SAS® Retrieval Agent Manager on G-Cloud

Service Definition

Get fast, trusted insights from unstructured enterprise data through no-code, scalable, agent-driven AI automation.

SAS® Retrieval Agent Manager on G-Cloud

UK Government departments operate in an environment where the scale, diversity, and operational criticality of data are increasing faster than traditional systems can handle. Every day, they must integrate information from legacy estates, modern cloud platforms, case management systems, operational databases, and third-party sources, all while meeting strict requirements for security, auditability, and governance.

SAS® Retrieval Agent Manager provides a secure, automated, and highly governed way to extract, synchronise, and operationalise data from complex, distributed systems, ensuring that public sector organisations can make trusted information available for analytics, investigations, mission-critical operations, and cross-government collaboration. With Retrieval Agent Manager, departments gain a resilient and controlled method for turning raw operational data into reliable intelligence, supporting evidence-based decisions that protect the public, enhance service delivery, and strengthen national resilience.

SAS® Retrieval Agent Manager helps government departments make better decisions by providing scalable, efficient access to unstructured enterprise data. It removes the inefficiencies of manual data extraction and addresses the challenges of traditional RAG approaches.

Built on the Retrieval Augmented Generation (RAG) framework, it streamlines complexity into a scalable solution that delivers fast, accurate, and context-aware AI responses from unstructured data. This allows departments to integrate advanced AI seamlessly into existing systems with confidence in reliability, control, and performance at scale.

Why is SAS Retrieval Agent Manager Important?

Over 80% of enterprise data is locked in unstructured formats across diverse documents and it's growing rapidly.

Enterprises are overwhelmed by this growing unstructured data, as relying on labor-intensive information retrieval processes is inefficient, time-consuming, error-prone, and exposes organisations to compliance risks, while simultaneously draining productivity and delaying critical insights.

While Retrieval-Augmented Generation (RAG) offers a way forward, traditional implementations are code-heavy, fragmented, and inflexible. They depend on technical teams, create vendor lock-ins, and lack agentic capabilities for automation, leaving workflows slow, manual, and difficult to scale across departments.

Traditional RAG implementations can slow GenAI adoption due to potential challenges such as:

- Heavy coding requirements create barriers for non-technical users, increasing dependency on IT resources slowing GenAI-driven decisioning
- Inefficient API strategy creates integration challenges with enterprise applications, restrict access to diverse data sources, and reinforce vendor lock-ins, constraining flexibility and scalability across use cases
- Lack of human oversight in validating built-in auto-evaluations can raise trust issues, undermine transparency and confidence, and increase the risk of unreliable outputs and non-compliance
- Limited adoption of optimisation methods such as model compression and ONNX acceleration leads to slower, more expensive AI operations
- Poor security standards such as access controls, privacy policies and audit trails elevate security and compliance risks
- Absence of an agent orchestration strategy causes agents to operate in isolation, limits automation of complex workflows, and leads to repeated tasks and loss of context.

What does SAS Retrieval Agent Manager do?

SAS Retrieval Agent Manager helps organisations make better decisions by providing scalable, efficient access to unstructured enterprise data. It removes the inefficiencies of manual data extraction and addresses the challenges of traditional RAG approaches.

Built on the Retrieval Augmented Generation (RAG) framework, it streamlines complexity into a scalable solution that delivers fast, accurate, and context-aware AI responses from unstructured data. This allows organisations to integrate advanced AI seamlessly into existing systems with confidence in reliability, control, and performance at scale.

The SAS Difference

- **Intuitive no-code experience:** Empowers non-technical users to build and manage workflows, supporting diverse document types, multilingual OCR, and usage monitoring.
- **Trustworthy and transparent AI**: Delivers citation-backed, source-grounded responses with built-in evaluation and human-in-the-loop oversight for transparency, traceability, and trust.
- **Modular, plug-and-play architecture**: Supports leading model providers and vector databases with no vendor lock-in, including native integration with file systems and Git for seamless onboarding.
- **High performance and enterprise-ready integration:** Delivers fast, efficient AI through model acceleration techniques and integrates into enterprise systems via robust APIs and secure, scalable deployment.
- **Autonomous agent orchestration**: Orchestrates agents that retrieve, reason, and act across systems, automating complex, high-value workflows with precision, integration, and parallel execution.

Business Impact

SAS delivers organisations a transformative solution for faster, more accurate, and scalable access to unstructured enterprise data, addressing the challenges of manual data sifting and the limitations of traditional RAG approaches. SAS Retrieval Agent Manager is driving measurable impact across industries:

- **Accelerate time to value**: No-code AI and optimised performance enable chatbots and virtual assistants to deliver more relevant responses, boosting customer satisfaction by up to 35%.
- **Enhance trust**: Verified grounding, human oversight, and built-in evaluation improve reliability, achieving up to a 40% increase in content accuracy over traditional AI models.
- **Scale efficiently**: Flexible AI supports multiple LLMs, vector databases, and API integrations, reducing integration effort without requiring system re-architecture.
- **Automate decisions**: Autonomous agents streamline workflows, cutting research time in enterprise applications by 50%.

Benefits

- SAS Retrieval Agent Manager addresses the enterprise “context gap” and the challenge of applying general-purpose LLMs to internal, unstructured proprietary data. By automatically vectorising, chunking, and organising documents into searchable collections, RAM enables accurate, context-aware question answering and content generation. This reduces the need for manual data preparation, complex custom coding, and the maintenance overhead typically associated with traditional RAG implementations.
- Faster Decision-Making: Delivering precise, context-rich answers via a no-code interface, optimised AI, and scalable deployment to accelerate enterprise AI adoption and value.
- Trusted AI Outcomes: Ensuring accuracy and transparency with citation-backed responses, human oversight, and robust security and governance.
- Cost-Effective Scalability: Efficiently scaling AI across use cases with flexible integrations, plug-and-play LLMs and databases avoiding vendor lock-in and supporting broad data compatibility.
- High Operational Efficiency: Automating complex workflows with intelligent, goal-driven agents that seamlessly integrate and collaborate to boost productivity.

Overview

SAS Retrieval Agent Manager provides:

- Intuitive no-code experience: Empowers nontechnical users with a no-code interface to build and manage workflows, with support for diverse document types, multilingual OCR and usage monitoring.
- Trustworthy & transparent AI: Delivers citation-backed, source-grounded responses with built-in evaluation and human-in-the-loop oversight to ensure transparency, traceability, and trust.
- Modular, plug-and-play architecture: Supports multiple third-party LLMs and vector databases with no vendor lock-in, and includes native integration with file systems and Git for seamless enterprise onboarding.
- High-performance & enterprise-ready integration: Delivers fast, efficient AI through model acceleration techniques and offers seamless integration into enterprise systems via robust APIs and secure, scalable deployment.
- Autonomous agent orchestration: Supports orchestration of agents that retrieve, reason and act across systems, automating complex, high-value Agentic AI workflows with precision, integration and parallel execution.
- Flexible deployment options: Provides robust support for on-premises deployments giving organisations full control over their data with easy extension to cloud or hybrid environments.

Learn More

Learn more at SAS® Retrieval Agent Manager