

**Industries:** Real estate

**Region:** USA

## PROJECT TYPE

Enterprise RAG platform

## TECHNOLOGIES

Azure OpenAI GPT-4.1,  
Pinecone, Azure AI Search,  
Python, FastAPI, React

## DURATION

4 months

## METHODOLOGY

Scrum

## TEAM

1 AI Architect  
1 Backend Engineer  
1 Frontend Engineer  
1 Data Engineer  
1 QA Engineer  
1 Business Analyst

# RAG-Based Knowledge Platform for a Commercial Real Estate Operator

An internal AI-powered knowledge platform for a commercial real estate operator managing a multi-property portfolio. The solution unified lease, vendor, maintenance, and compliance documentation into one retrieval layer with citation-based answers and role-based access.

## Project background

The Client is a US-based commercial real estate operator managing a portfolio of office buildings, mixed-use properties, and business parks across several cities.

As the portfolio expanded, critical operational knowledge became scattered across multiple sources. The Client needed to make this knowledge usable in daily work across properties and teams.

## Project Distinctive Features

- ✓ Citation-based answer generation
- ✓ Role-based retrieval and access control
- ✓ Separation of portfolio-wide and local procedures
- ✓ Hybrid retrieval architecture
- ✓ Fallback responses when information is unavailable
- ✓ Support for lease, vendor, maintenance, and compliance documentation
- ✓ Monitoring and answer-quality tracing
- ✓ Web interface for internal operational teams

## Business challenge

The Client wanted to reduce delays in handling tenant, contractor, and property-related requests caused by scattered operational knowledge across multiple systems and teams.

### Additional challenges:

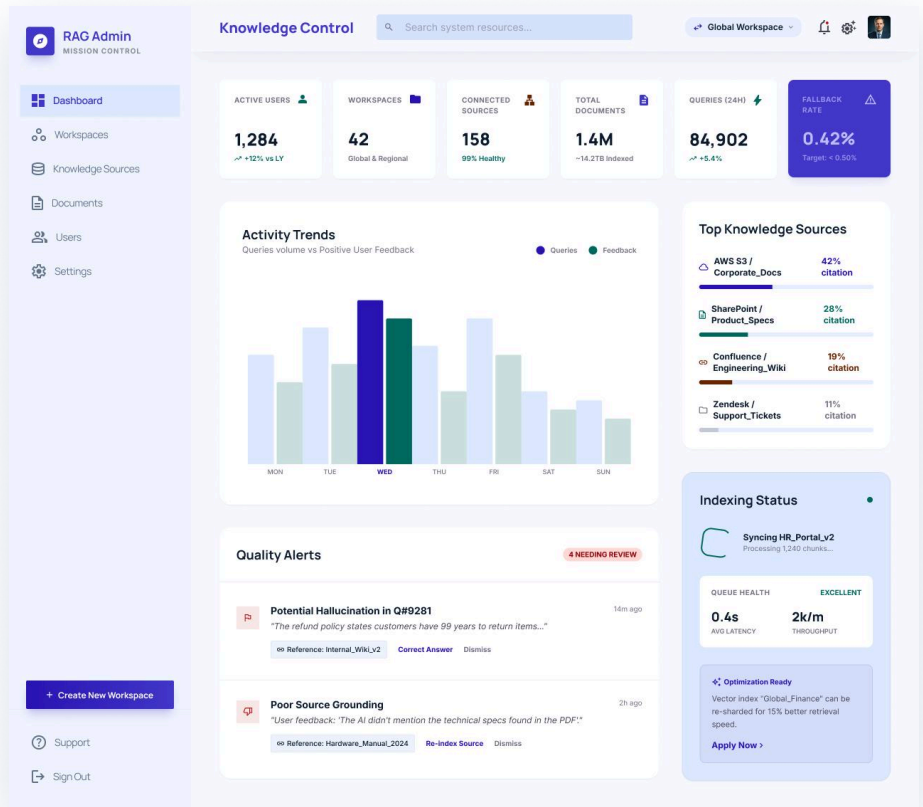
We delivered an AI/GenAI readiness assessment that defined business opportunities, designed the target architecture, and established a clear path from experimentation to implementation.

# Our solution

SumatoSoft developed a RAG-based operational knowledge platform for the Client's internal teams. The solution connected all the company's documentation into one retrieval layer and enabled employees to ask natural-language questions and receive answers grounded in internal documents. The platform facilitates daily operational use across property management, facilities, and support functions, with access controls applied at the retrieval level.



SumatoSoft helped us turn fragmented operational knowledge into a working internal tool that supports our teams every day. The platform gave us a practical way to use AI across multiple properties while keeping information access structured, controlled, and aligned with how our business operates.



## The biggest part – knowledge ingestion and preparation

SumatoSoft began by collecting operational materials from five Client-side groups involved in daily property operations: property management, facilities, tenant operations, compliance, and central operations. We hold online working sessions with responsible stakeholders, followed by targeted follow-up calls and document handover.

At this stage, SumatoSoft focused on the materials actually used in day-to-day work: lease-related documents, vendor and maintenance records, compliance instructions, escalation procedures, and internal operating guidelines. The team reviewed these materials together with the Client, removed duplicates and outdated versions, and clarified which sources were used as the practical reference for different operational scenarios.

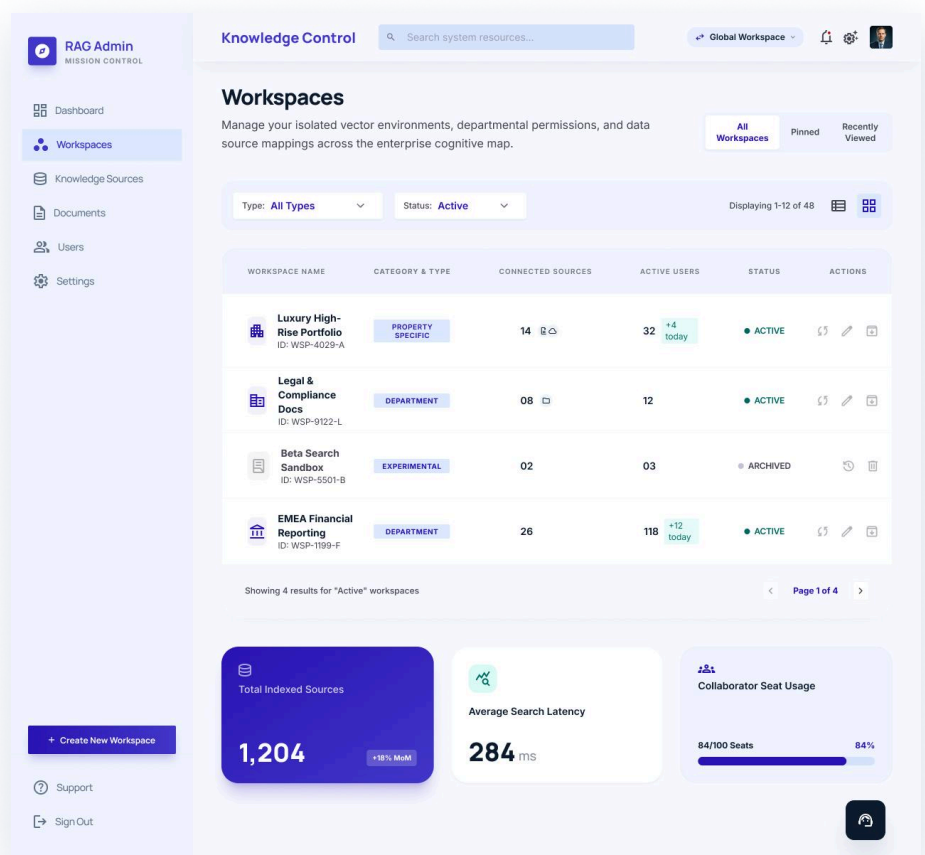
A separate part of the work addressed differences in procedures across buildings. These workflows were not merged into one standard flow, because that would have reduced operational accuracy. Instead, SumatoSoft separated portfolio-wide rules from building-specific procedures and preserved local operating logic where required. As a result, the system was prepared to return answers in the context of a specific property rather than provide an averaged response across the whole portfolio.

The same approach was applied to information owned by different teams. SumatoSoft worked with the Client to define which content should be available for each user role and how access boundaries should be applied before retrieval.

This stage prepared the knowledge base for indexing and created the foundation for property-aware retrieval and grounded answer generation.

## System components

The solution consisted of three main parts: a knowledge preparation and indexing layer, a retrieval and answer generation layer, and a web interface for internal teams. Together, these components turned property-specific operational materials into a working system for daily use across multiple buildings and teams.



## Component 1: Retrieval architecture

To support practical operational use, SumatoSoft implemented a retrieval layer that combined semantic search with structured filtering. The solution used vector indexing through Pinecone and additional search capabilities through Azure AI Search. Retrieval logic was built to distinguish between portfolio-wide rules and building-specific procedures, account for document categories, and apply access boundaries before content reached the model.

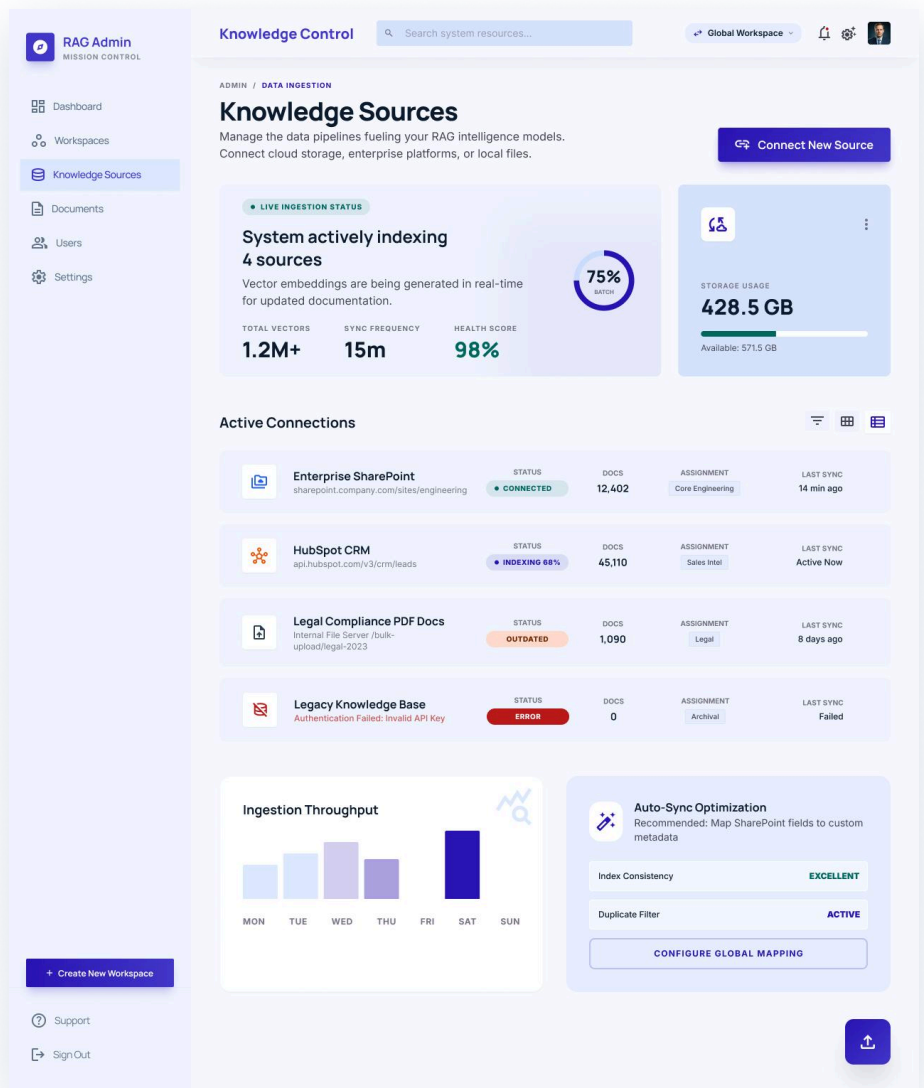
This allowed employees to work through one search flow and receive information relevant to the specific property, topic, and user role instead of manually checking multiple repositories and teams.

## Component 2: AI answer generation

On top of the retrieval layer, SumatoSoft implemented a controlled answer generation flow powered by Azure OpenAI GPT-4.1. The model was used within a grounded pipeline rather than as a general-purpose assistant. Answers were generated only from retrieved internal context and returned with citations, which made the output traceable and suitable for operational use.

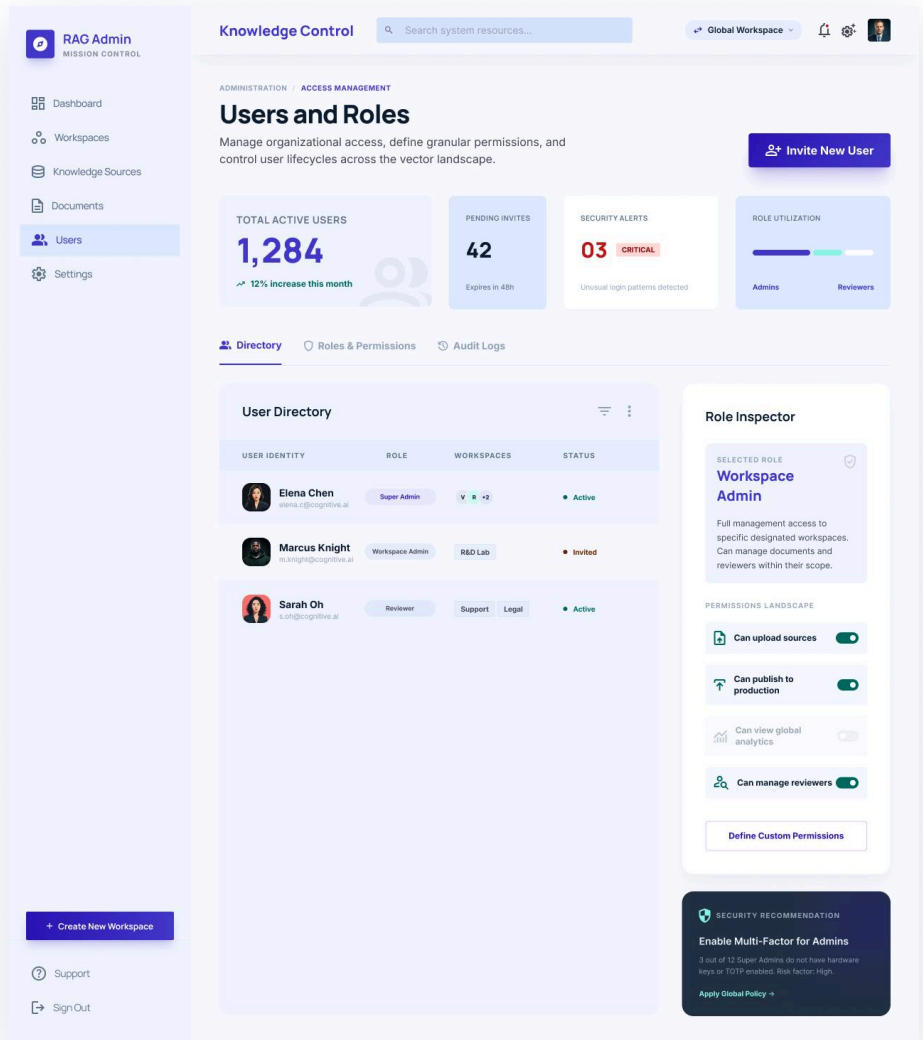
Special attention was given to hallucination control. The system relied on approved retrieved fragments, limited open-ended generation, and returned fallback responses when the available context was insufficient for a reliable answer. This helped reduce unsupported responses in workflows where accuracy and source visibility were critical.

SumatoSoft also addressed token usage as part of production design. Since the platform was expected to support frequent internal requests across multiple properties, retrieval and prompting logic were tuned to keep context focused and avoid unnecessary token consumption. Relevant content was filtered and assembled before generation, which helped keep usage predictable as adoption expanded.



## Component 3: Application layer

The user-facing application was built as a web interface for internal teams. Employees could search operational knowledge, review cited sources, and work with information in the context of a specific property through one interface. As a result, the Client received a system that reduced dependency on manual search and made distributed operational knowledge usable across the portfolio.



## Customer's benefits

The solution was deployed across 18 commercial properties and used by property management, facilities, and support teams through one internal interface. The Client reduced time spent on operational information retrieval by 45% and created a scalable knowledge layer for portfolio-wide operations.