

Automated Decision Making Standard

Approved by:	Leadership Team
Standard Owner:	General Manager Information
Review date:	1 March 2023

1 Definitions

- 1.1. **Automation** is the use of systems or components of systems to replace repeatable processes in order to reduce dependency on manual actions or interventions.
- 1.2. Processes can be automated based on the application of:
 - (i) known business rules, and/or
 - (ii) data-based algorithms without involvement or assessment by a human, including statistically or analytically derived patterns in machine learning or Artificial Intelligence.
- 1.3. A **decision** for the purpose of this standard is the action of choosing between two or more possible actions and may be derived from legislative, cabinet or other legal authority or can be operational, and may be discretionary or non-discretionary.
- 1.4. An **automated decision** for the purpose of this standard is a decision within an automated process where there is no substantial human involvement in making the decision.
- 1.5. **Discretionary decisions** require an exercise of judgment to choose between two or more possible actions.
- 1.6. A non-discretionary decision **does not** require any exercise of judgement to determine the appropriate action.
- 1.7. A **Business Owner** is the person who is accountable for the automated process at any given time.
- 1.8. For the purposes of this standard, "**bias**" refers to the tendency of an automated decision process to create unfair and unjustified outcomes, such as favouring or disfavouring one group over others.
- 1.9. Automated decisions may be biased because, for instance, the **datasets** they rely on are biased, potentially as a result of how data was collected in the past, or because **social conditions** mean that some groups are overrepresented in some risk groups.
- 1.10. The prohibited grounds of discrimination are set out in the **Human Rights Act 1993 Section 21**: sex, marital status, religious belief, ethical belief, colour, race, ethnic or national origins, disability, age, political opinion, employment status, family status and sexual orientation.
- 1.11. Discrimination on these grounds can be justified under the **Bill of Rights Act 1990 Section 5**, but only to such reasonable limits that are lawful and can be clearly and ethically justified.

2 Applicability

- 2.1 This standard **must** be applied using the operational guidance when:
- (i) there is a proposal to automate a decision (as defined in sections 1.3 and 1.4), **AND**
 - (ii) the automated decision has the **potential to affect**, an individual's entitlement, obligations, or eligibility status for support delivered or funded by the Ministry of Social Development (the Ministry).
- 2.2 Where a complex algorithm is being proposed, the Model Development Lifecycle **must** be used.
- 2.3 Any exception to this standard **must** be approved by the Chief Executive before automated decision-making can be implemented.

3 Standard Requirements

3.1 General

- 3.1.1 Automated decision-making **must**:
- (i) improve efficiencies and effectiveness of decision making and balance factors such as cost, accuracy, reliability and safeguarding the wellbeing of those affected.
 - (ii) comply with all applicable Ministry policies and standards that relate to the privacy, security and management of information.
- 3.1.2 Automated decision-making **must not** create inefficiencies for those the decisions directly affect, for example, adding in manual workarounds for a client to enable automation, or unnecessarily increasing time from application to notification of a decision than would otherwise occur if it was manually completed.
- 3.1.3 There **must** be clear, relevant, and accessible guidance for users who are required to input or provide data to be used in automated decision-making, for example, a service user entering their information in MyMSD.

3.2 Accuracy, bias and discrimination

- 3.2.1 Accuracy and reliability **must** be assessed before automated decision-making is implemented to ensure, insofar as possible, that automated decision-making is producing expected results, that automated decisions do not deny clients full and correct entitlement (FACE), and bias and discrimination is well managed.
- 3.2.2 Based on the assessment carried out under 3.2.1, where evidence suggests that automated decision-making has resulted in unintended bias, steps **must** be taken to identify and remove or mitigate the unintended bias, and any residual risk **must** be accepted by the Business Owner.
- 3.2.3 Where unintended bias cannot be removed or sufficiently mitigated, substantial human involvement **must** be included in the process. This would then mean that the decision is no longer an automated decision.

3.3 Policy, fraud and legal considerations

- 3.3.1 Automated decisions **must** be lawful and align with policy intent.

- 3.3.2 An assessment must be undertaken to determine whether any proposed automated decision-making has the potential to:
- (i) increase the likelihood that people will commit internal or external fraud or client non-compliance; or
 - (ii) increase the scale or size of potential internal or external fraud or client non-compliance.
- 3.3.3 Any increased risk of fraud **must** be accepted by the Business Owner before automated decision-making can be implemented.
- 3.3.4 Prior to automating discretionary decisions, you **must** ensure that any legal risk(s) are identified and mitigated or accepted by the Business Owner before automated decision-making can be implemented.

3.4 Transparency

- 3.4.1 The Ministry **must** make information publicly available about:
- (i) what policies and processes are used to identify and mitigate risks associated with automated decision-making, in particular those that relate to human rights and ethics; and
 - (ii) information about what decisions are made using automation.
- 3.4.2 The Ministry **must** provide as much transparency as possible, while minimising the risk of fraud, to clearly explain how a decision has been made through the use of automation, including the role of humans in automating the decision and who is accountable for the process and the decision made.
- 3.4.3 If a lawful restriction prevents explanation, the Ministry **must** provide as much explanation as possible to the individual and clearly outline what details have been withheld and why.
- 3.4.4 The use of automated decision-making **must** be communicated to the individual in a way that is easy to understand and clearly shows a decision was made using automation, the outcome of that decision, and the process for challenging or appealing decisions.

3.5 Human oversight

- 3.5.1 A visible and accessible point of contact **must** be nominated for public inquiries about decisions made using automation.
- 3.5.2 The Ministry **must** provide a channel for challenging or appealing decisions made using automation and this channel **must** be made easily visible and accessible to the individual(s) impacted by the decision.
- 3.5.3 The process to review an automated decision that has been challenged or appealed **must not** itself be an automated process.

3.6 Compliance and assurance

- 3.6.1 Compliance with this standard **must** be verified for all new uses of automated decision-making through the existing Security, Privacy, Human Rights and Ethics Certification and Accreditation process.
- 3.6.2 Regular monitoring **must** be carried out to ensure that the automated decision-making continues to produce expected results and to ensure bias and discrimination are well managed.

- 3.6.3 A compliance review **must** be carried out at least once every three years or more frequently (based on the nature and level of risk connected to the process) to ensure that any automated decision-making that is approved under this standard continues to meet the requirements of the standard.

4 References

- 4.1.1 Key artefacts used as inputs in the development of this Standard.

[Principles for Safe and Effective Use of Data and Analytics](#)

[Algorithm Charter for Aotearoa New Zealand](#)

[Data Protection and Use Policy](#)

[Privacy, Human Rights and Ethics Policy \[in development\]](#)

- 4.1.2 Artefacts that directly support the application of this Standard.

[Operational Guidance](#)

[Data Model Lifecycle](#)

[PHRaE guidance: Operational analytics and automation](#)

APPROVED 15 FEB 2022