Submission to the Office of the Australian Information Commissioner

Issues Paper: Automated Decision-Making Transparency Obligation (APP 1)

Submitter	[redacted]
Capacity	[redacted] **AegisGuard (pre-pilot SaaS in development)**
Sector	**Mine-site safety and compliance, Western Australia**
Date	**27 May 2026**
Status	**Public submission — no confidential sections**
Contact	[redacted]

About this submission

This submission responds to a subset of the questions in the OAIC's Issues Paper on the ADM Transparency Obligation (APP 1). The author is a non-lawyer founder building safety-decision tooling for the Australian mining sector. The views expressed are personal and reflect the author's perspective as a practitioner building software in which AI agents generate safety recommendations that workers and supervisors then act on. Nothing in this submission is intended as legal advice, as a statement of fact about any present or future product, or as a commitment binding any future entity. References to AegisGuard describe a design philosophy and product intent, not an operational product or company.

This submission addresses Questions 1, 2, 5, and 10. The author has not commented on the remaining questions either because they fall outside the author's area of expertise or because the author does not have a settled view at this stage.

Question 1 — Substantially and directly related to making a decision

The OAIC asks which factors are relevant in assessing whether a computer program substantially facilitates and is directly connected to a human decision-maker's decision, and invites ranking on a 1 to 5 scale.

In my view, the listed factors should be ranked as follows:

Factor	Rank	Comment
Degree of reliance on the ADM system output	**5**	Where humans habitually accept the output without independent review, the ADM is effectively making the decision regardless of formal labels.

Factor	Rank	Comment
Ability and likelihood of human override over an ADM decision	**5**	Ability alone is too weak. Many systems formally permit override that does not occur in practice. Likelihood should be weighted equally.
Nature of the output (advisory vs determinative)	**4**	Useful as a signal but easy to game. A label of advisory does not change the underlying influence of the output.
Integration of ADM into decision-making workflow	**4**	Where the ADM output is the only structured input the human sees, integration is high regardless of nominal status.
Transparency and explainability of outputs	**3**	Important for governance generally, but a separate question from whether the ADM is substantially and directly related to the decision.

The author would respectfully suggest three additional factors warrant inclusion:

1. Time available to the human decision-maker. In time-pressured contexts (safety calls in mining, clinical triage, real-time fraud screening) a human may have seconds to consider an ADM output before acting. A theoretical right to override is meaningless if the operational window for review is too short for meaningful consideration. Guidance that ignores this factor will under-capture ADM systems that dominate decisions despite formal human oversight.
2. Whether overrides are recorded, justified, and reviewed. Systems that capture override events with worker reasoning, and that surface override patterns to supervisors, demonstrate a genuine human role. Systems that do not record overrides cannot demonstrate one. The presence or absence of override-capture infrastructure is, in the author's view, observable evidence of whether the human role is substantive or decorative.
3. The measurable override rate in practice. Once a system is operating, the actual rate at which humans diverge from ADM output is the most reliable indicator of human role. Entities relying on the advisory-only framing should be expected to be able to point to override-rate data.

These three additional factors share a theme: the test for substantially and directly related should be functional rather than formal. Labels and policy commitments are necessary but not sufficient.

Question 2 — Fictional Edge Case: NACIA and GPTea

The OAIC asks whether the described use of GPTea by NACIA staff would be substantially and directly related to making decisions.

Yes — most likely, on the facts described.

The author's reasoning:

The edge case states that NACIA staff always have oversight and make the final decision based on the recommendation provided by the chatbot. Taken at face value, this is formal oversight. Whether it is substantive oversight depends on factors the edge case does not specify: how often NACIA staff diverge from the chatbot's recommendation; what time and information they have to review the underlying candidate profiles independently; what training and tooling they have to challenge the recommendation; whether their performance metrics reward or penalise diverging from the recommendation.

In contexts where a chatbot provides a structured recommendation alongside a summary of evidence the human did not independently read, the practical effect is often that the human becomes an editor of the chatbot's conclusion rather than an independent decision-maker. If that is the operating reality at NACIA, the ADM is substantially and directly related to the eligibility decision regardless of the formal designation of the chatbot as advisory.

The author would respectfully suggest the OAIC guidance treat fact patterns like NACIA as presumptively in scope unless the entity can evidence that the human role is substantive on the factors set out in Question 1. Placing the evidentiary burden on the deploying entity, rather than on the affected individual or the regulator, reflects the asymmetry of information between the parties.

Question 5 — Meaning of significant service or support

The OAIC asks what should be considered a significant service or support and invites ranking of a list.

In my view, the listed examples should be ranked as follows:

Service or support	Rank (1-5)
Physical assistance / support	**5**
Financial assistance	**5**
Access to essential banking and credit services	**5**
Access to essential utilities	**5**
Access to educational services	**4**
In-home services	**4**
Access to telecommunications	**4**

The author would respectfully suggest three additional categories should be expressly captured:

4. Workplace safety determinations. Decisions made by ADM about whether a worker is fit to work, permitted to enter a hazardous area, or authorised to operate a piece of equipment have direct physical-safety consequences. In high-hazard sectors (mining, construction, heavy industry, transport, emergency services) these decisions affect life

and limb. The author would suggest these be ranked equivalent to physical assistance / support — that is, 5 — and that guidance expressly name workplace safety access and authorisation decisions as in scope.

5. Employment access and continuation. Decisions made or substantially supported by ADM that affect whether an individual is hired, retained, suspended, or dismissed have profound effects on individual rights and interests. This is partially captured by the discriminatory job ad edge case in Question 8 but warrants standalone treatment.
6. Regulatory or licensing status. Decisions affecting whether an individual holds, retains, or loses a licence, registration, certification, or other regulatory status are significant in their own right and frequently downstream of automated screening.

The author would respectfully suggest the test for significant service or support should not be limited to consumer-facing transactions. Many of the highest-stakes ADM decisions for individuals occur in workplace and regulatory contexts, where the affected individual has even less practical ability to challenge or contest the decision than in a consumer setting.

Question 10 — Extent of disclosure

The OAIC sets out five considerations for meaningful and accessible disclosure and asks whether there is anything additional entities should disclose.

The author would respectfully suggest the following additions to the OAIC's view of effective disclosure:

7. The categories of potential adverse consequences for the individual. A disclosure that the ADM is used to support eligibility assessments tells the individual nothing about what could go wrong for them. Naming the categories of adverse outcome (denial of service, delay, escalation, referral, restriction of access) lets the individual understand what is at stake.
8. The procedural pathway for challenge or contestation. Where guidance recognises that effective disclosure should be framed in a way that allows the information and the decision to be challenged or contested, the entity should be expected to set out the actual pathway for doing so. A vague right to contact us is less useful than a specific procedure (timeframes, response standards, escalation steps).
9. The role of the human in the decision loop. Where the entity asserts that a human meaningfully reviews ADM output, the entity should disclose what the human reviews, what authority the human has to depart from the ADM, and any constraints on the human's ability to do so. This is the consumer-facing counterpart to the test under Question 1.
10. Whether the ADM output is recorded for audit and how long it is retained. Individuals should be told whether the decision and the contributing ADM output will be preserved, so they understand whether records are available if they later seek to contest the decision or make a complaint.
11. Whether a non-ADM alternative is available. In contexts where the entity offers any non-automated pathway (manual review, in-person assessment, paper application), the existence of this pathway should be disclosed.
12. For high-stakes contexts, evidence of independent testing or assurance. Where the ADM is applied to decisions in safety, employment, finance, or other high-stakes categories, disclosure should include whether the system has been subject to

independent testing, audit, or assurance, and where the consumer can find that information.

The common thread is that effective disclosure should equip the individual to act — to understand what is at stake, to challenge the decision, and to know what evidence is preserved. Disclosure framed as a compliance exercise rather than a tool for individual agency is, in the author's view, likely to fail the underlying transparency purpose.

Closing

The author thanks the OAIC for the opportunity to contribute to the development of this guidance and would be available to clarify any point in this submission if useful.

Submitted: 27 May 2026.

[redacted] AegisGuard (pre-pilot SaaS in development; mine-site safety and compliance, Western Australia).

Contact: [redacted]