



SHI RAG Studio

Purpose-built to bring trust, speed,
and clarity to enterprise generative AI.



SHI's RAG Studio empowers organizations to operationalize the full potential of generative AI with confidence, securely integrating large language models (LLMs) with enterprise-specific data. Through retrieval-augmented generation (RAG), SHI enables AI systems to reference trusted knowledge sources in real time – dramatically reducing hallucinations and enhancing factual accuracy. This solution accelerates the transition from experimentation to production, bringing transparency, auditability, and measurable ROI to every phase of your AI journey.

What is retrieval-augmented generation?

RAG is an AI technique that enhances LLM responses by grounding them in external, domain-specific data. Rather than relying solely on pre-trained knowledge, a RAG pipeline ingests and retrieves relevant information from your enterprise content – ensuring responses are context-aware, accurate, and traceable. This approach is essential for organizations that demand relevancy, accuracy, compliance, and trust in AI outputs.

Why RAG matters to your business

Retrieval-augmented generation (RAG) helps generative AI deliver answers grounded in your organization's trusted data – not just what the model was trained on. This means fewer hallucinations, more accurate responses, and better alignment with your policies and workflows. For business leaders, RAG translates into faster decision-making, stronger compliance, and increased confidence in AI outputs.

Key capabilities

- **Visual benchmarking:** Side-by-side comparison of base LLM vs. RAG-enhanced outputs.
- **Customizable experimentation:** Choose your LLM, embedding model, chunking strategy, and data source.
- **Secure and compliant:** No fine-tuning or data leakage – your data remains fully under your control.
- **Built-in observability:** Evaluate output accuracy, similarity thresholds, and response latency.
- **Plug-and-play integration:** Compatible with NVIDIA NIMs, LangChain, Facebook AI Similarity Search (FAISS), Pinecone, and more.

Business outcomes and value

- Up to 85% reduction in hallucination rates compared to base LLMs.
- Domain-specific accuracy exceeding 90% using customer data and tuned parameters.
- Working proof of concept in under 7 days – accelerating time to value.
- Increased stakeholder trust and adoption through transparent grounding and auditability.
- Lower development costs and faster deployment cycles than traditional AI buildouts.

SHI RAG Studio

Purpose-built to bring trust, speed,
and clarity to enterprise generative AI.

Ideal use cases across verticals

- **Public sector:** Legal research, policy Q&A, and RFP assistant.
- **Education:** Financial aid copilot, curriculum advisor, and AI front desk.
- **Healthcare:** Clinical policy advisor and staff knowledge assistant.
- **Legal:** Case law interpreter and regulatory search companion.
- **Finance:** Investment briefings and compliance updates.
- **Retail:** Product advisor and internal FAQ engine.

How to engage

Start with SHI's hosted sandbox environment featuring curated datasets. Then, switch to 'build your own' mode where you can upload your content, run experiments, and compare different configurations — all within a secure SHI AI Lab environment or your on-premises infrastructure. SHI's AI experts will guide you from initial experimentation to full-scale production deployment.

Why SHI?

SHI is not just a global IT solutions provider; we are your AI transformation partner. Our AI Lab infrastructure, in partnership with NVIDIA, is purpose-built for enterprise-grade ideation, experimentation, and adoption. This aligns directly with our strategic approach: Imagine, experiment, adopt. We empower organizations to imagine what's possible with AI, experiment in secure, scalable environments designed for rapid prototyping and production rollout, and adopt solutions with confidence — backed by our deep expertise in infrastructure, data governance, and enterprise AI.

Whether you're a CIO seeking strategic clarity or a developer building the next-gen AI app, SHI brings the technical depth and operational rigor needed to succeed. We help you navigate complexity, reduce risk, and unlock real business value.



Scan to
learn more

Speak to an AI expert. Contact us at AI@SHI.com.

Learn more at SHI.com/AI.

SHI International Corp. \\ SHI.com \\ 888-764-8888 \\ 2