Western Australia AI Hub®

Submission to the Office of the Australian Information Commissioner

Automated Decision-Making Transparency Obligation (APP 1)

Submitted by: Western Australian AI Hub Management Committee

Management Committee: [redacted]

Addressee: Office of the Australian Information Commissioner (OAIC)

Subject: Consultation on guidance for transparency in automated decision-making under Australian Privacy Principle 1

Publication preference: The Western Australian AI Hub is content for this submission to be published.

1. About the Western Australian AI Hub

The Western Australian AI Hub supports responsible, practical and humanity-centred integration of artificial intelligence and automated decision-making across Australia. Our perspective combines applied technology, privacy and data governance, responsible innovation, humanity-centred design, organisational assurance, and the needs of industry, government, research and community stakeholders.

Western Australia's unique landscape should be a key factor in shaping national guidance. With our vast geography, the complexity of delivering services to regional and remote areas, and a diverse economic base, from mining and energy to health and education, our context is distinct. Importantly, for the communities we serve, things like accessibility, cultural safety, and digital inclusion aren't just optional add-ons; they are essential to achieving real transparency in privacy.

Our work is grounded in humanity-centred design principles aligned with Australia's Human-Centred AI Ethics Framework (HCAIEF) and informed by international standards including ISO/IEC 42001 (AI management systems) and the NIST AI Risk Management Framework. We are committed to ensuring that guidance on ADM transparency reflects the needs of First Nations communities, consistent with the Maiam Nayri Wingara Indigenous Data Sovereignty Principles and the CARE Principles for Indigenous Data Governance, and addresses the particular challenges of deploying automated systems in regional, remote and

culturally diverse settings.

2. Executive position

We, at the WA AI Hub, want the OAIC's guidance on automated decision-making (ADM) to be practical and truly meaningful for the people it impacts. Guidance should not reduce the obligation to a generic statement that an Australian Privacy Principle (APP) entity "uses AI" or "uses automated decision-making". It should require decision-centred, layered and plain language transparency about when personal information is used in substantially automated decisions that have a legal or similarly significant effect on individuals.

Our central recommendation is that the OAIC adopt a **substance-over-form** approach. The question should not be whether a human appears somewhere in the process, but whether automation materially shapes, determines, gates, ranks, scores, recommends or otherwise influences the outcome. Human involvement should only take a decision outside the core concern where that involvement is timely, informed, independent, authorised and capable of changing the outcome.

The new obligation should help people answer **four** practical questions:

1. Is a decision about me being made or materially influenced by automation?
2. What personal information, including inferred or generated information, is used?
3. What kinds of effects could the decision have on me?
4. How can I seek clarification, correction, review or human reconsideration?

The guidance should also support APP entities by providing templates, decision trees, sector examples, and model privacy policy clauses. This is especially important for small and medium entities and for organisations procuring third-party AI or ADM systems.

3. Summary of recommendations

No.	Recommendation
1	Define "substantially automated" by the materiality of automation in the decision process, not by whether a human formally approves the final decision.
2	Clarify that meaningful human involvement requires authority, competence, time, evidence, independence and a real ability to change the outcome.
3	Confirm that personal information includes inferred, derived, generated, scored or predicted information about an identifiable individual where it is used in the decision

No.	Recommendation
	process.
4	Interpret "legal or similarly significant effect" broadly, contextually and by reference to practical consequences for the individual, not merely by formal legal status.
5	Require layered transparency: a public privacy policy overview, decision-specific notices where practicable, and more detailed information on request or review.
6	Provide an ADM disclosure table template for privacy policies, including decision type, purpose, personal information inputs, role of automation, human oversight, likely effects and review pathway.
7	Do not require source code or disclosure that compromises security, but do require intelligible information about the main factors, data sources, system limitations and safeguards.
8	Include specific examples covering employment, credit, insurance, housing, education, health, social services, fraud detection, risk scoring, triage, eligibility, pricing and access to essential services.
9	Make clear that vendor opacity is not a defence. APP entities remain accountable for transparency about ADM systems they deploy or procure.
10	Align the guidance with privacy impact assessment practice, data quality obligations, security obligations, recordkeeping, contestability, review and procurement governance.
11	Address vulnerable and disproportionately affected groups, including children, Aboriginal and Torres Strait Islander peoples, people with disability, culturally and linguistically diverse communities, and people in regional and remote Australia.
12	Provide implementation artefacts: a decision tree, glossary, sample clauses, ADM register template, risk-based examples and a readiness checklist.

4. Responses to consultation questions

Question 1: Substantially and directly related to making a decision

Answer:

We support a practical, risk-based and decision-centred approach, provided the guidance sets a clear minimum standard for meaningful transparency.

The decisive factor is whether the human decision-maker retains meaningful, evaluative control over the outcome, or merely ratifies the automated output. Where the human genuinely evaluates and can readily depart from the system's result, the connection to ADM is weaker; where the output is accepted as the default, the decision is properly treated as substantially automated. In order of importance, the factors that determine this connection are: (1) the degree of meaningful human override and evaluative control; (2) the degree of reliance on the system output; (3) the level of automation and integration into a high-throughput workflow; (4) the explainability of the output to the affected individual; and (5) the advisory versus determinative nature of the output.

This ordering is reflected in the detailed ranking at point C in response to Question 1 below.

A. The OAIC should frame the APP 1 obligation as part of open and transparent management of personal information, rather than as a narrow AI disclosure exercise. The purpose is not merely to tell individuals that automation exists. The purpose is to help individuals understand when personal information about them is being used in a decision process that may materially affect their rights, opportunities, access to services, safety, finances, reputation or life chances.

   The guidance should therefore distinguish between:

   - **Organisational transparency**: an entity's privacy policy explains categories of ADM use across the organisation.
   - **Decision transparency**: an individual can understand that a particular class of decision may be automated or substantially automated.
   - **Actionable transparency**: an individual knows how to seek correction, explanation, review or escalation to a human decision-maker.

B. We recommend that OAIC guidance expressly discourage generic statements such as "we may use automated tools" unless accompanied by decision categories, personal information inputs, purposes, likely effects and review pathways.

C. Ranking of factors relevant to assessing whether a computer program substantially facilitates and is directly connected to a human decision-maker's decision (1 = least

important, 5 = most important):

- **Degree of reliance on the ADM system output: 5**. The materiality of the automated output in shaping the final decision is the most critical factor. Where an entity relies on the system output as the primary or default basis for a decision, the obligation should clearly apply.
- **Ability and likelihood of human override over an ADM decision: 5.** This factor is equally critical. A system where human override is technically possible but practically unlikely, discouraged or unsupported does not represent meaningful human involvement.
- **Nature of the output (advisory vs determinative): 4**. Whether the output is a recommendation, a binding determination or a gatekeeping filter matters, but the practical effect on the individual is more important than the formal label.
- **Transparency and explainability of outputs: 4.** The capacity of the system to produce intelligible explanations is important for assessing the quality of any human review, but transparency is primarily an obligation that follows from the threshold being met, not a factor that determines the threshold itself.
- **Integration of ADM into decision-making workflow: 4**. How deeply embedded the automated component is in the end-to-end process affects whether the human reviewer can meaningfully depart from the system output. Workflow design often determines whether human oversight is real or nominal.
- **Consequences for the individual: 5**. The WA AI Hub considers the severity and reversibility of the decision's consequences for the affected individual to be a factor of equal importance. A decision that affects access to employment, housing, credit, health services, education, government benefits or essential services should be treated as more likely to require transparency, regardless of the technical architecture of the system.

Question 2: Substantially and directly related to making a decision

Answer:

Yes. NACIA's use of GPTea is substantially and directly related to making decisions about aged-care eligibility, and it engages the APP 1 transparency obligation.

Free text comments

"Substantially and directly related" should be interpreted by reference to the role automation plays in shaping the outcome, not by the formal label applied to the tool, the nominal presence of a human reviewer, or the organisation's characterisation of the system as merely "advisory."

On the facts of this scenario, GPTea summarises candidate profiles and generates eligibility recommendations that NACIA staff rely upon when making final decisions. The human role is downstream of the automated output and is exercised within the constraints it creates. That is sufficient to engage the transparency obligation. The fact that staff "always have oversight" and "make the final decision" does not, by itself, take the system outside APP 1: oversight exercised without genuine authority, competence, time, or independence to depart from the automated recommendation is procedural, not substantive.

A decision should be treated as substantially automated where an automated system materially determines, recommends, ranks, scores, classifies, predicts, prioritises, triages, filters, flags, or gates an outcome, and the human role is limited, constrained, or structurally unlikely to alter the result.

The OAIC guidance should provide worked examples confirming that automation is substantial even where a human is formally involved. These include:

- A staff member accepts a system recommendation by default unless exceptional circumstances are expressly recorded.
- A risk score determines eligibility for manual review, or the intensity of scrutiny applied to an individual.
- A system ranks applicants and only those above a threshold are presented to a human reviewer.
- A fraud or compliance model triggers suspension, delay, investigation, or adverse treatment without independent human initiation.
- A model-generated classification determines price, access, service level, or priority queue position.
- A generative AI tool summarises evidence and recommends an outcome in a high-volume, time-pressured decision environment — as GPTea does here.

Conversely, automation is less likely to be substantial where a tool is genuinely administrative or supportive,for example, where it collates information, checks arithmetic, schedules a workflow, or presents options without weighting, ranking, or influencing the substantive outcome.

The real test is not what a tool is called, but what it actually does to the decision. The question is whether the human reviewer has, in practice, the time, authority, information, and institutional freedom to depart from the automated output — or whether they are functioning as a ratification step rather than a decision-maker.

In the NACIA scenario, elderly individuals seeking access to life-changing aged-care support are among the most vulnerable persons under any reasonable definition. The combination of a

vulnerable cohort, a life-defining outcome, and a generative AI system operating at volume in a time-pressured environment makes this an unambiguous case where the transparency obligation applies. Describing the human role as "oversight" does not change the underlying reality: GPTea shapes what staff see, how they frame the decision, and what outcome they are primed to reach. That influence is both substantial and direct.

The WA AI Hub recommends that the OAIC guidance make clear that **the label "advisory" cannot, by itself, exclude a system from the transparency obligation.** Where a recommendation is generated by an AI system and presented to a human reviewer operating under volume, time, or organisational pressure, the advisory framing is likely to describe the form of the output rather than the practical reality of its influence on the decision.

Question 3: Meaning of significantly affect rights or interests

Answer:

The WA AI Hub ranks the following factors by the degree to which they amplify impact on an individual's rights or interests:

1. **Use of sensitive information** (health, biometrics, protected characteristics such as race, disability, or religion)
   5 - Strongly agree. Sensitive information carries an inherent power asymmetry: the individual rarely controls how it is collected, retained, or inferred. Its use in ADM creates compounded risk because errors or biases in sensitive data are both harder to detect and more damaging to correct.
2. **Impact on vulnerable persons** (children, First Nations peoples, the elderly, people in crisis, people with disability)
   5 - Strongly agree. Vulnerability amplifies harm because affected individuals are less likely to have the resources, literacy, or institutional access to understand, challenge, or recover from an adverse automated decision. The OAIC guidance should recognise intersecting vulnerability, i.e. a person may belong to multiple vulnerable groups simultaneously, which compounds risk.
3. **Intrusive practices** (continuous surveillance, persistent behavioural profiling, real-time monitoring in sensitive contexts)
   5 - Strongly agree. The ongoing and aggregative nature of intrusive data collection means that its effect on rights or interests is not a single event but a sustained encroachment. Longitudinal profiling can alter a person's opportunities over time in ways that are invisible to them.
4. **Financial or life-defining outcomes** (denial of credit, housing, essential employment, welfare entitlements, aged-care eligibility, immigration status)
   5 - Strongly agree. These decisions determine a person's material security and future

trajectory. An error in an ADM outcome here is not a minor inconvenience — it can be practically irreversible without significant legal or financial resources.

5. **Cumulative or systemic effect** (automated patterns of exclusion that entrench structural inequality across a cohort or community)
   5 - Strongly agree.
   Individual impact assessments can miss population-level harm. The OAIC guidance should explicitly account for ADM systems that may be low-impact on any single individual yet cause systemic discrimination at scale — this is particularly important for First Nations communities and other historically marginalised groups.
6. **Additional factor: Opacity and irreversibility of the decision pathway**
   5 - Strongly agree.
   Where an individual cannot access a meaningful explanation of why a decision was reached and has no practical avenue to challenge it, the severity of impact on their rights is independently elevated, regardless of the substantive outcome. Opacity prevents remedy and must itself be treated as a rights-affecting factor.

An ADM system that combines sensitive information about a vulnerable individual with a life-defining outcome will almost always significantly affect that person's rights or interests and warrants the highest level of transparency under APP 1. Critically, the presence of high-impact factors cannot be neutralised by nominal human involvement in the decision process.

The mere presence of a human reviewer does not, by itself, establish that a person has "contributed to" a decision in any meaningful sense. APP entities may otherwise inadvertently create what we term **"human-in-the-loop theatre"** where a human is formally present but the automated output effectively determines the matter. This would hollow out the purpose of the transparency obligation.

The OAIC guidance should make clear that nominal human involvement is insufficient. The WA AI Hub recommends that the OAIC define **meaningful human involvement** through six criteria:

Criterion	Description
Authority	The person has genuine authority to change, reject, or override the automated output, not merely to endorse it.
Competence	The person has sufficient training, domain expertise, and understanding of the system and its limitations.
Time	The person has adequate time to review the individual case, rather than processing a high-volume queue under productivity pressure.
Information	The person can access relevant evidence, uncertainty indicators, and known system limitations, not only the automated recommendation.

Independence	The person is free from operational, cultural, or performance-management pressure to accept the system output by default.
Accountability	In significant cases, the person records their reasons whether they accept, vary, or reject the automated output.

These criteria reflect a substantive, not merely procedural, standard of human participation. Their absence should be treated as a factor that increases not decreases the likelihood that an ADM decision significantly affects rights or interests, because the individual is, in effect, subject to a fully automated outcome without safeguard.

The WA AI Hub also recommends that the OAIC periodically publish sector-specific guidance on which combinations of the above factors are known to produce high-impact ADM outcomes, for example, in aged-care, employment screening, and child protection to help entities self-assess their obligations proactively rather than reactively.

Question 4: Meaning of significantly affect rights or interests

Answer:

Classes of person warranting heightened protection under an ADM transparency framework include: children and young people; Aboriginal and Torres Strait Islander peoples; people with disability; older people; culturally and linguistically diverse communities; people experiencing family violence, financial hardship or homelessness; and people in regional, remote and very remote communities. These groups are named as prompts for closer scrutiny, not as an exhaustive list, and not as a safe harbour for entities whose systems affect people outside it.

Vulnerability is contextual, not categorical

The OAIC should not frame vulnerability as membership of a fixed class of people. Anyone can be made vulnerable to an automated decision where they face a power or information asymmetry, have limited ability to detect or contest an error, are under time or financial pressure, or have no practical alternative to the service in question. A static list risks two errors simultaneously: missing people made vulnerable by their circumstances, and stigmatising groups by implying they are inherently less capable.

The test the OAIC guidance should require entities to apply is: who could be materially disadvantaged by this system, in this context, and why? That question should be asked before deployment, at regular review intervals, and whenever the system or the population it affects changes materially.

Groups warranting heightened attention
While the test is contextual, experience shows that certain groups are more frequently and more severely affected by ADM errors and should be specifically named as triggers for proportionate review:

- Children and young people, including those in contact with child protection systems
- Aboriginal and Torres Strait Islander peoples
- People with disability, including cognitive or communication disability
- Older people, particularly those accessing aged care, health, or financial services
- Culturally and linguistically diverse communities, including people with limited English proficiency
- People experiencing family violence, financial hardship, housing instability, or homelessness
- People in regional, remote, and very remote communities

Naming these groups creates an obligation to examine impact more carefully. It does not create a presumption that everyone outside the list is unaffected.

Western Australian and remote context
Western Australia's geography makes this question particularly acute. A decision that appears low-impact in a metropolitan setting can be severe in a remote community where alternative providers are absent, connectivity is unreliable, digital literacy varies, and service delays compound across multiple systems simultaneously. The OAIC guidance should require entities to assess practical impact in the worst-served settings they operate in, not the average case, and not the metropolitan default.

For decisions affecting Aboriginal and Torres Strait Islander peoples, impact assessment should be consistent with Indigenous data sovereignty, including the Maiam Nayri Wingara Data Sovereignty Principles and the CARE Principles for Indigenous Data Governance. Entities should not treat community-level data as a resource to be used without community involvement in how it is applied.

Children and young people
Privacy notices alone are not adequate for this cohort. Entities whose ADM systems affect children should provide age-appropriate, layered explanations of how decisions are made; support involvement of a parent, guardian, or advocate without extinguishing the young person's own agency; and apply a best-interests-of-the-child standard consistent with the emerging Children's Online Privacy Code. Where defaults must be set, the most protective option should be the default, not the most commercially convenient one.

Proxy information and group-based data
ADM systems frequently use information the individual did not directly provide: postcode, device type, purchasing behaviour, health service usage, employer, school, or demographic groupings that function as proxies for protected characteristics. Where such variables materially influence a decision about an identified or reasonably identifiable individual, they should be treated as personal information for transparency purposes regardless of whether they are individually identifying in isolation. Entities should be required to consider whether proxy variables, inferred attributes, or segmentation techniques are functioning as personal information in context and to disclose that use.

Amplifying factors
To make vulnerability assessment practical, the guidance should direct entities to weigh the following in order of their capacity to increase harm:

1. **Use of sensitive information:** health data, biometrics, or protected characteristics such as race, disability, or religion
2. Impact on any of the groups identified above
3. **Intrusive practices:** continuous monitoring, persistent behavioural profiling, or real-time surveillance in sensitive contexts
4. **Financial or life-defining outcomes:** denial of credit, housing, employment, welfare, aged-care eligibility, or immigration status
5. **Cumulative or systemic effects:** automated patterns of exclusion that entrench existing inequality across a cohort or community over time

Future-proofing
As automated systems increasingly act with limited human oversight, disproportionate-impact assessment must be continuous rather than a one-time exercise at deployment. This requires monitoring for differential error rates across groups, equity-focused testing before and after deployment, and an accessible, non-automated pathway to human review for any individual who believes a system has affected them unfairly. The objective is not only to protect today's identified groups. It is to build systems that remain fair and contestable for people whose vulnerability cannot yet be foreseen.

Question 5: Meaning of significantly affect rights or interests

Answer:

The phrase "significantly affect rights or interests" should be defined to include both formal legal effects and practical consequences that materially alter a person's access to, or treatment in relation to, important opportunities, services or protections. It should be interpreted broadly and contextually, with attention to real-world outcomes for the

individual rather than narrowly confined to changes in legal status.

A legal effect includes decisions that determine, alter or affect legal rights, obligations, entitlements, liabilities or status. A similarly significant effect should include decisions that may materially affect an individual's access to, or treatment in relation to, important opportunities, services or protections.

Examples should include decisions concerning:

- employment, recruitment, rostering, performance management or termination;
- credit, lending, insurance, pricing, eligibility or risk classification;
- housing, tenancy screening or accommodation access;
- education admissions, assessment, progression, scholarship or misconduct processes;
- health, disability, aged care, triage, prioritisation or service allocation;
- government benefits, permits, compliance, enforcement or service access;
- identity verification, fraud detection, account suspension or de-platforming;
- children's services and decisions affecting vulnerable individuals or communities.

The effect should not need to be adverse in every case. A decision class should fall within the guidance where it is reasonably capable of producing a significant effect for some individuals, including through denial, delay, additional scrutiny, materially different pricing, reduced service quality, reputational harm or loss of opportunity.

Where the OAIC wishes to prioritise its examples and oversight, the WA AI Hub recommends assessing each service and support against the question "automated handling in this sector significantly affects an individual's rights or interests," rated on a five-point agreement scale (1 = Strongly disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly agree).

The ratings below are derived from a contextual weighting of five best-practice factors drawn from this submission:

1. impact on fundamental rights
2. the essential rather than discretionary nature of the service
3. potential for systemic bias or discrimination
4. the economic cost to the consumer
5. the individual's practical ability to detect and challenge the decision.

On that basis:

1. **Education and child protection - 5 (Strongly agree)**, as automated handling directly shapes a person's life trajectory and safety and affects individuals with the least capacity to detect or challenge it
2. **Healthcare, aged care and disability support - 5 (Strongly agree)**, as it affects fundamental wellbeing, dignity and often life-critical outcomes for highly vulnerable

individuals

3. **Banking, credit and finance - 4 (Agree)**, as it determines financial security and independence and carries high systemic-bias risk, while some recourse mechanisms exist
4. **Essential utilities and telecommunications - 4 (Agree)**, as it enables basic participation in modern society, with impact escalating where regional or remote individuals have limited alternatives
5. **General retail and physical assistance - 3 (Neutral)**, as it concerns largely discretionary commerce yet, consistent with this submission, is not excluded and can still produce a significant effect for a particular individual. This weighting is intended to guide proportionate attention rather than to exclude any sector: a lower-ranked service can still significantly affect a particular individual, especially in a regional or remote setting where alternatives are limited.

Question 6: Meaning of significantly affect rights or interests – other legal frameworks and policies

Answer:

Yes. The OAIC should not read the new obligation in isolation but should map it against the established legal and policy landscape. The most relevant instruments are the Privacy Act 1988 (Cth) and its Australian Privacy Principles; the Commonwealth anti-discrimination framework, comprising the Age Discrimination Act 2004, the Racial Discrimination Act 1975, the Sex Discrimination Act 1984 and the Disability Discrimination Act 1992; and sector-specific safeguards such as the NDIS Quality and Safeguards framework and the Aged Care Quality Standards. For decisions affecting Aboriginal and Torres Strait Islander peoples, the CARE Principles for Indigenous Data Governance and the Maiam nayri Wingara Indigenous Data Sovereignty principles should guide both assessment and consultation. International comparators, including the OECD AI Principles and the EU AI Act, provide useful reference points, and Western Australian deployers should be able to align their disclosures with the information privacy obligations under the WA Privacy and Responsible Information Sharing (PRIS) Act so that they can, in practice, comply once across both regimes.

As an operational test, ADM significantly affects rights or interests where automation materially influences who receives service, opportunity, protection or cost in a way that is not trivial. Guidance should avoid capturing low-impact administrative automation while ensuring that risk scoring, profiling, triage and access decisions are not wrongly treated as low-risk.

We support a proportionate approach. Low-risk automation that does not materially affect

individuals should not carry extensive disclosure obligations, for example spell-checking, duplicate detection, appointment reminders, enquiry routing or non-personal analytics.

A use is not low-risk merely because it is internal, probabilistic, preliminary or one step removed from the final decision. A triage score, eligibility flag or risk classification can be highly significant where it determines whether a person receives service, review, investigation, opportunity or delay.

The OAIC should provide a decision tree that asks:

1. Is personal information used, including inferred or generated information?
2. Does automation materially influence the outcome, pathway, priority, eligibility, price, risk treatment or scrutiny?
3. Could the decision class have legal or similarly significant consequences for any affected group?
4. Is human involvement meaningful under the criteria in this guidance?
5. What level of disclosure is proportionate to the risk?

Question 7: Meaning of significantly affect rights or interests

Answer:

Yes. Differential pricing can significantly affect an individual's interests where automated profiling sets materially different prices for the same essential good or service. The significance lies not in the dollar figure alone but in whether the pricing targets essential goods or services, produces discriminatory patterns, or excludes a person from fair market access.

Worked example (differential pricing): Consider a digital marketplace that uses predictive profiling to set personalised prices, so that two individuals are quoted materially different amounts for the same essential good or service.

Determination: this is a significant ADM where it targets essential goods or services or produces discriminatory patterns, the significance lying not only in the dollar amount but in algorithmic exclusion from fair market access. As an indicative working threshold, a price variance of more than approximately 10 per cent on an essential good or service should be treated as presumptively material (the 10 per cent figure is an indicative benchmark proposed by the WA AI Hub to aid discussion; it is not legally binding, and the OAIC may calibrate or replace it through consultation).

The factors that make such pricing most material are, in order: impact on fundamental rights;

potential for systemic bias or discrimination; the essential rather than discretionary nature of the product; the economic cost to the consumer; and the individual's practical ability to detect and challenge the differential.

Applying this to the Daintree edge case, the WA AI Hub considers that the inflation of the price did significantly affect Marco's interests. A book is discretionary rather than essential, so on the product alone this case would not meet the indicative threshold above. The significant effect arises instead from the automated decision itself: Daintree used Marco's address as a postcode proxy for wealth to set a personalised, inflated price. That inference was opaque, undetectable to Marco, and inaccurate, since he is a freelance artist of limited means wrongly profiled as affluent, producing algorithmic exclusion from fair market access through a discriminatory, location-based inference.

On amount, we do not endorse a fixed 5 per cent or 20 per cent rule; consistent with the indicative threshold above, a variance of more than approximately 10 per cent should be treated as presumptively material. Marco is charged nearly three times the price (about a 186 per cent premium over the $35 benchmark), far beyond any reasonable threshold. The trigger is contextual rather than a single number: a percentage that is trivial on a discretionary item may be significant for an essential good or a financially vulnerable individual.

Using the Issues Paper scale of 1 (not important) to 5 (very important), we rate the dollar amount of differential pricing as 4, because absolute cost shapes impact but should not be assessed in isolation, and the significance of the product or service as 5, because its essential rather than discretionary nature is among the most decisive factors; as Marco's case shows, however, it is not the only path to significance, since a discriminatory or inaccurate proxy-based inference can render even discretionary pricing significant.

Question 8: Meaning of making a decision

Answer:

Yes. The automated non-delivery of the Manscon job ad to Rafqa should be treated as a decision.

"Making a decision" should be read functionally and purposefully. It covers any automated determination, selection, ranking, prioritisation, filtering or exclusion that materially influences a person's access to an opportunity, service, right or interest. A decision not to act, such as withholding or not displaying an opportunity, is as much a decision as a positive act. What matters is that automation, rather than the individual, determined the outcome.

On these facts, Opaque Window's algorithm made an affirmative determination: using gender as a metric, it selected who would and would not be shown the role and prioritised male graduates. Rafqa's non-receipt of the ad is the direct output of that automated selection, not mere inaction or chance. It is therefore a decision, and a significant one, because it uses a protected attribute (sex) as a proxy and curtails her access to employment and equal opportunity.

Both parties fall within scope. Opaque Window makes the automated targeting decision, and Manscon arranges for that decision to be made by procuring the platform and setting the campaign. The transparency obligation should reach the substance of who is included or excluded, regardless of how the step is labelled.

To keep the test workable, a genuinely trivial, non-consequential personalisation with no material effect on rights, opportunities or interests should not be a decision. The threshold is whether automation materially determines an outcome that significantly affects the individual, consistent with our answer to Question 7**.**

Question 9: Meaning of 'arranged for'

Answer:

Yes. Guidance on "arranged for" is needed, because the highest transparency risk now sits with automated decisions that an APP entity does not build itself, but procures, embeds or outsources, where the entity cannot see the third-party logic, yet remains accountable for it.

The reason guidance is required is that "arranged for" is where accountability most easily falls through the gap. An entity can satisfy the letter of its policy while the actual decision is made by a vendor's black-box system it neither controls nor understands. Guidance should confirm that arranging for an automated decision does not transfer responsibility: the entity that selects, configures or relies on the third-party system must still be able to explain, in plain language, what the system does and how it affects individuals, and should secure that explainability through procurement and contract terms.

We consider the following scenarios most in need of guidance: (1) software-as-a-service and white-label tools, where the deployer cannot see or audit the third-party logic; (2) recommendation, ranking and targeting engines that determine which individuals are shown an opportunity, service or price; (3) embedded or device-level automation with inaccessible logic; and (4) decisions by omission, where not showing a service, price or advertisement is itself the decision. Each is an "arranged for" arrangement in which the affected person may never know an automated decision occurred.

Worked example: an employer uses a third-party platform whose algorithm decides which users are shown a job advertisement, so some are never given the opportunity to apply. This is an arranged-for automated decision attracting transparency obligations. A final human sign-off on the shortlist does not extinguish the obligation where visibility or ranking has already been determined automatically, and excluding people on inferred attributes such as age, sex or location may also engage anti-discrimination law and the Fair Work Act 2009 (Cth).

To prevent compliance paralysis, the OAIC should set proportionate, safe-harbour due-diligence expectations for the arranged-for context, so smaller deployers such as micro, small and medium businesses and schools know what procurement-stage transparency steps are sufficient.

Question 10: Extent of disclosure

Answer:

Yes. Beyond the five considerations listed, entities should disclose four further elements to make ADM disclosures genuinely effective and meaningful.

First, the fact that ADM is being used at all. The most basic gap is that individuals often do not know an automated decision occurred. Entities should disclose, at or before the point of effect, that a decision about them is wholly or substantially automated, so the other transparency considerations have something to attach to.

Second, the role of automation and of any human reviewer. Disclosures should state whether the system decides, scores, ranks or only assists, and whether a human can and does meaningfully override the output. This prevents "human-in-the-loop" claims from obscuring decisions that are automated in substance.

Third, the main inputs and their consequences. Entities should disclose the main categories of personal information used, including inferred or third-party data, the principal factors that influence the outcome, the system's known limitations, and the likely effects on the individual. This is what makes a disclosure "specific enough to be meaningful" without a technical dump.

Fourth, accountability and recourse. To allow a decision to be genuinely contested, entities should disclose whether the system has been tested for accuracy and bias, who is accountable where a third-party or vendor system is used, and a clear, non-automated channel to seek human review or correction.

These additions should be delivered through layered disclosure: a concise plain-language notice for every affected person, with deeper detail available on request, and calibrated proportionately to the risk of the decision so smaller deployers such as micro, small and medium businesses and schools are not overwhelmed. To support consistency, the OAIC's guidance should illustrate these elements with practical, cross-sector examples, including borderline cases such as human rubber-stamping, administrative-only sorting, risk scores that trigger human review, and vendor-provided black-box systems.

5. Proposed model wording for OAIC guidance

We suggest the OAIC consider language to the following effect:

- An APP entity should not rely on the mere presence of a human reviewer to conclude that a decision is not substantially automated. Human involvement is meaningful only where the human reviewer has the authority, competence, time, information and independence necessary to assess the automated output and alter the outcome where appropriate.
- An APP entity's privacy policy should identify the categories of substantially automated decisions it makes, or arranges to be made, using personal information where those decisions may have a legal or similarly significant effect on individuals. The policy should describe the kinds of personal information used, including inferred or generated information, the role of automation, the likely effects of the decision, key safeguards and how individuals may seek correction, review or further information.
- An APP entity should take reasonable steps to ensure that disclosures about automated decision-making are current, accessible and understandable to the individuals likely to be affected. The level of detail should be proportionate to the significance of the decision and the reasonable needs of individuals to understand and act on the information.

We note that the guidance must also account for the rapidly evolving trajectory of agentic AI systems, which increasingly act autonomously across multiple steps, call tools, spawn sub-agents and operate over extended timeframes. The Singapore IMDA Model AI Governance Framework for Agentic AI (January 2026, updated May 2026) recommends unique agent identities tied to supervising users, and the Five Eyes Careful Adoption of Agentic AI Services advisory (1 May 2026) recommends cryptographically anchored agent identities with short-lived credentials and reversibility. We urge the OAIC to adopt a forward-looking, effects-based reading of the new obligation that captures these emerging deployment patterns.

Finally, we emphasise that vulnerability assessment under the guidance should integrate an intersectional, capacity-based approach rather than relying on static lists. Aboriginal and Torres Strait Islander peoples, people with disability, older Australians, children, those experiencing financial hardship, people from culturally and linguistically diverse backgrounds, those in

regional and remote communities, and survivors of domestic violence all warrant particular consideration. The guidance should recognise that geographic and connectivity disadvantage, cultural and linguistic factors, and cumulative or systemic effects of automation compound vulnerability in ways that static categories cannot capture.

We further recommend that the OAIC guidance require APP entities to provide at least one actionable counterfactual explanation where practicable, enabling individuals to understand what factors, if different, could have changed the outcome. For example, a credit decision disclosure might state that if an applicant's income had exceeded a particular threshold, the outcome would likely have differed. This approach makes transparency tangible and supports individuals in exercising their rights. Additionally, the guidance should encourage the adoption of machine-readable metadata schemas for automated decision logs to facilitate efficient regulatory auditing, interoperability across sectors and consistent record-keeping. Entities should maintain ADM documentation for a minimum retention period proportionate to the significance and reviewability of the decisions involved.

6. Conclusion

The WA AI Hub supports the OAIC's work to operationalise transparency for automated decision-making under APP 1. The guidance is an opportunity to set a national baseline that is practical for organisations and meaningful for people affected by ADM.

We encourage the OAIC to publish a guidance that is clear on scope, strong on substance, practical in implementation and attentive to the communities most likely to experience significant impacts. The test of success should be whether individuals and organisations can understand when ADM is used in consequential contexts and know what they can do next.

7. Disclosures and disclaimer

Authorship and submitting entity: The views expressed are those of the WA AI Hub and do not necessarily represent the views of any client, partner, member, university, government agency or other organisation with which the authors or the WA AI Hub are associated.

Interests and conflicts disclosure: **T**he WA AI Hub and its related entities provide AI governance training, strategic advisory and research services on a commercial basis, including the Responsible AI Governance Sprint™, AI for Me™, AI Business Index™ and related programs. Some recommendations in this submission concern areas in which the WA AI Hub

and its related entities may have a commercial interest. This submission is provided as a good-faith contribution to public consultation and is not made to obtain any specific commercial advantage. Other than these general commercial interests, the authors are not aware of any direct financial or personal interest that would materially affect the recommendations made.

Legal disclaimer: This submission is provided for general information and policy-consultation purposes only. It represents the considered views and opinions of the WA AI Hub as at the date of lodgement and does not constitute legal, regulatory, financial or professional advice, and must not be relied upon as such. It does not create any advisory, fiduciary or other relationship between the WA AI Hub (or its authors) and any reader. Recipients should obtain their own independent legal and professional advice before acting on any matter referred to in this submission.

While care has been taken to ensure the information is accurate at the time of writing, the WA AI Hub makes no representation or warranty, express or implied, as to its accuracy, completeness or currency. References to legislation, frameworks, third-party organisations, products, edge cases or examples are provided for illustration and discussion only and do not imply any endorsement, affiliation or assessment of any identified or hypothetical party. Any examples or fictional scenarios are illustrative and are not statements of fact about any real person or entity.

To the maximum extent permitted by law, the WA AI Hub, Customer Experience Insight Pty Ltd and the authors exclude all liability for any loss or damage arising from reliance on this submission. The WA AI Hub consents to the publication of this submission by the OAIC and authorises its inclusion on the public record. Where any part of this submission is considered confidential, that part has been separately identified to the OAIC. © 2026 Customer Experience Insight Pty Ltd trading as Western Australian AI Hub. All rights reserved.

Western Australia AI Hub®

Further recommendations (beyond the OAIC Issues Paper questions)

A. Interaction with accuracy, security, privacy impact assessment and accountability

The APP 1 transparency obligation should be connected to broader privacy governance. Transparency without assurance is insufficient.

ADM transparency depends on organisations having a clear understanding of their own systems. An APP entity cannot explain an ADM process it has not mapped, assessed or governed. We recommend that OAIC guidance link APP 1 transparency to:

- privacy impact assessments for high-impact ADM;
- data quality and correction processes for input data and derived scores;
- security controls for models, datasets, prompts, outputs and decision logs;
- records of decision logic, system changes, testing and monitoring;
- complaint handling and review processes;
- procurement controls for third-party ADM systems;
- ongoing monitoring for model drift, unexpected impacts and differential error rates.

The guidance should make clear that an entity's accountability is not displaced by outsourcing, cloud services, vendor products or model providers. If an APP entity uses personal information in a substantially automated decision process, it must be able to explain the use at an appropriate level of detail.

As set out in our response to Question 6, the OAIC should also map the transparency obligation against the broader legal and policy landscape – including the Privacy Act 1988 (Cth), the Commonwealth anti-discrimination framework, sector-specific safeguards, Indigenous data governance principles and relevant international comparators – and enable Western Australian deployers to align with the WA Privacy and Responsible Information Sharing (PRIS) Act so they can comply once across both regimes.

B. Practical implementation support required from the OAIC

The OAIC should provide practical tools that enable consistent implementation across sectors and entity sizes.

We recommend that the OAIC publish the following implementation resources with the guidance:

- a one-page decision tree for determining whether the obligation applies;

- a model ADM disclosure table for privacy policies;
- sample plain-language clauses for common ADM use cases;
- a glossary distinguishing automation, AI, ADM, profiling, recommendation, ranking, scoring, rules-based systems and generative AI;
- a checklist for assessing meaningful human involvement;
- a template internal ADM register;
- case studies for small business, public sector agencies, universities, health providers, finance, insurance and digital platforms;
- guidance for entities procuring ADM systems from vendors;
- examples of good, insufficient and misleading privacy policy disclosures.

These resources should be designed for practical integration, not only legal interpretation. Consistent templates will reduce compliance cost, improve comparability and help individuals find the information they need.

C. Additional comments on regulatory coherence and responsible innovation

The OAIC should position the guidance as an enabling framework for trustworthy innovation, not as a barrier to Responsible AI integration.

Clear transparency guidance supports innovation by setting predictable expectations. Organisations are more likely to deploy ADM responsibly when obligations are concrete, proportionate and aligned with operational practice. Ambiguity increases risk for individuals and uncertainty for organisations.

The guidance should be interoperable with existing and emerging AI governance practices, including privacy impact assessment, data governance, security risk management, procurement assurance, model documentation, testing, monitoring, auditability and review. It should avoid implying that transparency alone is sufficient to make an ADM system fair, accurate or lawful.

The WA AI Hub also recommends that the OAIC continue engagement with industry, government, academia and civil society after publication of the guidance, including through sector roundtables and practical implementation workshops. The regulated community will need examples, not just principles.

D. Standard ADM disclosure table (template for APP entities)

We recommend that OAIC guidance provide a standard table that APP entities can adopt. The following fields should be included as a minimum where an entity makes, or arranges for another party to make, substantially automated decisions using personal information that may have legal or similarly significant effects.

Disclosure field	Minimum content
Decision category	The type of decision, for example eligibility, pricing, prioritisation, risk assessment, fraud detection, recruitment or service allocation.
Purpose	Why the decision is made and the organisational function it supports.
Personal information used	Main categories of collected, inferred, derived, generated or predicted personal information used.
Role of automation	Whether the system determines, recommends, ranks, scores, flags, prioritises, triages or otherwise influences the decision.
Human involvement	Whether human review occurs, at what stage, and whether the reviewer can change the outcome.
Likely effects	The kinds of consequences the decision may have for individuals.
Safeguards	Key safeguards such as data quality checks, bias testing, monitoring, audit logs, review rights and escalation.
Third-party systems	Whether a vendor or external service materially supports the decision, without requiring disclosure of confidential source code.
How to seek review	Contact point and process for correction, clarification, complaint or human review.
Last updated	Date or version of the ADM disclosure.