

2026 NASCIO STATE IT RECOGNITION AWARDS

Category: Enterprise Technology Management

AI-Powered IT Spend Request Assessment: Bringing Consistency, Speed, and Data-Driven Rigor to Government IT Investment Decisions

State:	Hawaii
Agency:	Office of Enterprise Technology Services (ETS)
Project Title:	AI-Powered IT Spend Request Assessment
Project Dates:	January 2026 – Present (Phase 1 in production)
Contact:	Todd Omura, IT Governance Officer Office of Enterprise Technology Services State of Hawaii

Executive Summary

Hawaii law (HRS §27-43) charges the Office of Enterprise Technology Services (ETS) and its Chief Information Officer with statewide IT governance. Under the Governor's Budget Execution Policy (EM-25-03, Section 20.d), every IT acquisition of \$100,000 or more—hardware, software, projects, maintenance contracts, consultant services, and more—requires CIO approval. Departments submit an IT Spend Request through the state's IT portfolio management system with supporting documentation. Projects at \$1 million or above, or those designated by the CIO as enterprise initiatives, must additionally be reviewed and approved by the Project Advisory Council (PAC), whose voting members are the CIO, the DAGS Comptroller, the Director of Finance, and the Director of Human Resources Development (Administrative Directive 18-03).

This governance structure generates hundreds of IT Spend Requests each fiscal year. Every request requires evaluation across multiple dimensions—cost justification, strategic alignment with state IT goals and the department's multi-year IT roadmap, execution risk, impact of inaction, vendor history, and comparison with similar past initiatives. Until now, this evaluation has been entirely manual: analysts cross-reference data from the state's enterprise architecture platform, research vendors and market pricing independently, and produce narrative assessments without a standardized framework. The result was inconsistent evaluations, slow turnaround, and no institutional memory across assessment cycles.

Hawaii built an AI-powered assessment system that transforms this process. When a reviewer triggers an assessment, the system automatically retrieves the complete project record from the state's enterprise architecture platform, searches a knowledge base of over 1,400 historical IT Spend Requests for relevant precedents, researches current market pricing and vendor government contract history via web search, and synthesizes all findings into a comprehensive Word document assessment—delivered in two to four minutes.

Since deployment, the system has reduced assessment preparation time from hours of manual research to minutes of automated analysis, established consistent evaluation standards across all requests, and created institutional memory that improves with every project added to the knowledge base. The architecture is entirely cloud-native, using serverless computing, and the approach is replicable by any state with an enterprise architecture platform and a body of historical project data.

Idea

What problem or opportunity does the project address?

Under Hawaii's Budget Execution Policy (EM-25-03, Section 20), IT governance follows a tiered approval structure. Acquisitions under \$100,000 are delegated to department heads (Section 20.f). All IT acquisitions at \$100,000 or more require approval by the Chief Information Officer or designee (Section 20.d).

Departments submit an IT Spend Request—a project record maintained in the state's IT portfolio management system—along with supporting documentation. Each request must align with the department's multi-year IT roadmap per HRS §27-43. Projects at \$1,000,000 or above, legislatively mandated, or CIO-designated enterprise initiatives require additional approval by the Project Advisory Council under AD 18-03. ETS requires at least two work weeks to complete its review.

This governance process generates hundreds of IT Spend Requests each fiscal year, each requiring ETS's IT Governance staff to evaluate cost justification and budget reasonableness, alignment with state IT strategic goals and departmental roadmaps, execution risk, impact of inaction, vendor history and market pricing, and similarity to previously approved initiatives.

Before this project, every dimension was assessed manually. An analyst would open the project record in the state's enterprise architecture platform, read through dozens of fields, attempt to recall or search for similar past requests, independently research vendor pricing and government contract history, and then compose a narrative assessment. There was no standardized framework for what "good" looked like. Each analyst brought different expertise, different recall of institutional history, and different levels of rigor. The result:

- Inconsistent evaluations — the same request might receive substantially different assessments depending on who reviewed it.
- No institutional memory — lessons from past projects (failed vendor relationships, cost overruns on similar initiatives, successful patterns) existed only in individual analysts' recollection.
- Slow turnaround — manual research across multiple data sources consumed hours per request, creating bottlenecks during peak submission periods.
- No validation of self-reported scores — three of four inputs to the state's Priority Score are self-assessed by submitters, with no systematic way to verify accuracy across hundreds of requests.

Why does it matter?

IT investment decisions directly affect government service delivery to Hawaii's 1.4 million residents. AD 18-03 requires each request be evaluated for return on investment, alternatives, business justification, total costs, and procurement approach before public funds are committed. An inadequately evaluated request can lead to duplicative systems, vendor lock-in, or cost overruns. Conversely, an overly conservative review can delay critical infrastructure upgrades and citizen-facing digital services.

What makes it different?

Most government AI initiatives apply language models to generate content or answer questions. Without grounding, these models produce plausible-sounding text but fabricate project names, invent budget figures, and cite nonexistent precedents—unacceptable for investment decisions. This system solves the grounding problem with three layers that ensure every claim is backed by verifiable data:

- **Organizational Memory (RAG)** — A continuously updated knowledge base indexes 1,400+ historical IT Spend Requests. The system retrieves similar past projects by technology, vendor, budget range, and scope—from the state’s actual records, not AI memory.
- **Live Enterprise Data (MCP)** — AI agents connect directly to the enterprise architecture platform to pull current project details—budgets, relationships, lifecycle status, risk scores—ensuring analysis is never based on stale data.
- **Market Intelligence (Web Search)** — Current vendor pricing, government contract history, and market context are researched automatically for every evaluation.

Additionally, the system includes an anti-gaming capability unique among government AI assessment tools. It independently validates self-reported priority scores by cross-referencing claimed business value, project risk, and inaction risk against evidence in the project data, flagging inconsistencies for reviewer attention.

What makes it universal?

Every state evaluates IT investment requests and struggles with the same challenges: inconsistent standards, loss of institutional knowledge through turnover, and the impossibility of manually comparing each request against the full historical portfolio. This solution is directly replicable by any state with an enterprise architecture platform and historical project data. The multi-agent pattern can be adapted to grant applications, procurement reviews, and compliance assessments. It aligns with the State CIO Top Ten Priority of leveraging AI to improve government efficiency.

Implementation

What was the roadmap?

The project followed an iterative, prototype-driven approach rather than a traditional waterfall methodology. This was essential because the problem—applying AI to evaluate IT investment requests—had no established playbook in government.

- **Phase 1 (Q1 2026) — Requirements & Data Foundation:** Defined the assessment framework: identified key evaluation dimensions (cost, alignment, risk, inaction risk, vendor history, precedent comparison) and the data sources needed for each.
- **Phase 2 (Q1 2026) — Knowledge Base & RAG Pipeline:** Built the data pipeline to export all historical project records from the enterprise architecture platform into an AI-searchable knowledge base. Over 1,400 project records, each with 40+ fields, are converted into AI-optimized text format and indexed for semantic search.
- **Phase 3 (Q1 2026) — AI Agent Development:** Developed the multi-agent AI pipeline with three specialized agents (Orchestrator, Historian, Researcher) running in parallel. Integrated real-time data enrichment via the enterprise architecture platform’s API. Built the scoring validation engine to detect self-assessment gaming.
- **Phase 4 (Q2 2026–Present) — Production Deployment:** Deployed to production on a major cloud platform as a serverless service. Implemented automated nightly knowledge base refresh and pre-generation of assessments. Currently in active use by IT Governance reviewers.

Success is measured by three criteria: consistency of assessments across different requests (standardized framework applied uniformly), time to produce an assessment (target: under five minutes vs. hours of manual work), and grounding accuracy (every cited precedent and data point must be verifiable in the source systems).

Who was involved?

The project was initiated and championed by Hawaii's Chief Information Officer as part of a broader directive to explore how AI could improve IT governance processes—specifically the review and analysis workflow that ETS's IT Governance staff perform for every IT Spend Request under EM-25-03, Section 20.d and AD 18-03. The core development was led by a small team within ETS, working closely with IT Governance reviewers who provided continuous feedback on assessment quality and relevance.

Stakeholder buy-in was built through demonstration, not presentation. The team produced working prototypes within weeks rather than spending months on requirements documentation. IT Governance staff could compare AI-generated assessments against their own manual evaluations of the same requests, building confidence in the system's accuracy and identifying areas where the AI's analysis needed refinement.

How did you do it?

The system runs entirely on a major cloud platform using serverless infrastructure, requiring no dedicated servers or infrastructure management. Key resources:

- **Compute:** A single serverless compute service hosts the API and orchestrates the entire workflow.
- **AI Models:** Multiple AI foundation models are used through a managed AI platform: a reasoning-optimized model for the Orchestrator and Historian agents, and a speed-optimized model for the Researcher agent.
- **Search & Retrieval:** A managed search and retrieval service indexes the historical knowledge base and powers semantic search across 1,400+ past project records.
- **Automation:** Automated nightly jobs refresh the knowledge base and pre-generate assessments so reviewers receive instant results each morning.

The architecture prioritizes reliability: exponential backoff handles AI rate limits, smart caching prevents redundant processing for unchanged records, and the system is fully containerized for consistent deployment.

Impact

What did the project make better?

This project transformed IT Spend Request evaluation from an inconsistent, labor-intensive manual process into a standardized, data-driven, AI-augmented workflow. The improvements are both quantitative and qualitative:

- **Speed:** Assessment preparation reduced from hours of manual research per request to a two-to-four-minute automated analysis. Pre-generated assessments via nightly automation deliver instant results when reviewers begin their day.
- **Consistency:** Every request is now evaluated against the same framework, using the same data sources, with the same depth of analysis. The assessment no longer depends on which analyst happens to be assigned.

- **Institutional Memory:** The system searches across 1,400+ historical project records for every evaluation—a breadth of comparison no human analyst could replicate from memory. Past vendor performance, cost patterns, and outcome history are surfaced automatically.
- **Evidence-Based Analysis:** Every statement in the assessment is grounded in verifiable sources: actual historical project records, live data from the enterprise architecture platform, or current market data from web search. No claims are generated from the AI’s training data alone.
- **Scoring Integrity:** Self-reported Priority Scores are independently validated against project evidence. The system identifies when Business Value, Project Risk, or Inaction Risk scores appear inconsistent with the supporting data, providing reviewers with specific discrepancies to investigate.
- **Compounding Value:** The knowledge base improves with every project added. As the state’s project portfolio grows, the system’s ability to find relevant precedents and identify patterns strengthens—creating a compounding return on the initial investment.

How do you know?

The system’s impact is measurable across several dimensions:

Metric	Before (Manual)	After (AI-Augmented)
Assessment Time	Hours of manual research	2–4 minutes (automated) Instant with pre-cached results
Knowledge Base Coverage	Analyst memory (varies by individual)	1,400+ historical projects indexed and searchable
Evaluation Consistency	Variable by analyst	Standardized framework applied uniformly
Data Sources per Assessment	1–2 (manual lookup)	3 layers: historical RAG, live platform data, web search
Scoring Validation	None (self-reported accepted at face value)	AI cross-references each score against project evidence
Knowledge Retention	Lost with staff turnover	Permanent — knowledge base persists and grows

IT Governance reviewers have validated assessment quality by comparing AI-generated evaluations against their own manual assessments of the same requests. The AI consistently surfaces historical precedents and market data that reviewers had not identified in their manual review, while maintaining accuracy on all factual claims—a direct result of the three-layer grounding architecture. The anti-gaming capability has enabled reviewers to have targeted, evidence-based conversations with submitters when self-reported scores appear inconsistent with project evidence.

What now?

The system is in active production use. The longer-term roadmap includes:

- **Role-Aware Reports:** Dual-audience reports: analytical assessments for IT Governance reviewers and guidance-oriented versions for Spend Request authors, providing actionable suggestions to improve submissions before formal review.
- **Vendor Intelligence Agent:** A dedicated AI agent focused on vendor government contract history, past performance metrics, and financial health indicators—context that reviewers currently research manually.
- **Broader Application:** Applying the same multi-agent grounded-AI pattern to other state evaluation workflows: grant applications, procurement reviews, and regulatory compliance assessments.

- **Assessment Change Tracking:** When a project record is updated and re-assessed, automatically highlighting what changed in the data and how the AI analysis shifted—so reviewers see the delta, not a full re-read.

Hawaii's IT Spend Request assessment process is now faster, more consistent, more thorough, and more transparent than it has ever been—and the value compounds as the knowledge base grows with every project added.