GU004 - AI & Automation Risk and Opportunity Analysis Overview

Revenue NSW Wide

Notes:

A desktop review was completed to map out the operating environment, including legal and other requirements and guidance and best practice. This review was then developed into this risk analysis structured around the core components of the NSW AI Ethics Policy and the AI and Automation Integrated Lifecycle Stages.

The appropriate components of this risk analysis will be integrated into the relevant Standards and Guidance.

Community benefit

Considerations	Risks	Opportunities	Controls for consideration

NSW Ombudsman – legal/decision making (*'The New Machinery of Government: Using Machine Technology in Administrative Decision Making'*)
 	
DCS-2024-04 – Use of Artificial Intelligence by NSW Government Agencies
 	
NSW AI Strategy
 	
NSW AI Ethics Policy (V2)
	• **R.1 - Technology can produce systemic injustices at scale very quickly**
	  ○ **R.1.1** – Limited public or oversight body visibility
	  ○ **R.1.2** – Complexity of technology makes it harder to understand process or recognise errors
	  ○ **R.1.3** – Errors more likely to impact large volumes of people
	  ○ **R.1.4** – Greater risk to lower socio-economic groups
	• **R.2 –** Feedback and complaint mechanisms and fitness for purpose
	• **R.3** – Speed of automated decisions balanced with speed of human intervention.
		• **O.1** – Improved customer experience driven by a community benefit or government benefit
		• **O.2** – Efficiency and cost saving
		  ○ **O.2.1** – Increased productivity and refocus of resources on 'higher value' work
		  ○ **O.2.2** – Triage system based on authorisation, procedure, assessment, and documentation.
		  ○ **O**.2.3 – Consideration of all possible harms to come into risk assessment when considering impacts of projects
		• **O.3** – Dedication of teams to ensuring that there is sufficient oversight of machine design technology
		  ○ **O.3.1** – Involvement of Policy, Risk, TAB, and other associated teams from day one
		  ○ **O**.3.2 – Consideration of all potential risks, including adverse outcomes
			Overview:
			• **C.1** – Ongoing oversight on systems across the process from end-to-end to ensure community interests are being met
			• **C.2** – A 'customer first' approach when weighing up efficiencies and cost saving (should not come at the expense of staffing or manual review)
			• **C.3** – Analysis may be conducted of projects prior to implementation by a multi-disciplinary team to consider all angles
			Concept:
			• **C.1.1** – Policy oversight from beginning of process – is what we are trying to do permissible under legislation?
			• **C.1.2** – Internal legal team that can answer questions as to legislative compliance from a place of expertise – particularly for areas where there are no legal experts (e.g. on privacy issues).
			• **C.3.1** - General risk factor assessment considerations to include:
			  ○ Whether the AI is delivering a new or existing service;
			  ○ The potential to cause discrimination from unintended bias;
			  ○ Whether the AI system is a single point of failure for your service or policy;
			  ○ If there is a sufficient experienced oversight of the AI system;
			  ○ Over-reliance on the AI system or ignoring the system due to high rates of false alert;
			  ○ Whether the linkage between operating the AI system and the policy outcome is clear;
			  ○ The system's explainability and transparency regarding generated content and decisions; and
			  ○ Were other non-AI systems considered?
			Discovery:
			• **C.2.1** – Customer-centric design approach that results in increased productivity – people focusing on customer interaction and engagement
			Develop:
			• **C.3.2** – Consistent updates through development process on approvals, chains, and processes
			Deliver:
			• **C.1.1** – Consideration of if a mechanism appears to be a 'tick and flick' exercise – how would the public view our operational processes.
			Launch:
			• Nil
			Operate, Monitor, and Iterate:
			• **C.1.3** – Proactive and ongoing external testing and auditing of systems

Considerations	Risks	Opportunities	Controls for consideration

			• **C.1.4** – Reactive individual review rights
			• **C.1.5** – Authorisations measurement of machine-made determinations
			• **C.1.6** – Measurement of prescribed requirements metrics – visibility, avoidability, subjectivity to testing, explainable, accurate, subjectivity to audit, replicable, internally reviewable, externally reviewable, compensable, and secure
			• **C.1.7** – AI system to be assessed to ensure that it improves on existing approaches to deliver the outcomes described in:
			  ○ the Human Services Outcome Framework;
			  ○ the Smart Places Outcome Framework;
			  ○ NSW Treasury Budget Outcomes;
			  ○ your Agency Strategic Plan; or
			  ○ another relevant NSW outcomes framework
			• **C.3.3** - Register of decision-making systems across Revenue NSW
			• **C.2.2** – Cost saving needs to be measured against transparency and disclosure of the system, as well as the use of the ADM
			Retirement:
			• **C.3.4** – Register retained of retired systems and how these systems made their determinations (i.e. process breakdown)
			  ○

	•	•	

Legal compliance & administrative law

Considerations	Risks	Opportunities	Controls for consideration

NSW Ombudsman – statutory decision making (‘*Administrative law and automated decision making*’)	• **R.4** – Failure to comply with administrative law resulting in maladministration or invalidated decisions
	  ○ **R.4.1** – This may also result in a risk that one or more legislative requirements are not met
	  ○ **R**.4.2 – Consideration of all relevant legislation, including the PPIP Act (and associated codes and directions), HRIP Act, GIPA Act, Privacy Act (Commonwealth), and Data Sharing (Government Sector) Act
	• **R.5** – Maladministration as a holistic view (part maladministration may invalidate wider decisions)
	• **R.6** – Decision makers can only be people, not machines – and cannot simply ‘accept’ automated results without question
	  ○ **R.6.1** – This risk emphasised by KPIs and staff not having enough time to review decisions/machine outputs
	  ○ **R.6.2** – No full scope as to what systems can and cannot be used to make or assist decisions
	• **R.7**- Loss of ability to challenge decisions in a court as no ‘decision’ has been made by a machine
	  ○ **R.7.1** – Loss of natural justice because of this – serious maladministration potential
	  ○ **R**.**7.2** – Risk of additional key considerations being missed or overlooked if not considered (e.g. exceptional circumstances)
	• **R.8** – Parliamentary representation and the ‘Carltona’ principle – any decision made by a department is taken to be a
		• **O.4** – If a decision is not ‘discretionary’, there is an opportunity to fully automate the process.
		• **O.5** – Potential to increase consistency and repeatability by providing input to decision-making to reduce or identify human biases within the process
		• **O.6** – ADM systems can streamline or improve customer response times by helping the relevant officer in a decision-making process
		• **O.7** – Opportunity to improve court referrals and decision-making process chains via transparency
		• **O.8** – Opportunity to improve consistency in Ministerial approach and representation (ref **O.5**).
		  ○ **O.8.1** – Consistency of decisions for customers may also be enhanced
		  ○ **O**.8.2 – Accuracy of information (both in data consumption and where manual review of all items is not possible with FTE)
			Overview:
			• **C.4** – Assessment of legal requirements and discretionary decision making embedded into all current and previous automation and AI systems
			• **C.5** – Decisions, if assisted by machine learning, must consider discretion across the process to ensure that we have a balance re: human bias vs machine bias
			• **C.6** – Guidance for streamlining process to ensure there is a clear pathway for a decision-making pathway
			• **C.8** – Overarching principle and ‘front of mind’ that Revenue NSW is representing the Minister’s Office in all we do
			Concept:
			• **C.5.1** – The determination if a decision is discretionary should be the first step in the concept/assessment process
			Discovery:
			• **C.4.1** – Case-by-case consideration must be embedded by looking at the statute, the purpose of it, and context in which it applies – policy design must be embedded
			• **C.5.2** – Key requirements to be embedded in design, including authorisation to conduct a particular action, appropriate procedures that are well documented, appropriate assessment undertaken with all relevant material, analysis and considerations must be considered with documented hierarchy, and adequate documentation showing how the decision was made.
			Develop:
			• **C.4.2** – Needs to consider beyond business rules/our legislation and consider the broader legislative context when building a system/approach
			• **C.8.1** – Embedding of impartiality and structured approach to Ministerial impartiality
			• **C.5.3** – Clear articulation of ‘human-in-the-loop’ or ‘human-on -top’ processes
			  ○ **C.5.4** – Discretion and practical principles and key points around bias of accepting outputs of technology without critical review
			Deliver:
			• **C.6.1** – Partner with Ombudsman to deliver relevant training and awareness on the concept of decision making
			• **C.7.1** – Application of legal principles for decision making by operations: openness, fairness, accountability, consistency, rationality, legality, and impartiality
			• **C.8.2** – Exceptional circumstances may be able to be flagged for a greater level of review to ensure consistency and customer coverage
			Launch:
			• Nil
			Operate, Monitor, and Iterate:

Considerations	Risks	Opportunities	Controls for consideration

	decision made by the relevant Minister (Delegations)		• **C.6.2** – Framework to assess the relevancy of various pieces of legislation when collecting data, including:
			  ○ *Privacy and Personal Information Protection Act 1997* (NSW);
			  ○ *Anti-Discrimination Act 1977* (NSW);
			  ○ *Government Information (Public Access) Act 2009* (NSW);
			  ○ *Health Information and Information Privacy Act 2002* (NSW);
			  ○ *Public Health Act 2010* (NSW);
			  ○ relevant Acts for your agency or space; and
			  ○ *Police Act 1990* (NSW).
			• **C.6.3** – Consider potential for bias and adoption of bias, and compliance with the following:
			  ○ *State Records Act 1998* (NSW)**;**
			  ○ *Road Transport Act 2013* (NSW);
			  ○ *Anti-Discrimination Act 1977* (Cth)**;**
			  ○ *Disability Discrimination Act 1992* (Cth);
			  ○ *Workplace Surveillance Act 2005* (NSW);
			  ○ *Australian Human Rights Commission Act 1986* (Cth);
			  ○ *Age Discrimination Act 2004* (Cth);
			  ○ *Racial Discrimination Act 1975* (Cth);
			  ○ *Sex Discrimination Act 1984* (Cth);
			  ○ *Surveillance Devices Act 2007* (NSW); and
			  ○ *Telecommunications (Interception and Access) Act 1979* (Cth).
			• **C. 7.2 -** Control of decision mapping for any decision proceeding to court – meaning we have an audit trail if there is an administrative point of law contested
			• **C.7.3** – Consider uplift of processes that ‘appear’ automated (for instance, templates with minor changes to customers) to ensure people are utilising their ‘right to be heard’
			• **C.8.3** – Routine review of decisions to be conducted with criteria to determine a cohesive approach taken to system review
			Retirement:
			• **C.8.4** – Consideration of accuracy and completeness once a system has been retired prior to implementing full automation (how do we ensure that there is a focus on the data ingestion to ensure that future ADM systems have accuracy enshrined?)

Legislative reform - *‘The Role of Parliament in authorising machine technology’*	• **R.13** – Legislative reforms do not consider or may be ambiguous about technology, and use may even be prevented by legislation • **R.14** – If legislative reforms enable to authorise technology use, this may introduce complacency, unintended consequences, or inconsistencies in instruments	• **O.13** – Opportunity to amend or introduce legislation to remove ambiguity, enable use of technology, and provide clear boundaries/requirements • **O.14** – Alignment and codesign of technology and legislative instruments to remove ambiguity and align design and operation of technology within the legislation	**Overview:** • **C.13.1** – Ensuring that when a new project is stood up that requires legislation change that there is consideration of the technological requirements (if any) that need to be in the legislation to be included in drafting **Concept:** • **C.13.2** – Consideration of current legislative state should be established – needs to be a fully-formed picture of what Revenue NSW is permitted to do under legislation • **C.15.1** – Consideration of the following criteria around legislative change and risk/ethics standards and guides:

Considerations	Risks	Opportunities	Controls for consideration
	• **R.15** – Legislative instruments put in place may create practical issues where it is difficult or impossible for administrators to use discretion to challenge or overrule ADM decisions	• **O.15** – Legislative ability to create a 'human in the loop/human on top' to ensure accountability for decisions made by ADM.	○ Visibility/transparency ○ Avoidability (can you opt out?) ○ Testing & monitoring ○ Explainable ○ Accurate ○ Records kept & subject to audit ○ Replicable ○ Internal and external review ○ Compensable ○ Privacy and data secure **Discovery:** • **C.14.1** – Legislative mapping of process and projected process under new legislation to identify any differences and drafting instructions to be informed by this • **C.15.1** – Consider whether legislative amendment is prudent or necessary, or, alternatively, if an amendment will result in adverse impacts on staff or customers **Develop:** • **C.13.3** – Management of legislative parameters in consultation with digital (what can we do/can't we do) and looped in on the drafting • **C.14.2** – Consideration of risk-based choices/trade-offs in development and design by a 3rd party or vendor may not consider all legislative/other requirements or government contexts **Deliver:** • Nil **Launch:** • Nil **Operate, Monitor, and Iterate:** • **C.13.4** – Regular staff training/monitoring for legislative compliance and forums so that questions can be asked • **C.14.3 –** Review of legislation and any relevant required legislative changes that may need to be made if incidents arise. • **C.15.3 –** Review of decisions and the interaction of human/ADM decisions and balance (is this still working for business requirements?) **Retirement:** • **C.14.3** – Previous legislative interpretation and implementation to be mapped and a 'timeline' of previously established and accepted practice created.

Fairness

Considerations	Risks	Opportunities	Controls for consideration

NSW Ombudsman – appropriate procedures (*'Administrative law requirements for good decision making'*)
 	
AIAF
 	
NSW Automation Guide
	• **R.9** – Algorithmic bias
	  ○ **R.9.1** – Customers may face a bias when dealing with an algorithm that may give adverse outcomes based on statistical information
	  ○ **R.9.2** – Decisions could be challenged as biased or invalid based on Revenue NSW administered legislation
	• **R.10** – Inadequate documentation and record-keeping for AI processes
	• **R**.11 – Decisions not made 'on the merits of the case' where the line blurs between discretionary and non-discretionary
	• **R.12** – Lack of coverage across the business in relation to AI projects
		• **O.9 –** Opportunity to identify customers based on vulnerability criteria to offer them additional support
		  ○ **O.9.1** – ADM may assist in mitigating biases or conflicts of interest
		• **O.10** – Guidelines and communication to customers and stakeholders that provides adequate documentation
		• **O.11** – Delineation between discretionary and non-discretionary parts of a decision (particularly as a mapped process)
		• **O.12** – Opportunity to reconstruct projects from 'just IT' to multi-disciplinary (bridging the gap between IT, policy, operational, legal, and other expertise)
			Overview:
			• **C.9** – Parameters to be placed on how Revenue NSW defines and identifies vulnerable customers (particularly when taking punitive actions)
			• **C.10** – Framework to be communicated to customers and stakeholders (within ordinary bounds) to ensure coverage
			• **C.12.1 –** Risk framework across the business that considers the following:
			  ○ Consequences of inaccurate or incomplete data use;
			  ○ Poorly defined descriptions and indicators of 'fairness';
			  ○ Decisions to exclude or include outlier data;
			  ○ Informal or inconsistent data cleansing and repair protocol and processes;
			  ○ Informal bias detection methods (best practice includes automated testing);
			  ○ Inadvertently creating new associations when linking data and consequences;
			  ○ Likelihood that re-running scenarios could produce different results; and
			  ○ Differences in the data used for training compared to the data for intended use.
			• **C.12.2** – Further risk analysis that asks:
			  ○ Explanation as to why the selected data is being used
			  ○ Is the data that you need for the system available and of appropriate quality given the potential harms identified (AI systems may be impacted quite heavily by poor data quality)?
			  ○ Does the data reflect the population that will be impacted by your system?
			  ○ Have you considered how your AI system will address issues of diversity and inclusion (including geographic diversity)?
			  ○ Have you considered the impact regarding gender and on minority groups including how the system might impact different individuals in minority groups when developing this AI system (including – those with a disability, LGBT+ communities, people from CALD backgrounds, Aboriginal and Torres Strait Islanders, children and young people, and people from varying socio-economic backgrounds.
			  ○ Do you have appropriate performance measures and targets (including fairness ones) for your AI system, given the potential harms (accuracy and precision are readily quantifiable for most systems which predict or classify outcomes. This performance can be absolute or relative to existing systems)?
			  ○ Do you have a way to monitor and calibrate the performance (including fairness) of your AI system (operational uses of AI which are continuously updated/trained can quickly move out of performance thresholds. Supervisory systems can monitor system performance and alert where calibration is needed)?
			  ○ Elevated risk uses of AI should have clear performance monitoring and calibration schedules – for elevated risk uses of AI which are continuously training and adapting with moderate residual risks, weekly performance monitoring and calibration is recommended. For low risk, monthly evaluation and calibration is recommended. For operational systems with High Risk or Very High Risk, a custom evaluation and calibration is required.

Considerations	Risks	Opportunities	Controls for consideration
			○ **Concept:** • **C.9.1** – Embedding of procedural fairness principles from end-to-end for processes to cover the overarching principles of transparency, accountability, quality (robust and repeatable), and management of customer expectations • **C.12** – Construction of a 'base-level' team to ensure that coverage across disciplines is adequate **Discovery:** • **nil** **Develop:** • **C.10.1** – Inbuilt communication about machine-made decisions that informs customers when a decision is informed or made by technology (should be able to be easily described to a layperson where required) • **C.11.1** – Mapping of discretionary and non-discretionary elements of a process to ensure that each part of the process is understood and mapped **Deliver:** • **C.9.2** – Procedures and work instructions to clearly articulate key requirements and decision-making **Launch:** • **nil** **Operate, Monitor, and Iterate:** • **C.9.3** – Ensuring that in review processes we have ensured that the 'no bias rule' and the 'hearing rule' were appropriately managed for all processes • **C.10.2** – Decisions made by or with ADM supports need to be documented and appropriate records of the decision made, including: details for the decision, reasons for the decision, date of the decision, identity of the decision maker, copies of written notification, and file notes of non-written communication. • **C.11.2** – Operational staff to ensure, when keeping records, that, where reasons are required, they must be accurate, meaningful, and understandable • **C.11.3** – Discretionary and non-discretionary decisions to be clearly noted, and processes to be clarified to ensure coverage • **C.9.4** – Review of stakeholder engagement, including frontline awareness and understanding **Retirement:** • **C.10.3** – Ability to provide, even once a system is retired, when requested, reasons to those who were affected by a decision • **C.11.4 –** Potential to use the Automation Maturity Matrix as laid out by the NSW Automation Guide for review

Considerations	Risks	Opportunities	Controls for consideration

'Good practice for designing and implementing machine technology'	• **R.16 -** Cognitive bias of developers, owners and operators, and decision makers
	• **R.17 -** Code errors more likely to impact more people than individual actions by officers
	• **R.18** – Speed of automation likely to impact more customers in a shorter timeframe than manual processing
	• **R.19** – Errors may take longer to detect and correct than manual processing
	• **R.20** – Data errors may lead to unintended consequences
	• **R.21** – Training or testing data may lead to inbuilt bias or unintended consequences
		• **O.16** – Independent assurance processes to promote appropriate and balanced use of ADM
		• **O.17** – Code errors subject to additional layers of scrutiny and breaking through the digital layer to 'bake in' fairness
		• **O.18** – Speed of automation able to process the representations of more customers in a shorter timeframe than manual processing
		• **O.19** – Opportunity to ensure that batches are subject to checks to ensure that there is requisite oversight for potential errors
		• **O.20** – Opportunity to seek ingestion of better-quality data to ensure better outcomes
		• **O.21** – Training or testing data to be checked by impartial third party, or, alternatively, created without bias
			Overview:
			• **C.16.1** – AI use to be controlled and limited until the removal of bias is possible (if possible at all) and not used to discriminate against customers
			Concept:
			• **C.17.1** – Procurement or design to consider potential for bias – lowest risk system can be selected or created
			• **C.16.2** – Anti-discrimination needs to be embedded into testing, review, monitoring, and evaluation processes and tools
			Discovery:
			• **C.17.2** – Testing of code and systems to ensure that there is a code error assurance program
			• **C.21.1** - Consideration of including an algorithmic assessment tool or similar to allow for self-assessment of current state when creating a new system including ADM
			• **C.17.3** – Review sessions/interviews with those designing, implementing, and operating ADM systems to ensure that there is proper implementation of discretionary decision-making being informed by ADM inputs and discharge due diligence
			Develop:
			• **C.17.4** – Code errors to be examined via rigorous testing to ensure that there is an accounting for all possible and foreseeable failure
			• **C.20.1** – 'Fail safe' to be developed to ensure that, where there is a certain failure or error rate, that there will be a full review of the process and the implemented ADM
			Operate, Monitor, and Iterate:
			• **C.16.3** – Ensuring people are accountable for decisions to be made, and that people understand the concept of cognitive bias in dealing with matters
			• **C.18.1** – Speed of automation can be controlled through small batch processes with 100% QA applied prior to full implementation
			Monitoring & Review:
			• **C.16.4** – Introduction of legal testing – independent review of how legal requirements have been interpreted into code, and whenever amendments are made. This needs to cover:
			  ○ Asking right question to get the right answer; and
			  ○ Review whenever legislative amendments or system changes are made
			• **C.17.5** – Risk assessment to guide the need and nature/scope of independent legal audit carried out prior to implementation and after implementation
			• **C.19.1** – Test, monitoring, and review plan – developed and implemented to define types, criteria/KPIs, and frequency of testing & monitoring, as well as alert and escalation processes to ensure any issues are managed as soon as possible
			• **C.20.2 –** Independent review of any reported data errors at a tactical layer an then theme/trend at the operational layer, followed by systemic insights at an organisational layer.
			• **C.21.2** – Validation and accuracy testing: verify and validate machine outputs against human decisions, and external review/challenge or remedy options

Considerations	Risks	Opportunities	Controls for consideration
			Retirement: • **C.21.3** – Evaluation plan to be developed and implemented to carry out evaluation of outcomes and independent assurance of the implementation of controls (as a retrospective) • **C.19.2** – Auditing/assurance completed to allow for fully human decision making independent of ADM to verify appropriate outcomes

Privacy and Security

Considerations	Risks	Opportunities	Controls for consideration

'Privacy, information, and data legislative and other requirements'
 	
Privacy and Personal Information Protection Act 1998 (and associated codes and instruments)
 	
Health Records and Information Privacy Act 2002
 	
Government Information (Public Access) Act 2009
 	
Privacy Act 1988 (Cth)
 	
Data Sharing (Government Sector) Act 2015
	• **R.22** – Disposal of health records may not be rigorous or strong enough, which may result in information being used by AI
	• **R.23** – Completeness of information (i.e. where there are fields that have not been filled in, how will the ADM treat this?)
	• **R.24** – Transparency of privacy information, however, risk of putting out full PIAs and the public finding out more than is reasonable about systems and processes
	• **R.25** – No current limitations on some processes results in retention and possible ADM use of information that should no longer be relevant
	• **R.26** – Consideration of ingestion of sensitive data – what constitutes this and how do we ensure we are not worsening circumstances for vulnerable people
		• **O.22** – Ability to make systems scrutinise information or uplift portals to warn people about entering health information
		• **O.23** – Train ADM to identify fields that have not been filled in or have not been filled in correctly to query or raise for manual oversight
		• **O.24** – Opportunity to share reasonable amounts of information with the public and increase public trust
		• **O.25** – Legislative change able to be exacted to ensure limitation periods are adequately applied
		• **O.26** – Data anonymisation or quarantining of material that has been deemed sensitive to ensure best results
			Overview:
			• **C.22.1** – Collection should not be excessive and 'Privacy by Design' should be considered when building a new process assisted by ADM
			Concept:
			• **C.23.1** – Information fields to be strictly defined and concept to ensure that there is a top-level approach to fields that are required when inputting data
			• **C.24.1** – Mandatory PIAs to be undertaken prior to any project being implemented (even if not consuming new information)
			Discovery:
			• **Nil**
			Develop:
			• **C.22.2 –** System to consider in design whether it will reject certain types of information when attempt is made to input it
			• **C.25.1** – Legislative change to ensure limitation periods being implemented into system solutions
			• **C.25.2 –** Consideration of all IPPs and best practice to fulfil their requirements as laid out in the legislation (see detailed literature review for further information)
			• **C.26.1** – Design a system to identify and triage any sensitive information, including but not limited to the following:
			  ○ Children;
			  ○ Religious individuals;
			  ○ Racially or ethnically diverse individuals;
			  ○ Individuals with political opinions or associations;
			  ○ Individuals with trade union membership or associations;
			  ○ Gender and/or sexually diverse individuals;
			  ○ Individuals with a criminal record;
			  ○ Specific health or genetic information;
			  ○ Personal biometric information; and
			  ○ Other sensitive person-centred data.
			Deliver:
			• **C.23.2 –** Rejection of data could be made automatic where all fields are not filled in
			• **C.25.3** – Notification if an item is approaching a limitation period and has not been culled or requires further review
			• **C.26.2 –** Consideration of the following risk-centred questions in development:
			  ○ Have you applied the 'Privacy by design' and 'security by design' principles in your system?
			  ○ Have you completed a privacy impact assessment (either third party or self-assessed)?

Considerations	Risks	Opportunities	Controls for consideration
			○ If you are using information about individuals who are reasonably identifiable, have you sought consent from citizens about using their data for this particular purpose? ○ Does your system adhere to the mandatory requirements in the NSW Cyber Security Policy? Have you considered end-to-end Security Principles for your system? ○ Does your dataset include using sensitive data subjects as described by section 19 of the NSW Privacy and Personal Information Protection Act 1998? **Launch:** • **Nil** **Operate, Monitor, and Iterate:** • **C.22.3** – Operational assessment and referral of information that may be irrelevant or an 'overcollection' (with the addition of implementing a process whereby redaction can be automated) • **C.23.3 –** Operational reports of any instances where bad or incomplete data has been received or ingested **Monitoring & Review:** • **C.22.4** – Audit and triage of incidents where there has been over-collection of information and data to ensure future changes can be made • **C.24.2** – Privacy impact assessments to be published to some extent – perhaps with redactions or at a high-level following implementation? **Retirement:** • **C.24.2** – Publication of approaches to dealing with retrospectives for retired and ongoing projects (essentially as assurance reports) • **C.22.5** – Self-assessment tool that would allow for review (either at retirement of old projects or concept stage)

Transparency

Considerations	Risks	Opportunities	Controls for consideration
'Good practice for designing and implementing machine technology'	• **R.27** - ADM decisions may be challenged by customers on the basis that they are not 'human' decisions • **R.28** - Loss of community trust if not providing adequate transparency of use of ADM • **R.29** - Increased customer feedback or complaints if all details of ADM provided • **R.30** - If publishing ADM details or code details – increased ability for threat actors to exploit ADM for fraud or general misuse • **R.31** - If details provided on the system, they may lead to discrepancies between information provided on decisions made by ADM vs by people • **R.32** - No trail for ADM assisted decisions – may result in issues when explaining reasoning	• **R.27** - Promotion of transparency as key ethics principle leading to increase in stakeholder and customer trust • **R.28** - Opportunity to communicate to customer improved outcomes as a result of ADM – e.g. 'Your transaction response has been improved by XX days' or '% of efficiency improved' by ADM • **R.29** - Opportunity to provide 'plain English' reasons or explanations to support decisions • **R.30** - Publishing source code or draft specifications can lead to input from stakeholders and improvement of code and outcomes • **R.31** - System details being published improves transparency and allows for an increase in public trust • **R**.32 – Opportunity to introduce an approach that fully encompasses process of ADM assisted decisions	**Overview:** ○ **C.27.1** - Strong audit trail to be incorporated into all Revenue NSW projects – evidence and trail of a decision to be clear for all projects ○ **C.28.1** - Risk assessment to consider: ○ Have you consulted with the relevant community that will benefit from (or be impacted by) the system? ○ Are the scope and the goals of the project publicly available, and have you communicated how safeguards have been put in place to mitigate any potential harms (explore diverse approaches to instil confidence within communities regarding your AI utilisation. This may entail targeted communication strategies or maintaining public registers. Offer concise and straightforward explanations of your AI usage to those potentially affected, especially for elevated risk. Ensure these explanations foster trust without generating confusion)? ○ Is there an easy and cost-effective way for people to appeal a decision that has been informed by your system (individuals have the right to raise concerns or appeal decisions. Ensure the use of simple and easily understandable language to facilitate this process)? ○ Does the AI system allow for transparent explanation of the factors leading to a decision or insight? **Concept:** • **C.31.1** - Early-stage transparency provisions if considering a third-party engagement with the design to ensure procurement specifications are covered. Alternatively, consider including GIPA requirements with 3rd party as part of procurement and contract development and management. **Discovery:** • **Nil** **Develop:** • **C.30.1** - Build into design centralised monitoring tool or register to cover key assurance activity, error, or other tolerance criteria, including reporting on accuracy or success rate, feedback loops (including customer complaints etc.) • **C.32.1** - Ensuring that there are no 'black box systems' in development • **C.29.1** - Consideration of the 'plain English' explanations to ensure ADM is adequately supported with the decisions made (e.g. 'path of reasoning') • **C.29.2** - Reasoning should be developed to ensure that people can understand how and why the decision was made **Deliver:** • **C.32.2** - Register or other document to define upstream and downstream actions/workflows • **C.29.3** - Standard communication/transparency clauses to be included in processes utilising ADM where appropriate – a 'Statement of Reasons' to be developed by the human administrator utilising the system for ADM.

Considerations	Risks	Opportunities	Controls for consideration
			• **C.28.2** - 'Hearing rule' to be embedded into correspondence by ensuring a right of response, review, or complaint are defined and communicated to customers at key stages in the process. Accountability (i.e. to a human) needs to be included in this approach as well **Launch:** • **Nil** **Operate, Monitor, and Iterate:** **Monitoring & Review:** • **C.32.3** - A register of ADM systems with version history, including descriptions of changes, and reasons for changes. Alternatively, other documentation could be considered to define code so it is available for oversight • **C.30.2** - Transparency mechanism, including details of feedback loops, and how these things are considered in the development and change management of ADM • **C.29.4** - Customer consultation on decisions and feedback **Retirement:** • **C.32.4** - Proactive review of ADM/AI/Automation on RNSW website under GIPA (Open Access Information – suggested by NSW Information Commissioner) • **C.28.3** - Partial release of information on retirement without adverse disclosure
'Government agency use – transparency and visibility'	• **R.33** - Lack of understanding and consolidated view/line of sight of use and appropriate governance and risk management • **R.34** - Lack of understanding and capability in the sector and regulatory agencies to support balanced and appropriate use of technology and oversight – both internal and external	• **O.33** - Opportunity to work with the Ombudsman and other regulators and policy bodies to provide greater understanding of benefits, risks, process controls, monitoring & assurance to increase trust • **O.34** - Opportunity to include independent members/oversight embedded into governance in a proactive way	**Overview:** **Concept:** **Discovery:** • **C.33** - Consideration for how and when to engage with regulators to enable input and oversight, while growing capability and joint understanding **Develop:** **Deliver:** **Launch:** • **Nil** **Operate, Monitor, and Iterate:** • C.34 - Use cases or examples to explain the types of ADM, AI, and Automation that is used, governance, risk management, monitoring & assurance, and oversight of them **Retirement:**

Accountability

Considerations	Risks	Opportunities	Controls for consideration

Legislative reform - *'The Role of Parliament in authorising machine technology'*	• **R.35** - Costs and resources needed to govern and monitor technology are not embedded into costings and budgets
	• **R.36** - No capability to govern, monitor, and provide assurance over technologies
	• **R.37** - Code writers or projects may not understand or may not have the skills to adequately interpret accountability requirements into code
	• **R**.38 – Relevant responsible individuals for decisions may be obscured by ADM processes
	• **R**.39 – Relevant government code of conduct and legislation may not be applied (including Integrity, Service, Trust, Accountability)
		• **O.35** – Opportunity to introduce a costs model that considers technology at the forefront
		• **O.36** – Assurance framework can be inbuilt to new ADM systems from day one
		• **O.37** – Training and upskilling should be made available to code writers or projects
		• **O.38** – Opportunity for audit trail to be embedded in all ADM processes and other processes
		• **O**.39 – Refresher on code of conduct and higher level education on what this means when combined with AI and ADM
			Overview:
			• **C.36.1** - Team of multi-disciplinary SMTs to be developed to involve lawyers, policy makers, and operational experts, as well as technicians, with roles and responsibilities that are clearly defined (both for implementation and review)
			• **C.37.1** - Independent advisors to be included in the loop, as well as change monitoring and reporting
			• **C.36.2 Risk assessment to consider the following**:
			  ○ Have you established who is responsible for:
			  ○ use of the AI outputs, insights, and decisions?
			  ○ policy/outcomes associated with the AI systems?
			  ○ monitoring the performance of the AI system?
			  ○ data governance?
			  ○ technical solution governance?
			  ○ appeal and redress process?
			  ○ Have you established a clear process to intervene if a relevant stakeholder finds concerns with insights, decisions, or content generated (appeal and redress)? Ensure you do not get overconfident or over reliant on the AI system?
			  ○ Consider if you are buying an AI system (high supplier control, low agency control), embedding an AI system (or co-train – shared supplier and agency control), or developing an AI system (or train – no supplier control, all agency control).
			Concept:
			• **C.35.1 -** Costing and resource models need to consider resources needed for oversight, testing, monitoring, and assurance
			• **C.35.2** – Prior to implementation, heed must be given to:
			  ○ **Large projects** – any project or service using AI AND DRF funded OR total cost >$5million = self assess against AIAF and then reviewed by a NSW AI review body.
			  ○ Small projects – any project or service using AI AND is funded from agency budget and cost <$5million = self assess against AIAF AND seek approval from an appropriate senior officer in the agency.
			• **C.39.1** – Risk analysis to include code of conduct and if public sector values can be adequately applied.
			Discovery:
			• **C.36.3** - When building on a project, consideration of bias, data, and inputs in the development of technology to ensure accountability
			• C.37.2 - Bias in code or rules to be identified and assurance/review mechanisms to be applied to this as required
			Develop:

Considerations	Risks	Opportunities	Controls for consideration
			• **C.38.1** - Built-in parameters around the appropriate level of human involvement in the decision-making process, including having regard to the function and legislation in question **Deliver:** ○ **C.38.2** - Escalations to be examined, reported, and heard and accounted for when reported **Launch:** • **Nil** **Operate, Monitor, and Iterate:** ○ **C.38.3** - Ensuring that people have adequate training, understanding, and mechanisms to escalate and review ○ **C.36.4** - Records of system mechanisms and 'trails' for decision making **Retirement:** ○ **C.35.3** – Review of costs structure and any future proofing required in other projects