclassification[12]) is increasingly perceived as necessary.

Integrating human judgment at various stages - during the design, real-time monitoring, or post-decision audits - can help ensure that ADM systems align with ethical standards, societal values, and regulations.

However, simply adding a human within the decision-making process does not inherently ensure better outcomes, nor should it serve as a means to deflect accountability for the system's Decisions.[13] In fact, just including a human is unlikely to prevent systems from producing wrongful or harmful outcomes for individuals. This is frequently due to inadequate implementations or lack of control over the system. Several flawed assumptions exist about human oversight of ADM. Both system deployers and affected individuals may erroneously assume that the system will invariably perform reliably, believing that any anomalies will inevitably be detected and addressed by human operators. Such assumptions can lead to unrealistic expectations and foster overreliance on the system.[14] As a result of this we argue that this particular criterion ranks as a 2 amongst the others.

Recommendation: Given the fact that just including a human is unlikely to prevent systems from producing wrongful or harmful outcomes for individuals due to inadequate implementations or lack of control over the system as well as the several flawed assumptions exist about human oversight of ADM we somewhat disagree and recommend that this factor be ranked as 2.

---

[11] European Union, 'Human Oversight of Automated Decision-Making' (Tech Dispatch, 2025) <https://www.aepd.es/documento/25-09-15-techdispatch-human-oversight-en.pdf> accessed 15 June 2026.
[12] ADM misclassification occurs when an ADM system assigns an incorrect label or category to input data - such as flagging a legitimate transaction as fraud. European Union, 'Human Oversight of Automated Decision-Making' (Tech Dispatch, 2025) <https://www.aepd.es/documento/25-09-15-techdispatch-human-oversight-en.pdf> accessed 15 June 2026.
[13] European Union, 'Human Oversight of Automated Decision-Making' (Tech Dispatch, 2025) < https://www.aepd.es/documento/25-09-15-techdispatch-human-oversight-en.pdf> accessed 15 June 2026.

[14] For some examples of these assumptions see European Union, 'Human Oversight of Automated Decision-Making' (Tech Dispatch, 2025) <https://www.aepd.es/documento/25-09-15-techdispatch-human-oversight-en.pdf> accessed 15 June 2026.

2.4 Transparency and Explainability of Outputs

Transparency in ADM systems often focuses on providing explanations of individual decisions, highlighting relevant factors or comparing the individual to the relevant population. However, this fails to address the deeper issues concerning ADM's implementation.[15] Transparency and explainability of inputs are a vital part of determining how substantially and directly related an ADM is to a decision. However, the problem with transparency is that as a means of providing explanations of individual decisions, it may fall into the "ripple effect trap", failing to acknowledge its impact upon the broader societal context in which it operates.[16] For example, explaining why someone was denied benefits may not reveal systemic biases in the data or the influence of private vendors. However, in the public sector, this narrow view risks overlooking how ADM reshapes accountability, fairness, and public trust at a structural level. Therefore, individual explanations will not de-mystify the broader ramifications of ADM's introduction upon existing social structures and nor can the mathematical formalism that underpins ADM accurately capture contextual, often contested, social concepts like fairness.[17]

This background undermines claims that transparency can ensure meaningful accountability and can respond to potential issues. As a result, current approaches to transparency may not appropriately foster accountability nor address the full spectrum of potential negative outcomes.

Recommendation: In light of our explanation above, we simply agree that this is a very relevant factor and recommend that this factor be ranked as 4.

Fictional Edge Case Study: Generative AI with human oversight - Substantially and Directly Related to Making a Decision

The National Aged-Care Insurance Agency (NACIA) makes decisions about eligibility for elderly people trying to access life-changing aged-care support. To make eligibility assessments quicker, NACIA staff use a generative AI chatbot, GPTea, to summarise candidate profiles and receive recommendations about eligibility. NACIA's human staff always have oversight and make the final decision about eligibility based on the recommendation provided by the chatbot.

Our assessment of the case study is that we would consider NACIA's described uses of GPTea to be substantially and directly related to making decisions.

---

[15] Hannah Louise Smith, 'If transparency is the solution, are we really addressing the problems of automated decision-making?' Internet Policy Review (24 July 2025) https://policyreview.info/articles/news/transparency-automated-decision-making accessed 15 June 2026

[16] McKane Andrus and others, 'AI Development for the Public Interest: From Abstraction Traps to Sociotechnical Risks' (4 February 2021) <https://arxiv.org/pdf/2102.04255> accessed 15 June 2026.

[17] See Selbst, Andrew D. and Boyd, Danah and Friedler, Sorelle A. and Venkatasubramanian, Suresh and Vertesi, Janet, Fairness and Abstraction in Sociotechnical Systems (August 23, 2018). 2019 ACM Conference on Fairness, Accountability, and Transparency (FAT*), 59-68, Available at SSRN: https://ssrn.com/abstract=3265913; see also Hannah Louise Smith, 'If transparency is the solution, are we really addressing the problems of automated decision-making?' Internet Policy Review (24 July 2025)https://policyreview.info/articles/news/transparency-automated-decision-making.

Reason

APP 1.7 states that privacy policies must contain certain information if the entity has arranged for a computer program to make, or do a thing that is substantially and directly related to making a decision. In defining substantially and directly, paragraph 337 of the Explanatory Memorandum states "substantially means where it is a key factor in facilitating the human's decision making; and directly means where the thing has a direct connection with making the decision."

In the scenario above, the first inquiry is whether the recommendations of the AI Chatbot GPTea is a key factor in the final decision taken by NACIA staff in deciding who is eligible to access life changing aged care support. Our answer is yes because despite having human oversight, the final decision is taken based on the recommendation provided by the chatbot. This shows that there is a significant degree of reliance by the staff on the ADM system output.

We also note that the word directly which means where the thing has a direct connection with making the decision is applicable in this context. The Chatbot GPTEA by summarising the candidates' profile and making recommendations, which the staff relies on, to make their final decision, should be considered a program that is doing a thing directly connected with making the decision.

Factors that Increase the Likelihood that a Decision could Affect an Individual's Rights or Interests

- Sensitive information - We strongly agree, ranking this as a 5, that this is a decision that could affect an individual's rights or interests because sensitive information gives insight into peculiar characteristics of a person that is highly likely to create unintended bias towards them could lead to harm and discriminatory practices in areas such as hiring, law enforcement, and healthcare. This is a main reason why the information in the first place is sensitive.
- Vulnerable persons - We agree, ranking this as 4, as this is a decision that could affect an individual's rights or interests because vulnerable persons are those at heightened risk of harm, exploitation, or discrimination due to age, disability, or cognitive/socioeconomic factors.
- Intrusive practices - We are neutral on this factor and therefore rank it as a 3.
- Financial outcomes - We agree, and rank this as a 4 because financial outcomes based applications are highly competitive and heavily in demand.

Recommendation: Additional Factors for the OAIC to Consider

- Nature of the Decision: the nature of the decision can increase the likelihood that a decision can affect an individual's right and interest. This is the case if the system determines
- Differential Treatment: Algorithms that drive targeted advertising or profiling can significantly affect interests if they result in restricted access to goods, employment opportunities, or differential pricing
- Human Oversight vs. Sole Automation: The likelihood of adverse effects increases exponentially with solely automated decision-making (decisions made by machines without human involvement). If an AI tool is purely advisory and a human meaningfully reviews the outcome, the risk is lower.
- Transparency and Redress: Opaque decision-making processes inherently increase risk, as they prevent individuals from understanding how a decision was reached and make it harder to contest unfair or biased outcomes.

Classes of persons we consider to be vulnerable for the purposes of considering a decision's expected effect on a person's rights or interests, and reasons.

The digital technology employed to digitise government services, guide or make automated decisions aims to optimise administration. Therefore, it treats all citizens as statistical data or categories: "average citizens" who have a reasonable understanding of the law, have the ability to fill in forms, can understand how government websites work, and can thus use them. These systems simplify the law and the language that is used but they are not overly patronising. "Average citizens" can engage independently with technology and require little assistance.The "average citizen" for the technology industry developing software does not necessarily match the statistical reality, especially in racially, ethnically, and income diverse countries like the United States. In 2018, the Washington Post ran a profile of the "average citizen" based on Bureau data and recent polling and concluded that the average American was a white, woman in her fifties, with a Bachelor's degree, average income, who lives in her own house with her spouse.[18]

- Children: Under regulations like the European Union's General Data Protection Regulation (GDPR), children are recognized as vulnerable persons. The GDPR restricts fully automated decisions (without human intervention) that produce legal or significant effects on the child, and mandates explicit consent mechanisms for data processing.
  There are also human rights impacts. The UNICEF Office of Research – Innocenti actively highlights how digital ADM can interact with a child's fundamental rights, including their right to privacy, education, and protection from exploitation.

[18] Philip Bump, This is What the Average American Looks Like in 2018, Washington Post (Aug. 13, 2018), https://www.washingtonpost.com/news/politics/wp/2018/08/13/this-is-what-theaverage-american-looks-like-in-2018/. accessed 15 June 2026.

- Low literacy individuals:individuals with low literacy cannot understand how new digital processes work, cannot translate their needs in the most accurate way, consume news that may not be precise, and cannot convey their views or requests accurately in writing. Because of their reduced digital skills, limited literacy, financial or ethnic background or mentality, these citizens are not treated equally and experience that the digitalization of public services limits their ability to exercise their rights before public bodies.
- Low income citizens
- low/no digital literacy
- Ethnicity - ethnic minorities: existing scholarship shows that ethnic minorities are discriminated by algorithmic decision-making systems that score them negatively, suggest they will commit more often crimes and welfare fraud, and surveil them more closely than other segments of the population.[19]
- Education
- Migrants

Conclusion

We strongly commend the OAIC for the opportunity to provide our responses and recommendations in response to the Issues Paper. We hope our comments are useful as the OAIC develops Guidance on the ADM transparency obligations under APP 1.7–1.9. supporting responsible and trustworthy uses of ADM and AI technologies. We welcome the opportunity to discuss these recommendations further and to continue to engage with the OAIC as the Guidance is further developed.

Yours sincerely,

[redacted] Hakakire Legal Services

[19] See Sofia Ranchordás & Ymre Schuurmans, Outsourcing the Welfare State: The Role of Private Actors in Welfare Fraud Investigations, 7 European Journal of Comparative Law and Governance 5, 31-33 (2020).