Australian Government | GovAI

Assessment Assistant

Assessing data against criteria using GenAI

Department of Finance

Executive Summary

Problem Faced

GovAI required an app to demonstrate the potential of Generative Artificial Intelligence (Gen AI) in addressing common challenges within the Australian Public Service (APS). Through consultation with senior executives, grant evaluation was chosen as an appropriate subject matter to showcase the app's capability.

Current evaluation approaches in the APS can be labour-intensive, inconsistent, and inefficient. Additionally, these methods can result in delays and reduced service quality. There is a need for solutions that enhance consistency, productivity, transparency, and accuracy in government evaluations.

Solution Overview

GovAI developed the Assessments Assistant app to showcase the potential of Gen AI in improving evaluation processes within the APS. The app assesses responses against a set of criteria, showing how AI can be leveraged to enhance the efficiency, accuracy, and consistency of evaluations.

Key Features:

- **Review Assistant:** Evaluates documents against predefined criteria, providing compliance assessments, scores, and information about criteria.
- **Chat Assistant:** Enables users to interact with AI to ask questions about the document and criteria using natural language.

Benefits and Impact

The app offers several key benefits:

- **Improved Efficiency:** By streamlining the evaluation process, the app reduces the time, effort, and costs associated with Government assessments.
- **Greater Consistency:** AI supported evaluations based on predefined criteria may reduce the likelihood of errors and variability in assessments.
- **Support for Decision-Making**: Compliance assessments and scores may contribute to more informed and evidence-based decision making.
- **Broader Impacts:** Fair and consistent evaluation processes are likely to have broader societal benefits.

Target Audience and Stakeholders

The Assessment Assistant app is designed for a range of users within the APS such as:

- APS employees who conduct evaluations
- Technical staff exploring AI capabilities
- Executives making decisions about AI adoption.

These stakeholders are impacted by the challenges in APS evaluations and stand to benefit from the improved efficiency, accuracy, and consistency demonstrated by the app.

App development was informed by consultations with APS executives. Although the current version is not production-ready, it is expected to serve as a valuable demonstration to stakeholders of the potential benefits of AI in government evaluations.

Risks and Mitigation Overview

To ensure the secure and ethical operation of the app, implemented measures include:

- Reliance on up-to-date static criteria documentation to ensure data quality and avoid the use of sensitive data.
- Disclaimer which must be read before logging in via the GovAI website, to prevent users uploading Official or Sensitive information.
- Features that augment rather than automate users' decision making.
- Governance and oversight mechanisms, including authentication via GovTEAMS accounts and developer console logs.

These measures ensure that the app operates with proper monitoring and control.

Use Case Status

Pilot

Use case timeline

Nov 2024: Development of Proof of Concept
Mar 2024: Demonstration to stakeholders (including Department of Finance Secretary)
Apr 2025: Demonstration version released

Additional Information

The application includes both synthetic data which imitates grant applications, and publicly available Australian Government grant documentation. It could be useful for any government service that needs to review submissions against specific criteria, including research proposals and procurement activities.

Lessons Learned

Key insights from implementing this AI use case include:

- Feedback indicated a desire for Generative AI use cases that support common work tasks, like evaluation.
- Regular updates to criteria documentation essential for maintaining accuracy in assessments.
- Subject matter expertise is key to effective application development

Contact information

Responsible Entity Name

Department of Finance

Area of Entity

GovAI
Government Services Branch
Business Enabling Services Group

Use Case Website/s

https://www.govai.gov.au/

Open for Collaboration?

Yes! Feel free to talk to us if you have any questions.

Use Case Contact

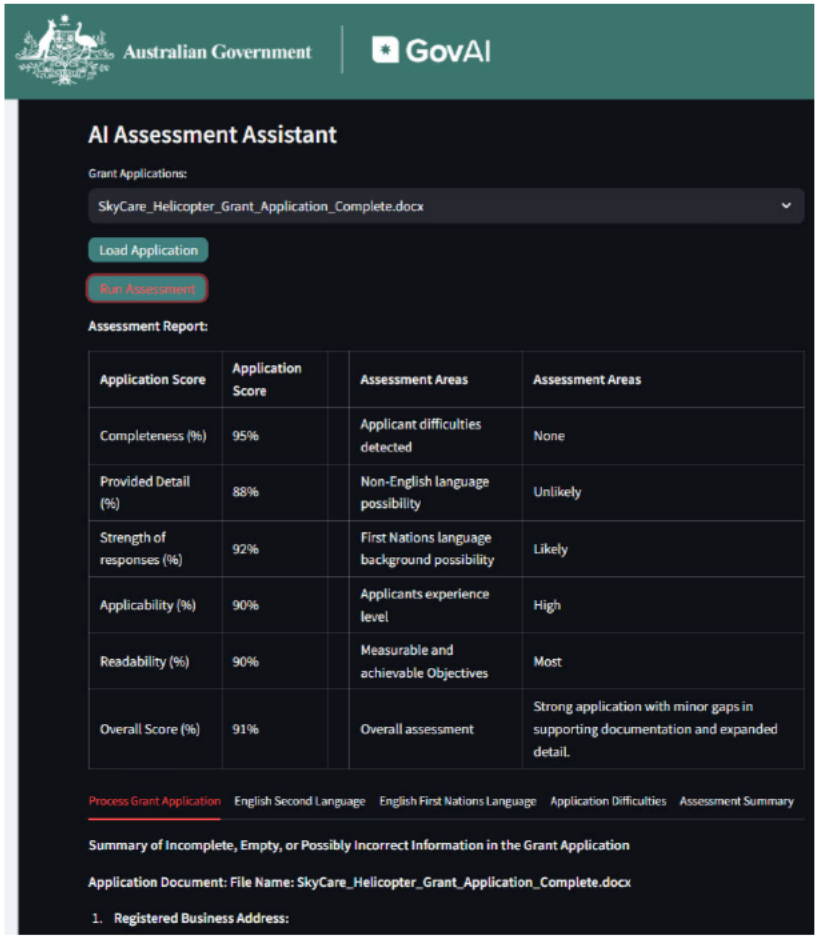
GovAI
GovAI@finance.gov.au

Use Case Owner

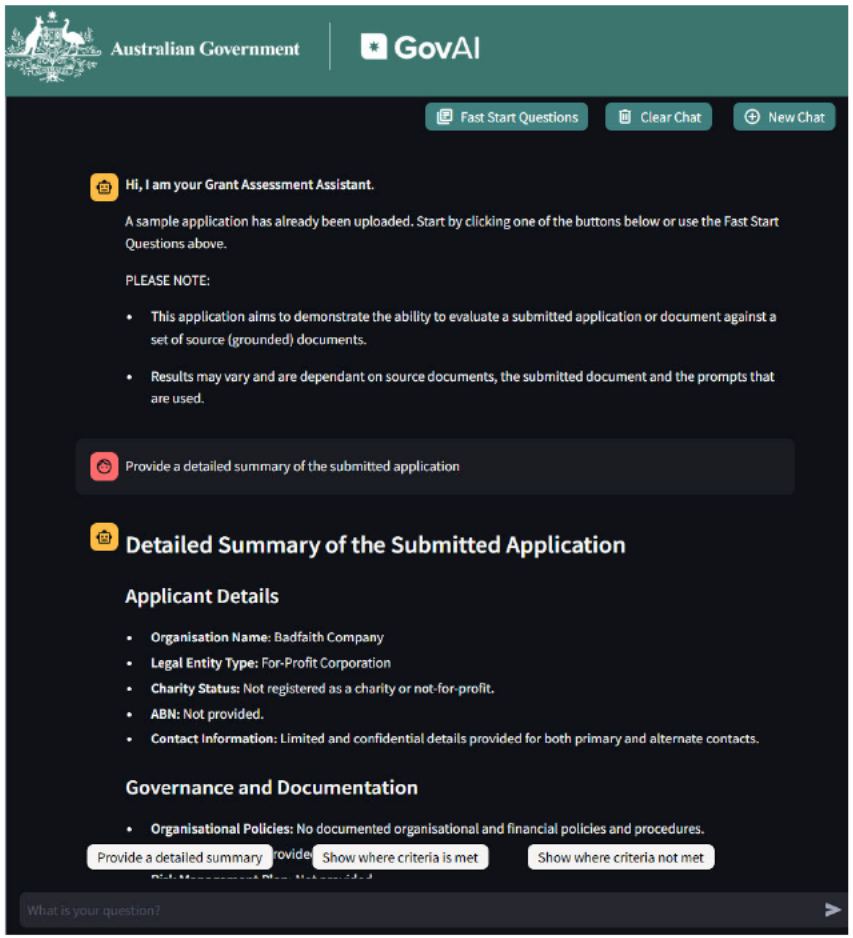
Angie Earl
GovAI Program Director
Government Services Branch

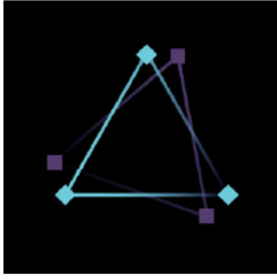
Screenshot/s

Review Assistant



Chat Assistant





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

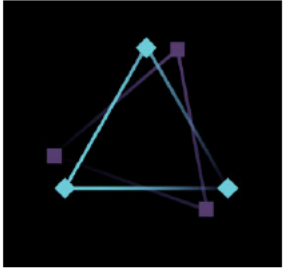
Index

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**	019 - General public services (other)
☐ **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	To ensure the AI system operates accurately, fairly, accessibly, and without discrimination, the Assessment Assistant employs a multi-faceted approach: • The system uses pre-defined criteria to maintain consistency and accuracy in evaluations. It leverages a curated set of documents, such as grant policies, to ensure the information provided is accurate and relevant. • It is designed to be representative by using publicly available and approved documentation, along with synthetic data that closely imitates existing Australian Government grants. Any bias in the system is likely to reflect bias in the assessment documentation.

	• Mechanisms are included for assessors to interact with AI using natural language, allowing for deeper insights and clarification of uncertainties.
Privacy	The Assessment Assistant upholds privacy rights and personal data protection by using pre-loaded documents for demonstration purposes. The system limits this preloaded information to publicly available grant policies and synthetic data. Users cannot upload their own documents, ensuring that no personal data is processed or stored. The controlled environment helps protect personal data and ensures privacy compliance. Users are advised to use only publicly accessible information, minimising privacy and personal data risks.
Rights of Users	Users can access the full grant criteria documentation cited in the responses, ensuring they can challenge AI decisions if needed. This transparency ensures that users are informed about their rights and can provide feedback or challenge the AI's decisions if necessary.

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	☒ Cost reduction ☒ Responsiveness of government operation

quality while reducing costs within administrative processes, systems, and services.	☒ 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 – Tasks that collect or analyse information that shapes agency policymaking.	☒ Information analysis processes ☐ Monitoring policy implementation ☐ 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 –	☒ Internal primary processes ☐ Internal support processes

Tasks that support agency management of resources, including employee management, procurement, and maintenance of technology systems.	☒ 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	☒ Natural Language Processing (NLP) ☒ Text Generation ☐ Text Mining

AI systems that process, interpret, and generate human language for interaction, comprehension, and automation.	☐ 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 "Assessment Assistant" application is hosted on Azure, leveraging a variety of Azure resources to deliver its functionalities. It is built using Python, utilising FastAPIs for the backend and Streamlit for the UI. The final solution utilises Docker containers and deployed for demonstration purposes using Azure Container Applications. Communication with the UI occurs through HTTPS (TLS 1.2) and authentication is backed by Azure Entra ID registered applications.

	The solution uses several Azure resources to support AI processing and storage components, including Azure Search for indexing and querying capabilities, Azure OpenAI for integrating powerful language models, and Azure Storage Accounts for scalable and secure data storage. The cost model involves leveraging Azure's pay-as-you-go pricing for cloud services. Identity service is managed through secure authentication mechanisms, such as GovTEAMS accounts.
Model / Algorithm used	The application utilises a Large Language Model (LLM) and Retrieval Augmented Generation (RAG) to provide its chat, question-answer, and assessment capabilities. The LLM is responsible for understanding and generating responses based on user queries, while the RAG component enhances the model's ability to retrieve relevant information from a specific set of internal unstructured data. The current LLM being utilised for the demonstration application is GPT 4o. The design of the application allows for flexibility in the backend, enabling it to be used with a variety of models that can be plugged in as needed. This modular approach ensures that the application can be easily updated or customised to meet specific requirements or to leverage advancements in AI technology.
Data Sources Select the types of data sources used and provide relevant details.	☐ Internal ☐ Third-party ☒ Public ☒ Synthetic **Details:** The application uses a set of curated documents that are pre-loaded. It is designed to use a diverse range of data sources as part of this curated set. In the case

	of this demonstration, the data used is as follows: • **Public**: Publicly available Australian Government grant policies documents. • **Synthetic**: Synthetic data that imitates existing Australian Government grants.
Risk Assessment and Mitigation Details	The application employs a comprehensive approach to risk analysis to ensure that its operation is smooth and secure. Key measures include: • **Authentication:** Using Entra ID, this ensures secure access to the application, protecting sensitive data and preventing unauthorized access. • **Sensitive Information Mitigation:** A disclaimer must be read before logging in via the GovAI website. • **Governance and Oversight:** Mechanisms in place to ensure safe and responsible use include authentication via GovTEAMS accounts.
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:** No Security and Compliance Frameworks have been implemented at this stage.
Assurance and Government Frameworks	Australia's AI Ethics Principles have been considered and applied during planning, however, no formal assessment against the AI Assurance framework has been undertaken at this time as this is not mandatory for low-risk use-cases.
Record maintenance	Documentation is maintained for model choices, training materials, testing results, and the curated document set. The platform also supports the exporting of conversations,

	allowing users to maintain records of their interactions, and review by administrators.
Disengagement	Models can be disconnected by admin.
Performance Metrics and Results	To monitor and evaluate the benefits of the AI solution, we are investigating the use of the following performance indicators and metrics: • **Access Frequency**: How often the application is accessed. • **User Types**: Differentiating between new and returning users. • **Daily Usage**: Overall daily interaction with the application. • **Cost and Usage**: Tracking the costs and usage of the application. The app also supports performance tracing and monitoring with Azure Application Insights.