

CASE STUDY

AMCA Digital Enablement Initiative

CONSTRUCTION INDUSTRY



Project: Agentic-RAG Chatbot Assurance and AI Maturity Roadmap



KEY OUTCOME 1

Validated chatbot release readiness



KEY OUTCOME 2

Strengthened AI governance foundations



KEY OUTCOME 3

Established ongoing evaluation capability

BACKGROUND

AMCA was pursuing a Digital Enablement Initiative to improve member access to workplace health and safety information through an AI-powered Agentic-RAG chatbot. The solution was designed to use AMCA's body of knowledge, including safety and project documentation, to provide members with practical guidance through an interactive interface.

KJR was engaged to support safe adoption of the solution by developing an AI Maturity Roadmap and performing independent assurance of the chatbot before member release.

CHALLENGE

AMCA needed confidence that its workplace health and safety chatbot would provide accurate, grounded and trustworthy responses while minimising risks associated with hallucination, harmful content, privacy leakage, unfair bias, poor retrieval performance and inappropriate responses to out-of-scope questions.

SOLUTION

KJR applied its Validation Driven Machine Learning (VDML) assurance methodology to assess the chatbot before release. KJR developed an AI Maturity Roadmap using the National AI Centre AI Adoption Guidelines and ISO/IEC 38500, then designed and executed a structured LLM evaluation program using curated test datasets. Testing assessed factual correctness, answer relevance, faithfulness, context relevance, harmful content, PII leakage, unfair bias, retrieval precision and recall, and refusal behaviour.

KJR also provided recommendations for developing overall organisational AI governance processes and ongoing monitoring, regression testing, drift detection and OWASP LLM threat monitoring for the chatbot solution.

CASE STUDY

AMCA Digital Enablement Initiative

CONSTRUCTION INDUSTRY



Project: Agentic-RAG Chatbot Assurance and AI Maturity Roadmap

DELIVERABLES

- AI Maturity Roadmap aligned to National AI Centre AI Adoption Guidelines and ISO/IEC 38500
 - Assessment of governance, data foundations, architecture, responsible AI, risk management, quality engineering and operating model maturity
- LLM evaluation strategy and test design
 - Curated evaluation datasets
 - Validation of factual correctness, answer relevance and groundedness
 - Testing for harmful content, PII leakage and unfair bias
 - Retrieval precision and recall assessment for the RAG implementation
 - Refusal-behaviour and out-of-scope request testing
- Release-readiness assurance reporting
- Recommendations for regression testing, drift detection and operational monitoring
- Handover of an evaluation platform to support ad hoc and scheduled chatbot evaluations
- OWASP LLM threat monitoring recommendations for production deployment

VALUE TO CLIENT

AMCA received independent evidence that the chatbot was operating within acceptable performance and safety boundaries before member release. The engagement reduced deployment risk, strengthened governance and responsible AI controls, and provided a practical roadmap for safe AI adoption, operational monitoring and ongoing improvement.

TOOLS & TECHNOLOGIES

- Agentic-RAG architecture
- Large Language Models (LLMs)
- Validation Driven Machine Learning (VDML)
- Arize AI Phoenix evaluation framework
- KJR GenAI evaluation platform
- OWASP LLM threat monitoring practices
- National AI Centre AI Adoption Guidelines
- ISO/IEC 38500 governance framework

