About the AI Assurance Framework



What is it?

The Framework will help you design, build and use AI technology appropriately. It contains questions that you will need to answer at every stage of your project and while you are operating an AI system. If you cannot answer the questions, the framework will let you know how to get help.

The aim of the Framework is to support the ACT Government innovate with AI technology while ensuring that we remain safe and secure and that there is clear accountability for the design and usage of AI systems.

Who should use it?

The Framework is intended to be used by:

- project teams who are using AI systems in their solutions
- operational teams who are managing AI systems
- Responsible officers who are accountable for the design and use of AI systems
- internal assessors conducting agency self-assessments
- the Al Advisory Group

When should I use it?

All AI systems and projects must be assessed against the framework throughout all stages of an AI project, from the initial planning to the final delivery. Regular reviews should also be conducted to review services that use AI systems.

Is applying this framework everything I need to do?

The framework is not a complete list of all requirements for AI projects. Project teams should comply with their directorate-specific AI processes, policy requirements and governance mechanisms as well.

Before you start

You must read this before you start any AI project:

ACT Government Artificial Intelligence Policy

When you do not need to apply this framework

You do not need to assess your product or service if:

- you are using an AI system that is a widely available commercial application, and
- you are not using ACT Government data to maintain, train or retrain the AI system's underlying model(s).

Examples: personal digital assistant, smart phones, smart watches, laptops, QR code reader, satnav system, smart card reader, smoke detector, digital thermometer.

Al systems developed on commercially available software platforms are not exempt.

How to conduct an Al assurance assessment



1. Assess risk factors

Consider and determine the risk factors for your AI or data driven project using the risk metrics in the Framework.

2. Answer questions & document reasons

Consider and capture your responses to the questions in the Framework. Decide about whether your project should:

- continue as-is
- continue with additional treatments
- stop.

Consider that any information you capture may be subject to Freedom of Information Act or public disclosure.

3. Self-assess or submit to the Al Advisory Group

Figure 1. The steps to conduct an AI assurance assessment.

This assessment is to be completed by (or the result confirmed with) the responsible officers.

Responsible Officers	Responsibilities	
Al System Owner	 Responsible for the AI insights / decisions Define AI system strategy and alignment with goals and deliverables. Ensure AI system meets framework requirements. 	
AI System Administrator	 Manage technical performance of the AI system. Monitor AI system performance and handle updates or changes. 	
Al System Data Custodian/Steward	 Responsible for data used in an AI system. Ensure data governance and management requirements are met. 	
AI System Project Managers	 Responsible for outcomes of the project (during development/upgrade of system) Manage the AI system project scope, goals, and deliverables. 	

These four roles have independent responsibilities and must not be held by the same person.

The responsible officers should be appropriately senior, skilled and qualified for the role.

When to submit your project to the Al Advisory Group



The Policy and Framework apply to *all* AI projects that:

- Use AI in commercially available products in new and novel ways, for example the solution is tailored for a new use specifically for the ACT Government (excludes configuration),
- Use AI solutions specifically developed for the ACT public service, internally or by external vendors, or
- Use <u>generative AI</u> capabilities, even if these capabilities are part of standard commercially available products and are not modified.

At the end of the self-assessment, the template will assign a risk rating (highest risk and total number of risks ranked medium or higher) to the different principles in your AI project. This rating will determine if your project should proceed as is (a score of 'low risk') or be submitted to the AI Advisory Group for consideration (a score of 'medium' or 'high' risk).

Evaluating AI benefits and risks



Benefits and risks

The ACT Government has a strong commitment to wellbeing and embracing emerging technologies.

This means you need to evaluate the potential risks of harms from deployment and operation of AI, as well as its benefits.

Currently, we use AI or data driven tools to:

- deliver insights that improve services and lives
- help directorates work more quickly and accurately

While there are many areas where AI can benefit the work we do, we need to engage with risks early and throughout the life lifecycle of the technology.

Cannot answer some questions?

It is important to make a note of questions you cannot answer as you progress through the assessment. It may be because information is not available or can only be answered once a pilot is undertaken.

If the project proceeds, treat these unanswered questions as representing midrange risk, commence with a pilot phase and closely monitor for harms and establish controls.

Evaluating and engaging with risk

This AI Assurance Framework is structured in sections that align to the eight Australian AI Ethics Principles.

Each section starts with a page that prompts you to consider the types of risk that your project may carry and helps shape your response to questions in that section with risk in mind. Refer to the ACT Insurance Agency (ACTIA) risk matrix and guidance for further information.

At the end of the self-assessment, you will assign a risk rating (highest risk and total number of risks ranked medium or higher) to the different Ethics Principles in your AI project. This rating will determine if your project should:

- proceed as is
- proceed, with additional risk mitigations
- stop.

Understanding the balance of benefits and risks

Some projects carry real risk, but are undertaken to improve existing processes, or because of a clear benefit to community.

Identifying and managing of these risks during the life of the project is an essential requirement, as is clarifying the benefits of the project.

Operational vs non-operational Al

ACT Government

Operational AI

Operational AI systems are those that have a real-world effect. The purpose is to generate an action, either prompting a human to act, or the system acting by itself. Operational AI systems often work in real time (or near real time) using a live environment for their source data.

Not all operational AI systems are high risk. An example of lower risk operational AI is the digital information boards that show the time of arrival of the next bus.

Operational AI that uses real-time data to recommend or make a decision that adversely impacts a human will likely be considered High or Very high risk.

Non-operational Al

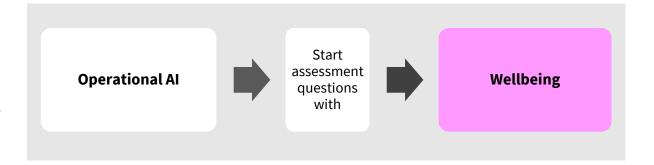
Non-operational AI systems do not use a live environment for their source data. Most frequently, they produce analysis and insight from historical data.

Non-operational AI often generate insights for policy or program design. However, this is considered a 'factor' in a decision-making process (made by a user), rather than an impactful decision made by an algorithm.

Non-operational AI typically represents a lower level of risk. However, the risk level needs to be carefully and consciously determined, especially where there is a possibility that AI insights and outputs may be used to influence important future policy positions.

Benefits identification

For all AI systems, the benefits of the AI project should be captured in a <u>Benefits</u> Realisation Plan before commencement.





Al risk factors exist on a spectrum



The key factor that determines risk is how the AI system is used, including whether it has an operational impact, or if it is not transparent or not explainable.

Very low risk
or N/A

Low
Midrange
High
high risk

Al generates insights for non-operational human use from non-sensitive data. For example, analytics package reporting on historical non-sensitive data. Al generates insights or alerts for operational human use with minimal potential for harm. For example, anomaly detection software; alarm system).

Al makes operational actions, decisions, or recommendations with no routine human oversight with minimal potential for harm. For example, automated door; biometric login with alternative login methods; automated phone menu, smart sign showing driver speed).

Al generates operational insights, decisions, or recommendations for humans to action with some potential for harm. For example, public facing chatbot, red light camera, intruder alert system.

Al generates insights for non-operational human use analysing sensitive data. For example, analytics package operating on data of vulnerable individuals. Al makes and implements operational decisions autonomously of human input in the interests of human safety and wellbeing. For example, anti-lock braking system.

Al makes and implements operational decisions within a specified range and refers exceptions for human to review and action. For example, loan application system, autonomous tram.

Al makes and implements operational decisions that can negatively affect human wellbeing without human input. For example, autonomous benefits eligibility reviews, judicial custodial sentence recommendations, unconstrained autonomous system and self-driving car.

Ethics framework



Mandatory Principles

There are eight principles that you must apply when using AI. These are mandated through the ACT Government AI Policy.

Human, societal, and environmental wellbeing

Al systems should benefit individuals, society, and the environment.

Human-centred values

Al systems should respect human rights, diversity, and the autonomy of individuals.

Fairness

Al systems should be inclusive and accessible and should not involve or result in unfair discrimination against individuals, communities or groups.

Reliability and safety

Al systems should reliably operate in accordance with their intended purpose.

Privacy protection and security

Al systems should respect and uphold privacy rights and data protection and ensure the security of data.

Accountability

People responsible for the different phases of the AI system lifecycle should be identifiable and accountable for the outcomes of the AI systems, and human oversight of AI systems should be enabled.

Transparency and explainability

There should be transparency and responsible disclosure so people can understand when they are being significantly impacted by AI and can find out when an AI system is engaging with them.

Contestability

When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or outcomes of the AI system.

More information can be found in the ACT Government Artificial Intelligence Policy. You must consider and apply this Policy when designing, implementing or running an AI System.

Wellbeing: Benefits



Consider the benefits associated with	None	Low	Medium	High	Very high
Delivering a better-quality existing service or outcome					
Reducing processing or delivery times					
Generating financial efficiencies or savings					
Providing an AI capability that could be used or adapted by other agencies					
Delivering a new service or outcome					
Enabling future innovations to existing services, or new services or outcomes					

Wellbeing: Risk Factors



Consider the risks associated with	None	Low	Medium	High	Very high
Delivering a better-quality existing service or outcome					
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 sufficient experienced human 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					

Al: Risk Factors



Consider the risks associated with	Very low risk None	Low Reversible at low cost	Medium Reversible at medium cost	High	Very high Irreversible
Physical harms					
Psychological harms					
Environmental harms or harms to the broader community					
Unauthorised use of health or sensitive personal information (SIP)					
Impact on right, privilege or entitlement					
Unintended identification or misidentification of an individual					

Al: Risk Factors continued



Consider the risks associated with	None	Low	Medium	High	Very high
Misapplication of a fine or penalty					
Other financial or commercial impact					
Incorrect advice or guidance					
Inconvenience or delay					

Other harms

Human, Societal and Environmental Wellbeing



Will your AI system contribute to improve the wellbeing domains from the ACT Wellbeing Framework?

•	Yes	Document your response and go to the next question.
•	No	Pause the project, consider how will the project affect wellbeing.

110	r dase the project, consider now with the project uncer wellseling.			
Click on the text to provide a response.				
otick of the text to provide a response.				

All AI projects should maintain a benefits register throughout the development of the project.

Any important highlights from the wellbeing framework.

Human, Societal and Environmental Wellbeing



Will the AI system improve on existing approaches to deliver the outcomes described in:

- a relevant framework, or
- your agency's strategic plans?

•	Yes	Document your response and go to the next question.
•	No	Pause the project, consider how will the project improve on existing approaches.

Click on the text to provide a response.

All AI projects should maintain a benefits register throughout the development of the project.

Human, Societal and Environmental Wellbeing



Has the environmental impact, such as carbon emissions created by computing power, been considered when planning the proposed AI system?

•	Yes	Document your response and go to the next question.
•	No	Pause the project, consider how will the project impact the environment.

Click on the text to provide a response.		

To learn more about the environmental implications, including emissions associated with AI systems, please consult with the Environment, Planning, and Sustainable Development Directorate (EPSDD).



Does this project and use of data align with relevant legislation?

- ACT Human Rights Act
- ACT Information Privacy Act
- ACT Public Sector Management Act

•	Yes	Document your response and go to the next question.
•	No	Pause the project.

Click on the text to provide a response.



Were other, non-AI systems considered?

•	Yes	cument your response and go to the next question.	
•	No	Pause the project.	
Informally After your pilot you must conduct a formal benefits review before scaling the project. Document your response.		After your pilot you must conduct a formal benefits review before scaling the project. Document your response.	

Click on the text to provide a response.	

For an AI project to be viable, AI must be the most appropriate system for your service delivery or policy problem.

Al systems can come with more risk and cost than traditional tools. You should use an Al system when it is the best system to maximise the benefit for the customer and the government.



Have you designed the AI system with stakeholder input?

•	Yes	cument your response and go to the next question.	
•	No	Pause the project.	

Click on the text to provide a response.	



Have you considered the impact of the AI system on ACTPS employees?

•	Yes	ocument your response and go to the next question.	
•	No	Pause the project.	

Click on the text to provide a response.

If this AI system will fundamentally change jobs or may result in the displacement of employees, genuine consultation with staff and unions must occur. To ensure effective consultation on workplace matters, follow the Consultation clause of the Enterprise Agreement.

Fairness: Risk Factors

to the data for intended use



Consider the risks associated with	Very low risk or N/A	Low	Medium	High	Very high risk
Using incomplete or inaccurate data					
Having poorly defined descriptions and indicators of "Fairness"					
Not ensuring ongoing monitoring of Fairness indicators					
Decisions to exclude outlier data					
Informal or inconsistent data cleansing and repair protocols and processes					
Using informal bias detection methods (best practice includes automated testing)					
The likelihood that re-running scenarios could produce different results					
Inadvertently creating new associations when linking data or metadata					
Differences in the data used for training compared					



Can you explain why you selected this data for your project and not others?

Consider how your data selection may impact the fairness or equity or your model.

•	Yes	Document your response and go to the next question.	
	 Unclear Consult with relevant stakeholders to identify alternative data sources or implement a data imp 		Consult with relevant stakeholders to identify alternative data sources or implement a data improvement strategy or redesign the
	Officieal	project.	
•	It's better than existing systems	Document your reasons. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to	
		pilot phase.	
•	No	Pause the project and consider how absent data or poor-quality data will impact your system.	

Cl	lick	on	the	text t	:0 r	provide	e a	res	ponse.	
•					.~ r	J. O	_ ~		P 0 C C .	

Your AI system may draw from multiple data sources to find new patterns and insights.

You need to determine if you can/should use the data for the AI system. This may be challenging for historical data that was collected for a different purpose.



Is the data that you need for this project available and of appropriate quality given the potential harms identified?

If your AI project is a data creation or data cleansing application answer according to the availability of any existing data that is needed for the project to succeed, for example, training datasets.

•	Yes Document your response and go to the next question.	
•	Unclear	Consult with relevant stakeholders to identify alternative data sources or implement a data improvement strategy or redesign the
	Officiear	project.
•	It's better than existing systems	Document your reasons. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to
		pilot phase.
•	No	Pause the project and consider how absent data or poor-quality data will impact your system.

Click on the text to provide a response.

Data quality is often described in terms of minimum requirements for accuracy, timeliness, completeness and consistency. Your AI system may be significantly impacted by poor quality data. It is important to understand how significant the impact is before relying on insights or decisions generated by the AI system. Absences of data may lead to unintended biases impacting insights generated by the AI system. Unbalanced data is a common problem when training AI systems.



Does your data reflect the population that will be impacted by your project or service?

•	Yes Document your response and go to the next question.	
	It's bottor than existing systems	You may need to seek advice from an ethics committee. You should clearly demonstrate that you have consulted with all relevant
	It's better than existing systems	stakeholders before proceeding to the pilot phase.
No or unclear Pause the project and address the gaps in your solution design.		Pause the project and address the gaps in your solution design.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

lick on the text to provide a response.		



Have you considered how your AI system will address issues of diversity and inclusion (including geographic diversity)?

Have you considered the impact to gender and minority groups including how the solution might impact different individuals in minority groups when developing this AI system?

Minority groups may include: – Those with a disability – LGBTIQA+ and gender fluid communities – Cultural and Linguistically Diverse groups – Aboriginal and Torres Strait Islander Peoples – Children and young people

•	Yes	Document your response and go to the next question.	
	It's better than existing systems	You may need to seek advice from an ethics committee. You should clearly demonstrate that you have consulted with all relevant	
		stakeholders before proceeding to the pilot phase. Consider a Human Rights Assessment.	
•	No or unclear Pause the project and address the gaps in your solution design.		
•	N/A	Document your reasons as to why this does not apply, then go to the next question.	

Click on the text to provide a response.

Services or decisions can impact different members of the relevant community in different ways.

Whether due to cultural sensitivities, or underrepresentation in training data sets. It is important to think deeply about everyone who might be impacted by AI systems.



Do you have appropriate fairness objective performance measures and targets?

Aspects of accuracy and precision are readily quantifiable for most systems which predict or classify outcomes. This performance can be absolute or relative to existing systems.

How would you characterise "Fairness" such as equity, respect, justice, in outcomes from an AI system? Which of these relate to, or are impacted using AI?

•	Yes	Document your response and go to the next question.
• No	No or uncloar	For operational AI systems pause the project until you have established performance measures and targets.
	No or unclear	For non-operational systems, results should be treated as indicative and not relied on.
N/A Document your reasons as to why this does not apply, then go to the next question.		

Click on the text to provide a response.

Good fairness metrics depends on the types of AI model(s) deployed in your AI system. Do measures like Accuracy, Precision, Sensitivity, and Specificity change significantly between cohorts.

Click on the text to provide a response.



Do you have a way to monitor and calibrate the fairness of your AI system?

Operational AI systems which are continuously updated / trained can quickly move outside of performance thresholds. Supervisory systems can monitor system performance and alert when calibration is needed.

•	Yes Document your response and go to the next question.	
	No or unclear	For operational AI systems pause the project until you have established performance measures and targets.
		For non-operational systems, results should be treated as indicative and not relied on.
•	N/A Document your reasons as to why this does not apply, then go to the next question.	

Operational AI systems should have clear performance monitoring and calibration schedules.

For operational AI systems which are continuously training and adapting with moderate residual risks, weekly performance monitoring and calibration is recommended.

For operational systems with high risk or very high risk, a custom evaluation and calibration will be required.

Reliability and Safety



Do you have appropriate performance measures for the reliability of your system?

Aspects of accuracy and precision are readily quantifiable for most systems which predict or classify outcomes. This performance can be absolute or relative to existing systems.

•	Yes Document your performance measures and go to the next question.	
	No or unclear	For operational AI systems pause the project until you have established performance measures and targets.
		For non-operational systems, results should be treated as indicative and not relied on.
•	N/A Document your reasons as to why this does not apply, then go to the next question.	

Click on the text to provide a response.			

Reliability and Safety

Click on the text to provide a response.



Do you have a way to monitor and calibrate the safety of your AI system?

Operational AI systems which are continuously updated / trained can quickly move outside of performance thresholds. Supervisory systems can monitor system performance and alert when calibration is needed.

•	Yes	Document your response and go to the next question.
• 1	No or unclear	For operational AI systems pause the project until you have established performance measures and targets.
		For non-operational systems, results should be treated as indicative and not relied on.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Operational AI systems should have clear performance monitoring and calibration schedules.

For operational AI systems which are continuously training and adapting with moderate residual risks, weekly performance monitoring and calibration is recommended.

For operational systems with high risk or very high risk, a custom evaluation and calibration will be required.

Reliability and Safety



Considering the planned mitigations, could the AI system cause significant or irreversible harms?

Operational AI systems which are continuously updated / trained can quickly move outside of performance thresholds. Supervisory systems can monitor system performance and alert when calibration is needed.

•	Yes Document your response and go to the next question.	
•	No or unclear	For operational AI systems pause the project until you have established performance measures and targets.
		For non-operational systems, results should be treated as indicative and not relied on.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Click on the text to provide a response.			



Have you embedded privacy and security into the design process of the system?

•	Yes	Document your response and go to the next question.	
•	No	se the project, review the project design.	

Click on the text to provide a response.			
•			



Have you completed a privacy impact assessment (either third party or self-assessed)?

•	Yes	Document your response and go to the next question.	
•	No	Pause the project.	
•	N/A	ocument why a PIA is not required.	

Click on the text to provide a response.		



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?

•	Yes	Document your response and go to the next question.
•	No	Pause the project.

Click on the text to provide a response.	



Does your AI System adhere to the mandatory requirements in the ACT Cyber Security Policy?

•	Yes	Document your response and go to the next question.
•	No	Pause the project.

110	r duse the project.				
Click on the text to provide a response.	Click on the text to provide a response				
click off the text to provide a response.					
I and the second					



Does your dataset include using sensitive data subjects?

•	Yes	Document why you need this data and consider seeking legal advice.
•	No	Document your response and go to the next question.

	<u> </u>	•	
Click on the text to provide a response.			
click off the text to provide a response.			

Accountability: Risk Factors



Consider the risks associated with	Very low risk or N/A	Low	Medium	High	Very high risk
Insufficient training of AI system operators					
Insufficient awareness of system limitations of Responsible Officers					
No or low documentation of performance targets or "Fairness" trade-offs					
No or limited mechanisms to record insight / Al System decision history					
The inability of third parties to accurately audit Al system insights/ decisions					

Accountability



Have you established who is responsible for:

- Use of the AI insights and decisions
- Policy/outcomes associated with the AI system
- Monitoring the performance of the Al system
- Data Governance

•	Yes, or N/A	Document your response and go to the next question.
•	No	Pause the project, review the project design.

Click on the text to provide a response.

Accountability



Have you established processes to:

- Intervene if a relevant stakeholder finds concerns with insights or decisions?
- Ensure you do not get overconfident or over reliant on the AI system?

•	Yes, or N/A	Document your response and go to the next question.
•	No	Pause the project, review the project design.

Click on the text to provide a response.

For operational AI systems, you must make sure that humans are accountable and can intervene. This may also apply to non-operational AI systems

Transparency: Risk Factors



Consider the risks associated with	Very low risk or N/A	Low	Medium	High	Very high risk
Incomplete documentation of AI system design, or implementation, or operation					
No or limited access to model's internal workings or source code					
Being unable to explain the output of a complex model					
A member of the public being unaware that they are interacting with an AI system					
No or low ability to incorporate user feedback into an AI system or model					



Have you consulted with the relevant community that will benefit from (or be impacted by) the AI system?

•	Yes	Document your response and go to the next question.		
		For AI systems intended to operate under legislation which allows use without community consultation, do not proceed unless you		
•	Authorised use	receive clear legal advice that allows this project to proceed. The project should be carefully monitored for harms during the pilot		
		phase.		
	. It's best outless suisting a set one	You may need to seek advice from an ethics committee. Document your reasons. You should clearly demonstrate that you have		
	It's better than existing systems	consulted with all relevant stakeholders before proceeding to pilot phase.		
•	Pause the project, develop a Community Engagement Plan and consult with the relevant community.			
•	N/A	Document your reasons as to why this does not apply, then go to the next question.		

Click on the text to provide a response.

You must consult with the relevant community when you design your AI system. This is particularly important for operational AI systems.

Communities have the right to influence government decision-making where those decisions, and the data on which they are based, will have an impact on them. For AI systems intended to operate under legislation which allows use without community consultation, the public benefits must be clear before proceeding to pilot phase.



Are the scope and goals of the project publicly available?

•	Yes	Document your response and go to the next question.
•	No	Make sure you communicate the scope and goals of the project to relevant stakeholders and the relevant community who are impacted before proceeding beyond pilot.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Click on the text to provide a response.

The ACT Government recognises we have important work to do to encourage public trust in AI, by ensuring Government is transparent and accountable, and that AI delivers positive outcomes to citizens.



Is there an easy and cost-effective way for people to appeal a decision that has been informed by your AI system?

•	Yes	Document your response and go to the next question.
•	No	Pause your project, consult with relevant stakeholders and establish an appeals process.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Click on the text to provide a response.						

Contestability – No person should ever lose a right, privilege or entitlement without right of appeal.

A basic requirement of Transparency is for an individual affected by a relevant decision to understand the basis of the decision, and to be able to effectively challenge it on the merits and/or if the decision was unlawful.

When planning your project, you must make sure no person could lose a right, privilege or entitlement without access to a review process or an effective way to challenge an AI generated or informed decision.

Click on the text to provide a response



Does the system using the AI allow for transparent explanations of the factors leading to the AI decision or insight?

•	Yes	Document your response and go to the next question.
•	No, but a person makes the final	Consult with relevant stakeholders and establish a process to readily reverse any decision or action made by the AI system.
	decision	Actively monitor for potential harms during the pilot phase.
•	No	Pause your project, consult with relevant stakeholders and establish an appeals process.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Explainability – As far as possible, you must have a way to clearly explain how a decision or outcome has been informed by AI.

If the system is a "black box" due to lack of access to the inner workings or is too complex to reasonably explain the factors leading to the insight generation, it is essential to consider the role of human judgement in intervening before an AI generated insight is acted on. It is important to formalise and document this human oversight process.

In low (or very low) risk environments, it may be sufficient to identify and document mechanisms to readily reverse any action arising from such an insight (e.g. a person overriding an automated barrier).

Contestability: Risk Factors



Consider the risks associated with	Very low risk or N/A	Low	Medium	High	Very high risk
Ensuring documentation is available for the AI system, allowing public review and challenge of its design, implementation, and operations					
No or limited access to the AI model's internal workings and source code to contest the challenges posed by "Black Box" systems.					
A member of the public understanding the AI system outputs to challenge decisions					
Ensuring people are aware when interacting with the AI system and that they can contest decisions					
No or low ability to modify the AI system in response to a contest made by members of the public					

Is a 'black box' system use, such as a large language model or generative AI, automatically high risk?

The inner workings of commercial AI systems and data driven tools are not always accessible and even if they are, they can be very complex to interpret. To address the risks this poses, think proactively about the role of human judgement in use of an "unexplainable" insight or decision. If you cannot explain the ways in which insights are outputted from an AI system, what are the potential harms that may arise? What's the likelihood of these harms and how readily they can be reversed? It is important that these considerations are documented. This is particularly important if midrange or higher risks are identified.

Comments: these responses should be considered as residual risks after mitigations are in place



Are there any channels or platforms available for members of the public to raise concerns or contest decisions?

•	Yes	Document your response and go to the next question.
•	No	Pause your project, consult with relevant stakeholders and establish a process.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Click on the text to provide a response.						

Contestability mechanisms – The technical complexity of AI systems can make it difficult for the general public to fully understand and contest their decision making.

When planning your project, providing clear and accessible mechanisms for individuals to contest and sort out any harmful or outcomes generated by AI decisions is essential.



Is the public informed about their rights to contest, who to contact, and the process for doing so?

•	Yes	Document your response and go to the next question.
•	No	Pause your project, develop accessible information on contestation rights and processes and plan for its distribution.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Click on the text to provide a response.						

Roles and responsibilities – The lack of transparency makes it hard for individuals to contest or challenge the decisions of the AI system.

When planning your project, establishing clear roles and responsibilities is essential to ensure contestability remains effective throughout the AI system's lifecycle.



Is there a clear process for members of the public to contest AI decisions?

•	Yes	Document your response and go to the next question.
•	No	Pause your project, consider establishing a formal process and document steps to implement.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Click on the text to provide a response.						

Develop guidance – Require your AI system to be designed with transparency and explainability, so that the AI system's internal working can be clearly communicated and understood.

Click on the text to provide a response



Are there any available educational resources about the AI system and how to contest decisions?

•	Yes	Document your response and go to the next question.
•	No	Pause your project, develop and disseminate resources to help people understand the system and their rights.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

offer on the text to provide a response.							

Develop guidance – Require your AI system to be designed with transparency and explainability, so that the AI system's internal working can be clearly communicated and understood.

- Clarify the lines of accountability for the AI system, including roles and responsibilities.
- Explain the process for submitting complaints and accessing dispute resolutions channels.
- Acknowledge the inherent complexity of AI and the challenges in making it fully transparent and explainable.



Will contested decisions be documented and used to improve the AI system over time?

•	Yes	Document your response and go to the next question.
•	No	Pause your project, establish documentation and feedback loop processes for contested decision to guide system enhancements.
•	N/A	Document your reasons as to why this does not apply, then go to the next question.

Click on the text to provide a response.							

Adaptation – Inform the public how the AI system will be continuously monitored and adapted over time.