

# CosmosAIGraph Hybrid RAG Approach

## User Inputs in an example AI Conversation:

1. What is the Python Flask Library
2. What are its dependences
3. What are the alternatives that use async processing
4. Who is the author
5. What other libraries did she write
6. Display a graph of all her libraries and their dependencies

DB RAG  
Graph RAG  
Vector RAG  
DB RAG  
Graph Rag  
Graph RAG



User  
Input



RAG  
Data



Prompt



LLM



Output  
JSON, TXT,  
etc

In-Memory  
RDF Graph  
Database,  
loaded from  
vCore



Cosmos DB  
vCore  
w/Vector  
data and  
search



### Application Logic:

- Determine Intent & RAG Strategy from User Intent
- Identifiy Entities
- Generate SPARQL query if Graph RAG
- Generate vCore query if DB RAG
- Vectorize user input if Vector RAG
- Execute the DB query to get Docs List
- Fetch Documents per List from Cosmos DB
- Craft the Prompt with the Document RAG Data
- Invoke the LLM with input & RAG data in prompt
- Parse the LLM response, and present in the web UI