Project Breakdown:

**Search Engine with the help of AI Agents.**

**Purpose**

Enable natural-language search over enterprise data using **Azure AI Search** and **AI Agent** (powered by Azure OpenAI or rules-based logic) that translates English queries into structured search queries.

**Flow:**  
Intent → Structured Query → Retrieve Results

**Goals**

* **Query Understanding**: Allow users to type queries in natural language (English). Use Azure OpenAI to extract intent and entities.
* **Query Transformation**: Convert the extracted meaning into a structured query format (e.g., OData) suitable for querying Azure AI Search.
* **Query Execution**: Execute the structured query against the indexed data and return relevant results to the user.

**Key Design Decisions**

* **Retrieval Engine**: Azure AI Search
* **Data Sources**:
  + Cosmos DB (SQL API)
  + Azure SQL Database (Synced to Azure AI Search via indexers or pipelines)
* **AI Agent**: Azure OpenAI (for intent understanding and query generation)
* **Observability**: Application Insights and logging for telemetry and monitoring

**Logical Architecture**

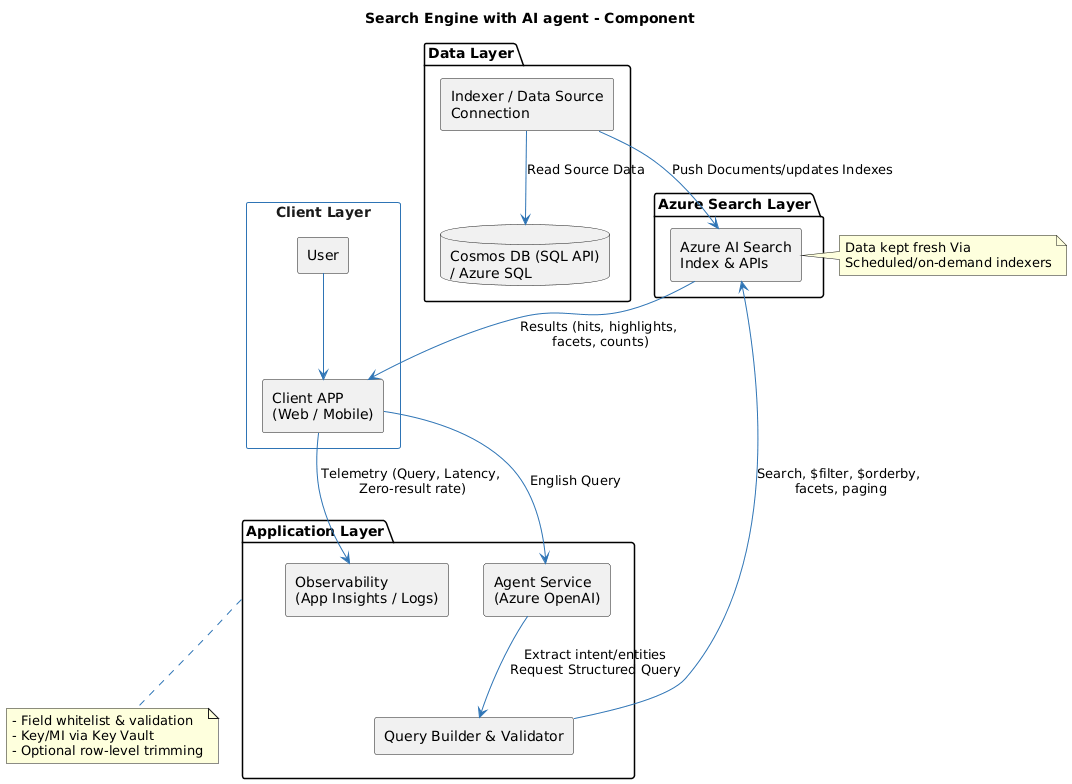
* **Client App**: Web or Mobile interface for user interaction
* **Agent Service**: Azure OpenAI-based service for NLP and query generation
* **Query Builder & Validator**: Converts intent into structured queries with validation
* **Azure AI Search**: Executes structured queries and returns results
* **Data Sources**: Cosmos DB / Azure SQL (source of truth)
* **Observability**: Logs, metrics, and telemetry for monitoring and debugging

**UML Diagrams**

**Component Diagram: Search Engine with AI Agent**

This diagram outlines the **logical architecture** of the system, broken down into four key layers:

* **Client Layer**: Represents the user interface (Web/Mobile) and user interaction.
* **Application Layer**: Contains the AI Agent (Azure OpenAI), Query Builder, and Observability components responsible for understanding queries, transforming them, and monitoring system behavior.
* **Azure Search Layer**: Handles the actual search operations using Azure AI Search Index & APIs.
* **Data Layer**: Includes the data sources (Cosmos DB / Azure SQL) and indexers that keep the search index up to date.



**Sequence Diagram: AI-Powered Search Flow**

This sequence diagram captures the **dynamic interaction** between components during a typical user search:

* The user submits a natural language query via the client app.
* The AI Agent interprets the query and collaborates with the Query Builder to generate a structured query.
* Azure AI Search executes the query and returns results.
* The system logs telemetry data for observability and performance tracking.

A diagram of search process

AI-generated content may be incorrect.