

Azure AI Search vs Azure Cosmos DB

High-Level Comparison for RAG and Graph RAG

1. RAG (Retrieval-Augmented Generation) Comparison

Aspect	Azure AI Search	Azure Cosmos DB
Primary Role	AI retrieval engine	General-purpose NoSQL DB
Designed for RAG	Yes (native)	Partial
Vector Search	Native & optimized	Supported
Hybrid Search	Built-in	Custom logic
Semantic Ranking	Supported	Not supported
Retrieval Quality	High	Medium
Engineering Effort	Low	Medium–High
Best Fit	RAG retrieval layer	Post-processing storage

Recommendation: Azure AI Search is the preferred service for RAG due to higher retrieval accuracy and native Azure OpenAI integration.

2. Graph RAG Comparison

Aspect	Azure AI Search	Azure Cosmos DB
Graph Support	No	Yes (Gremlin API)
Relationship Traversal	No	Yes
Multi-Hop Queries	No	Yes
Provenance Tracking	Limited	Strong
Transactional Updates	No	Yes
Vector Search	Yes	Limited
Best Fit	Text similarity retrieval	Graph RAG & explainability

Recommendation: Azure Cosmos DB is best for Graph RAG where relationship modeling and explainable AI are required.

Final Executive Summary

Azure AI Search is recommended for semantic document retrieval and RAG workloads. Azure Cosmos DB is recommended for Graph RAG, entity relationship modeling, and provenance tracking. Modern enterprise architectures often use both services together.