Australian Government | GovAI

Risk Foresighting Tool

Supporting the identification of emerging and under-considered risks

Department of Finance

Australian Government | GovAI

Executive Summary

Problem Faced

Government agencies are expected to respond to new and emerging risks, but current methods of identifying these risks often rely on manual review of large volumes of complex information. This can make it difficult to detect risks early, particularly when comparing insights across sectors or jurisdictions. Risk information is often siloed, inconsistently reported, and not easily synthesised at scale.

As a result, important signals may be missed, and policy responses may lag behind developments in the broader environment. There is a need for a more systematic, scalable, and timely approach to identifying emerging risks that matter to the Commonwealth and compare these with those already being considered.

Solution Overview

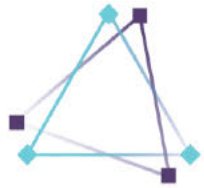
The Risk Policy Team at the Department of Finance developed a generative AI workflow that extracts and categorises risks from two sets of publicly available documents: Commonwealth sources (e.g. corporate plans, risk registers) and non-Commonwealth sources (e.g. industry reports, media articles).

Using structured frameworks such as PESTLE and incorporating human review at key stages, the tool compared the risks identified from each source to highlight where Commonwealth documentation did not reflect sector-based or emerging issues.

Benefits and Impact

The tool can improve APS risk identification by:

- **Efficiency**: Reducing the time required for gap analysis and workload for risk managers.
- **Consistency**: Supports standardised risk reporting by using frameworks like PESTLE.
- **Insight**: Identifies under-considered or missing risks from APS documents, helping agencies address blind spots early.
- **Accessibility**: Makes risk insights available to agencies lacking analytical capabilities.
- **Public value**: Supports more transparent, timely, evidence-based, and risk-informed policy development across government.
- **Informing stakeholders:** Informing the activities of the Commonwealth Risk Committee and providing the tool to Comcover Fund members.



Target Audience and Stakeholders

The tool is designed for use across the APS, including:

- **Policy officers and risk analysts** – to identify and conduct structured assessments of new and emerging risks.
- **Smaller agencies** – to access risk insights without needing dedicated analytic teams.
- **Regulators** – to detect regulatory burdens and support targeted interventions.

Stakeholder engagement has been central to the tool's development. It was tested by the Commonwealth Risk Committee, compared against Comcover benchmarking results, and refined based on Pilot feedback. Broader agency access is being considered.

Risks and Mitigation Overview

The tool is supported by a layered risk management approach:

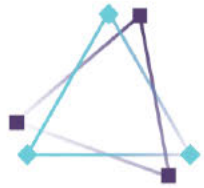
- **Bias** is reduced by drawing on diverse data sources and using AI looping to minimise skewed outputs.
- **Accuracy** is maintained through Human-in-the-loop review of all AI outputs, with each risk linked to its original source.
- **Privacy** is protected by using only publicly available data. No personal information is collected or stored.
- **Security** is ensured through restricted access for authorised APS users using secure authentication.
- **Compliance** is being considered against the DTA AI Assurance Framework and AI in Government Policy.

Use Case Status

Pilot

Use case timeline

Proof-of-concept stage completed. Entering automation stage, whereby a pilot informed by stakeholder engagement will be conducted of the tool. This is to be commenced in July 2025 and completed by August 2025.



Additional Information

The tool uses publicly available data only and does not collect or store personal information. Outputs are reviewed by human users and can be further analysed using the Dragonfly RRR framework, though access to Dragonfly is limited to authorised users.

It is proposed that the tool be hosted on the GovAI platform. Future development may include integration with SharePoint to streamline document ingestion and broader application across sectors.

Lessons Learned

- AI can help surface overlooked risks and improve whole-of-government awareness.
- Human review remains essential to ensure quality outputs.
- User testing highlighted the value of clear documentation and intuitive workflows.
- Model selection was a key consideration- Gemini 2.0 Flash was preferred for its ability to handle large volumes of information, while ChatGPT 4o was effective for categorisation.
- There are trade-offs between building automated workflows and developing a standalone application.

Contact information

Responsible Entity Name

Department of Finance

Open for Collaboration?

Yes

Area of Entity

Risk, Claims and Regulatory Reform Division, Risk and Insurance Branch, Risk Policy Team

Use Case Contact

s 22 [redacted]
Risk Policy Team,
Department of Finance
s 22 [redacted]

Use Case Website/s

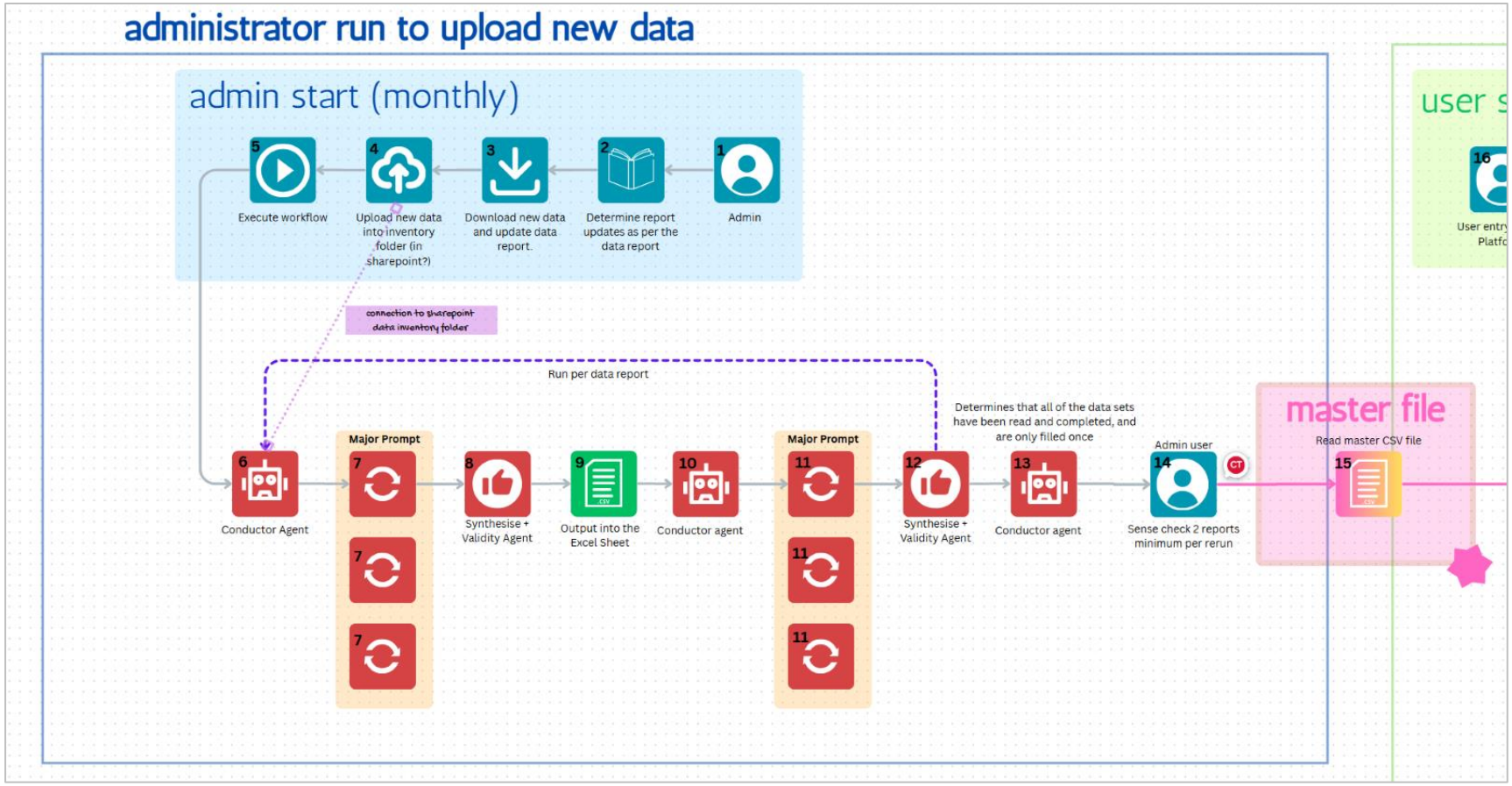
N/A

Use Case Owner

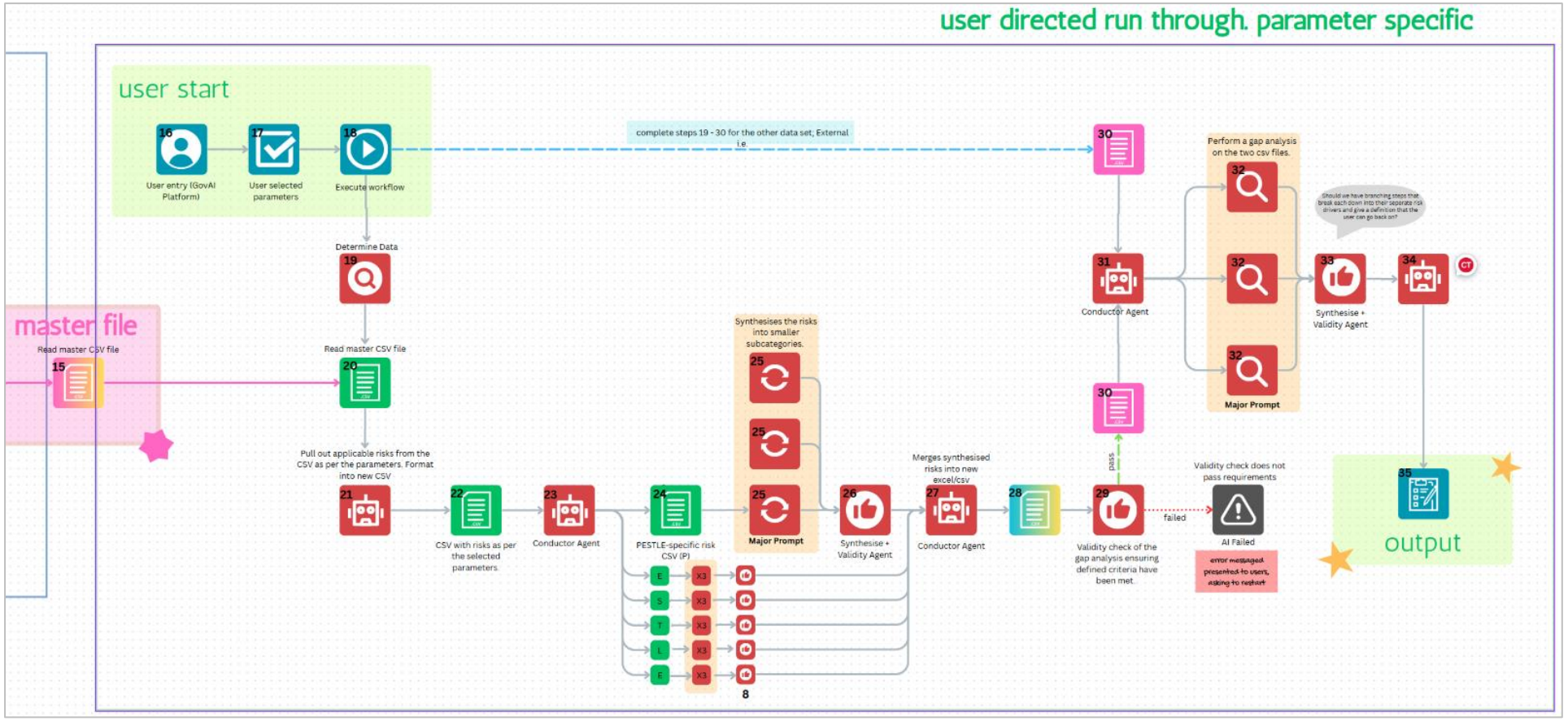
s 22 [redacted]
s 22 [redacted] Risk Policy Team,
Department of Finance
s 22 [redacted]

Screenshot/s

1. Automation Workflow Concept (Administrative Run)

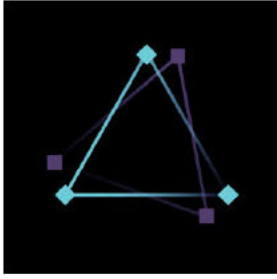


2. Automation Workflow Concept (User Directed Run)



3. Risk Gap Analysis Outcomes (Energy Pilot)

Risk Title	Risk Description	Is the Risk Being Considered (Yes/No)	Why or Why Not
Uncertainty in electric vehicle market growth	Consumer response, policy, and technology changes affecting EV market growth [1].	No	NCEs focus on CER integration challenges like voltage management and system stability [2, 3], but don't explicitly address EV market growth uncertainties. This could be an oversight, as EV adoption rates directly impact electricity demand and grid infrastructure needs.
Risks to renewables deployment	Supply chain pressures and permitting challenges hindering renewables growth [1].	Yes	The "Transition to Net-Zero" risk [4] acknowledges challenges in meeting climate transition targets, which implicitly includes risks to renewables deployment. However, the NCE document doesn't detail specific supply chain or permitting obstacles.
LNG surplus and price volatility	New LNG export capacity potentially creating a supply surplus and volatile energy demand patterns [5].	No	The NCE risks focus on the transition to net-zero and renewable energy goals [6], without explicitly addressing the impact of an LNG surplus. This is a gap, as LNG market dynamics can significantly influence Australia's energy export revenue and domestic gas prices.
Rising electricity demand from data centers and AI	Uncertainty in efficiency improvements affecting electricity demand from data centers and AI [5].	No	While NCEs recognize digital transformation challenges [7], they don't specifically highlight the rising electricity demand from data centers and AI. This is a potential blind spot, as these sectors are projected to significantly increase energy consumption, requiring proactive planning for grid capacity and efficiency improvements.
Impact of heat waves on electricity demand	Heat waves increasing electricity demand for cooling [8].	Yes	The NCE risks include "Extreme Weather Events" [9], which would encompass heat waves and their impact on infrastructure and communities. "Climate Change Impacts on Public Health & Communities" [10] explicitly recognizes the impact on public health and increasing humanitarian needs.



Detailed Overview

Version Control

Version	Date	Author	Description of Changes
1.0	3 Feb 2025	GovAI	Version 1 created
1.1	17 Mar 2025	GovAI	Modified based on feedback

Index

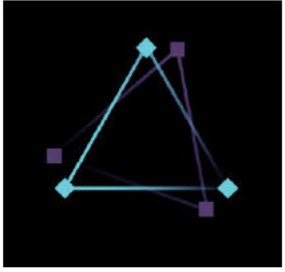
Note: For details about category items in the detailed overview, see *APS AI Use Case Repository Guidance- Guidance for Use Case Owners and Editors*.

Responsible Organisation Category

Select the Classification of the Functions of Government – Australia (COFOG-A) 3-digit category that best identifies the functional area associated with your AI use case.

☒ **01 - General Public Services**	014 - Planning and statistical services
☐ **02 - Defence**	Choose an item.
☐ **03 - Public Order and Safety**	Choose an item.
☐ **04 - Economic Affairs**	Choose an item.
☐ **05 - Environmental Protection**	Choose an item.
☐ **06 - Housing and Community Amenities**	Choose an item.
☐ **07 - Health**	Choose an item.
☐ **08 - Recreation, Culture, and Religion**	Choose an item.
☐ **09 - Education**	Choose an item.
☐ **10 - Social Protection**	Choose an item.
☐ **11 - Transport**	Choose an item.

Scope of the Use Case

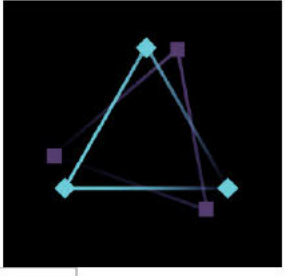


Use the dropdown menus below to identify the scope of your use case.

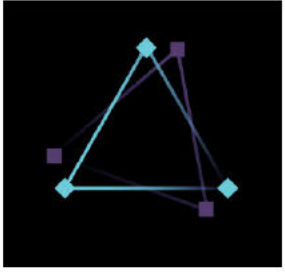
Geographical focus Choose the region for implementation from the dropdown list	National
Primary type of government interaction Choose the type of government interaction from the dropdown list	Government-to-government (G2G)
Cross-features - Sector Indicate if the use case describes a solution that can be used across sectors or in cross-sector scenarios (Yes/No).	Yes
Cross-features – Jurisdiction Indicate if the use case describes a solution that can be used across State/Federal borders or in cross-border scenarios (Yes/No)	Yes

Ethical Considerations

Accuracy, Fairness, Accessibility, Bias and Discrimination	The tool is designed to ensure accuracy, fairness, and accessibility through the following measures: • **Diverse data sources**: Draws from a wide range of publicly available Commonwealth and non-Commonwealth documents, including domestic and international sources, to support balanced and representative analysis. Data sources to be updated monthly to ensure ongoing accuracy and bias evaluation. • **Structured categorisation**: Uses frameworks such as PESTLE to ensure consistent and transparent classification of risks. • **AI reasoning with source traceability**: Outputs are paired with direct quotes from source documents to enable verification and reduce the risk of hallucination.



	• **Human-in-the-loop review**: All AI-generated outputs will be presented in a format that requires human selection. Users must review outputs before being used in further analysis or shared with stakeholders. • **Transparency**: The tool's processes (excluding Dragonfly, which is an optional part of the analysis) are documented to support accountability and continuous improvement. • **Accessibility**: Designed to be usable by agencies of all sizes, including those with limited analytical capacity.
Privacy	• The tool uses only publicly available data and does not collect, store, or process personal information. • During the pilot and feedback phases, any personal information provided by participants is handled in accordance with privacy obligations. • No user data is retained by the system, and all analysis is conducted on documents sourced from the public domain.
Rights of Users	• Users are informed that the tool is a decision support aid, not a final decision-maker. • Outputs are presented with source references and reasoning, allowing users to assess their relevance and accuracy. • Although users do not interact directly with the AI component, they are encouraged to provide feedback via the Comcover inbox or staff. • Users can challenge or clarify AI-generated insights, and their input is used to refine the tool and ensure it remains fit for purpose.



Value of the Use Case

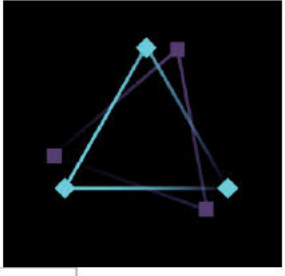
Identify the public value that the solution provides or is expected to provide. Select from the multi-select options.

Improved public service This category refers to solutions that enhance the services provided to end users, whether they are citizens or businesses.	☐ Personalised services ☐ Public (citizen)-centred services ☒ Increased quality of public information and services ☒ More responsive, efficient and cost-effective public services ☐ New services or channels
Improved administrative efficiency This category refers to solutions that increase efficiency, effectiveness, and quality while reducing costs within administrative processes, systems, and services.	☒ Cost reduction ☒ Responsiveness of government operation ☒ Improved management of public resources ☒ Increased quality of processes and systems ☒ Better collaboration and better communication ☐ Reduced risk of corruption and abuse of the law by public servants ☐ Greater fairness, honesty and equality enabled
Open government capabilities This category refers to solutions that enhance the level of openness, transparency, engagement, and communication within public organisations.	☐ Increased transparency of public sector operations ☐ Increased public participation in government actions and policymaking ☐ Improved public control of and influence on government actions and policies

AI Process Type

Select the types of tasks within government operations that the AI solution is performing or expected to perform

Supporting Decision Making- Tasks that support formal or informal agency decision-making on benefits or rights.	☐ Taking decisions on benefits ☐ Managing copyright and intellectual property rights
Analysis, monitoring and regulatory research -	☒ Information analysis processes ☐ Monitoring policy implementation

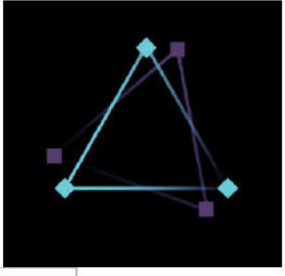


Tasks that collect or analyse information that shapes agency policymaking.	☒ Innovating public policy ☒ Prediction and planning
Enforcement – Tasks that identify or prioritise targets of agency enforcement action.	☐ Smart recognition processes ☐ Management of auditing and logging ☐ Predictive enforcement processes ☐ Supporting inspection processes ☐ Improving cybersecurity ☐ Registration and data notarisation processes ☐ Certification and validation processes
Internal management – Tasks that support agency management of resources, including employee management, procurement, and maintenance of technology systems.	☐ Internal primary processes ☒ Internal support processes ☒ Internal management processes ☐ Procurement management ☐ Financial management and support
Public services and engagement – Tasks that support the direct provision of services to the public or facilitate communication with the public for regulatory or other purposes.	☐ Engagement management ☐ Data-sharing management ☐ Governance and voting ☐ Payments and international transactions ☐ Supporting disintermediation ☐ Authentication of self-sovereign digital ID services ☐ Service integration ☐ Service personalisation ☐ Tracking of goods and assets along the supply chain

AI Technologies Utilised

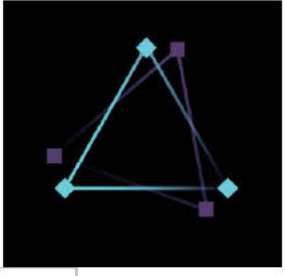
Select the types of AI technologies proposed / utilised to deliver the use case.

Reasoning or Knowledge Representation AI systems that store, structure, and process knowledge to make inferences, derive conclusions, or support decision-making.	☐ Knowledge Representation ☐ Automated Reasoning ☐ Commonsense Reasoning

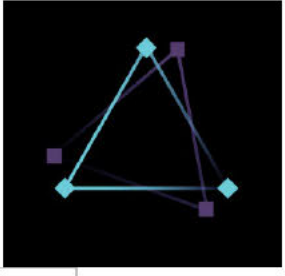


Planning and Optimisation AI techniques that generate, refine, and optimise action sequences or resource allocation to achieve specific goals efficiently.	☐ Planning and Scheduling ☐ Searching ☒ Optimisation
Learning and Adaptation AI systems that identify patterns, extract insights, and improve performance over time based on data.	☒ Machine Learning ☐ Deep Learning ☒ Generative AI
Communication and Natural Language Processing AI systems that process, interpret, and generate human language for interaction, comprehension, and automation.	☒ Natural Language Processing (NLP) ☒ Text Generation ☒ Text Mining ☐ Machine Translation
Perception through the Senses AI systems that process and interpret sensory data, such as visual, auditory, or tactile inputs, to understand and respond to their environment.	☐ Computer Vision ☐ Audio Processing
Integration and Interaction with the Environment AI systems that interact with physical or digital environments, including autonomous agents, robotics, and interconnected systems.	☐ Multi-agent Systems ☐ Robotics and Automation ☐ Connected and Automated Vehicles (CAVs)
AI as a Service AI capabilities delivered through cloud-based platforms, offering tools, models, and infrastructure for AI-powered applications.	☐ AI Services (e.g., cognitive computing, machine learning frameworks, bots) ☐ Infrastructure as a Service (IaaS) ☒ Platform as a Service (PaaS) ☐ Software as a Service (SaaS)
Additional Comments or Explanation:	If you have selected any of the subcategories above, feel free to provide more detailed comments or a description of how these elements apply to your specific use case.

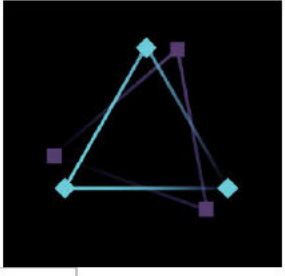
Technical Elements



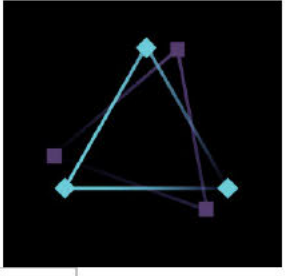
Platform implementation	The foresighting tool pilot was implemented as a manual workflow using publicly available generative AI tools accessed through consumer-facing web interfaces. It was not deployed as a standalone application or integrated into enterprise systems. • **Execution Environment**: The workflow was conducted using the chat interfaces of ChatGPT 4.0,Gemini Flash and NotebookLM. These tools were used to analyse risk-related content through structured prompts. Outputs were manually reviewed and synthesised by analysts. • **Hosting and Data Residency**: Both models were hosted on external infrastructure managed by OpenAI and Google. At the time of the pilot, there was no readily available option to select data residency or enforce Australian data sovereignty. The models processed data on overseas servers, and no sovereign or local hosting was used. • **Data Handling**: All data processed by the models was publicly available. Documents were manually compiled and stored in Microsoft 365 (e.g. SharePoint) and used as the basis for queries. No personal, sensitive, or classified information was uploaded or processed. • **Security**: The pilot was conducted in a secure internal environment. Outputs were reviewed before being shared or used in downstream analysis. The use of only public data reduced privacy and security risks. • **Identity Management**: No identity or access management was required during the pilot, as the tools were accessed by a small internal team using standard accounts. • **Cost Model**: The pilot leveraged free or existing licensed access to ChatGPT and



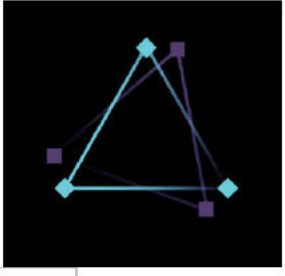
	Gemini. No dedicated infrastructure or vendor costs were incurred during this phase. This setup enabled rapid experimentation and validation of the concept while maintaining a clear boundary around data privacy, security, and responsible use.
Model / Algorithm used	The pilot used two publicly available generative AI models: • **Gemini 2.0 Flash** was used for data ingestion and synthesis. It was selected for its ability to process large volumes of unstructured content and extract specific risk insights. • **ChatGPT 4.0** was used for categorisation, particularly effective in applying structured frameworks like PESTLE and producing well-organised outputs. • **NotebookLM** was used for the comparative gap analysis between synthesised risk registers. Model selection was based on practical testing and performance in handling diverse document types. Prompt engineering was used to tailor model behaviour, and outputs were manually validated against source documents. No fine-tuning or custom training was applied.
Data Sources Select the types of data sources used and provide relevant details.	☐ Internal ☐ Third-party ☒ Public ☐ Synthetic **Details:** • Government reports (corporate reports 2025-2025). • External sources (e.g. industry risk reports, consulting reports, media articles). All data used in the pilot was publicly available and manually compiled by the team. No personal, sensitive, or classified information was used.



Risk Assessment and Mitigation Details	The pilot addressed several technical and operational risks associated with the tool's development and use: • **Prompt Reliability**: Prompts were iteratively refined to ensure they extracted relevant and accurate risks. Outputs were manually validated against source documents to mitigate hallucination or misclassification. • **Model Constraints**: The notebook-style interfaces used (ChatGPT and Gemini) had limited context retention. This was mitigated by breaking down PESTLE categories into subcategories and manually combining or separating risk categories as needed. • **Data Quality**: Only publicly available documents were used. Manual review ensured that AI-generated risks were grounded in verifiable source material. • **Bias and Coverage**: To reduce bias, the tool drew on diverse sources (e.g. industry reports, media, and Commonwealth documents). AI looping and human-in-the-loop review were used to ensure balanced outputs. • **Cybersecurity and Data Protection**: No personal or sensitive data was processed. The pilot was conducted in a secure internal environment, and outputs were reviewed before being shared or used in downstream analysis.	
Security and Compliance Frameworks Select the security and compliance frameworks and measures implemented. Provide details or additional artifacts if relevant.	☐ Authority to Operate (ATO) ☐ System Security Plan (SSP) ☐ Security Risk Management Plan (SRMP)	☐ Information Security Registered Assessors Program (IRAP) ☐ Penetration Testing
	Details: The pilot followed internal risk management practices.	



Assurance and Government Frameworks	A formal, comprehensive assessment against government AI assurance frameworks was not conducted during the pilot phase. However, the approach taken was designed to align with the intent and principles of the following frameworks: • **DTA AI Assurance Framework**: While not formally assessed, the pilot followed key principles of the framework, including human-in-the-loop oversight, transparency of AI outputs, and the exclusive use of publicly available data. An internal AI use register was maintained, and all staff involved completed AI fundamentals training. • **AI in Government Policy**: The proof-of-concept stage complied with core policy expectations by avoiding the use of personal or sensitive data, ensuring that AI outputs were reviewed by human analysts, and documenting the tool's purpose, limitations, and governance approach. • **National AI Assurance Framework**: The pilot aligned with the eight AI Ethics Principles, including fairness, accountability, reliability, and transparency. Human oversight was embedded through manual validation of AI outputs, and a documented audit trail was maintained throughout the pilot lifecycle. If the tool progresses toward production or broader deployment, a full evaluation against these frameworks – including formal risk assessments, transparency statements, and governance documentation – would be undertaken to ensure compliance and responsible AI use at scale.
Record maintenance	Documentation practices during the pilot included: • Maintaining a version-controlled prompt library.



	• Recording model selection rationale and testing outcomes. • Tracking data sources and preprocessing steps. • Capturing feedback from pilot participants and incorporating it into refinements. • Using a decision register to log key development choices and risk assessments. These practices supported transparency, auditability, and continuity throughout the pilot.
Disengagement	N/A – The pilot was exploratory and not deployed as a persistent or production system. No formal disengagement plan was required.
Performance Metrics and Results	• **Accuracy**: Risk gaps identified by the AI systems were manually validated against source documents. • **Specificity**: Risks were categorised using PESTLE and compared against Comcover benchmarking results. • **Timeliness**: A complete manual gap analysis could be conducted within approximately three hours (excluding data upload). • **Coverage**: The pilot successfully identified at least five under-considered risks not reflected in Commonwealth documentation. • **Qualitative Feedback**: The tool was used by the Commonwealth Risk Committee Secretariat. It was found effective in undertaking a gap analysis of risks identified in 2025-25 corporate plans and global research papers from external sources (WEF, OECD, UN, etc.).