Day - 2

10 December 2025 10:23

Azure Language Service:-

1. Sentiment
2. Key
3. Phrase
4. PII
5. NER
6. Summarization & Translation.

Azure Lang. Service → (Text Analysis) Azure AI service.

Input | document

1. Positive
2. Negative.
3. Neutral,
4. Mixed.

Opinion Mining → which aspect ?.

Room

4-5

At least one line positive, rest is neutral
↓
positive.

At least one line negative, rest is neutral
↓
Negative.

At least one line positive, At least one line negative
↓
Mixed.

All sentence is neutral → neutral.

KPE (Key Phrase Extraction):- Extract the Keyword / Phrase from the text.

I had a great experience with the Azure machine learning service and the customer support team.

Text Normalization

↳ [i, had, a - - -]

Tokenization ["i", "had", "a", "great" - - - -]

Context embedding → Token → Vector

Context embedding

What this sentence is all about?

① POS (Part of speech) :-

great → Adj
experience → Noun
Azure → Proper Noun
Customer support team → Noun Phrase.

② Dependency Parsing :-

Azure + Machine learning
Customer + Support + team.

③ Relevance Ranking :-

AML → High
CST → High

{ Azure Machine learning
Customer Support Team
great Experience.
}

Ex :- "I have a great learning experience in AI-102"

Normalization → i have a great - - -

Tokenization → ["i", "have", "a", "great" - - -]

Token → Vector ID (Numbers)

POS

Adjective → great
Noun phrase → learning experience.
Noun → AI-102

[AI 102
great learning experience]

{ lol → Neutral.
learning → Positive } → Positive

{
AI lol → Neur...
great learning experience → Posit...
}

PII:

Personal Identifical Information

Data that will directly / indirectly identify person.

ex → Name / Phone / email / Address.

(1) Detect PII (where, what)
(2) optionally Redact (mark, remove)

Hi, my name is Rohit Sharma, my Aadhaar number is 1234-5678-9012 and my phone number is 98765 43210. Please block my card ending 1234.

(1) Normalisation to tokenization text
(2) Pattern based detection - Regular Expression
country specific format.

(3) ML Classification.
(4) Merge + filter entity.

Summarization translation

↓ ↓

AI law. AI Translator

Type of Summarization:-

(1) Extractive - Pick the Phrase from sentence.
Abstractive - Recreate the summary.

1. Extractive
2. Abstractive. — Recreate the

"I ordered a refrigerator from ABC Electronics on 5th May. The product was delivered on 7th May but it was damaged from the back side. I immediately contacted customer support and raised a complaint. They promised replacement within 3 days, but no one came. I followed up multiple times through calls and emails but there was no proper response. It has been 15 days now and my issue is still open. I am extremely disappointed with the service."

**Azure openAI:-** Host the openAI solution.

1. Vnet
2. RBAC
3. AD Authentication.

**Model:-**

1. GPT - 4o = (Omni Model)
2. Text embedding Models
3. Whisper (speech)
4. DALL E (Image)
5. Assistant API

"Why did my credit card transaction fail yesterday?"

Normalization + Tokenization

why | did | my | credit | card | transaction | fail | yesterday | ? |

Domain — Banking
Detect Intent → failure
Detect temporary ref. → Yesterday.

**Embedding:-** Convert text into number format. (Vector)

Vector Allows.

1. Semantic Search.
2. Recommendation.
3. Similarity Matching
4. RAG (Retrival, Augmented Generation).

- "Server outage in Mumbai data center"
- "Production system down in India region"
- "I like coffee"

(1) Convert each line into Vector.
(2) Store in Vector DB
  (1) Cosmos DB Vector
  (2) Redis Vector
  (3) Azure AI search.
(3) User Search.
  "Outage in India"

## Assistant API:-

### Component:-

(1) Assitant → AI Brain.
(2) Thread → User Conversation
(3) Message → Each turn.
(4) Run → Execution
(5) Tool → Search, code, API

## Basic Fraud Assistant:-

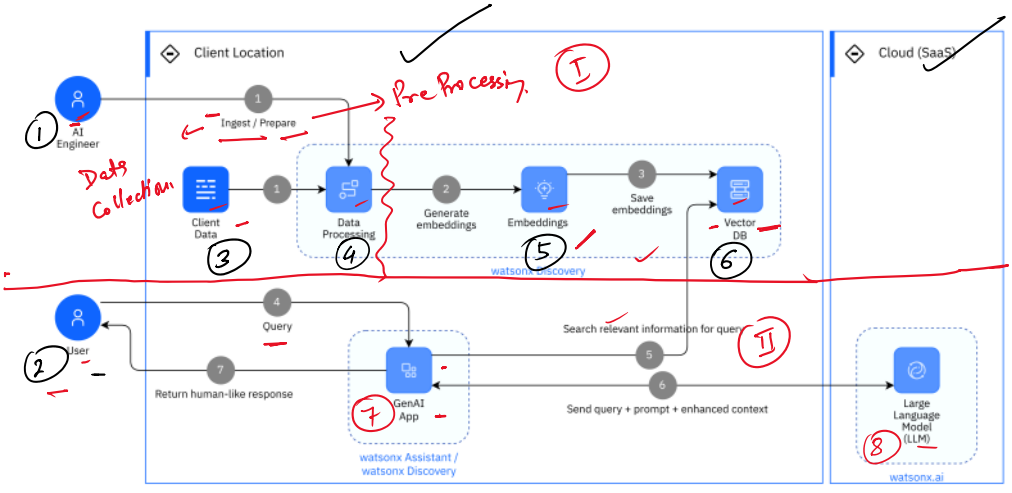
(1) Create Assistant
  Model - GPT - 4o
  tools - (file search, Interpreter)
(2) Upload Data
  (1) Transaction
  (2) Rule Book
  (3) Customer Complain.
(3) Create Thread -
  User - if transaction 1114 is fraud or not
(4) Tool execution
  Assistant:
    Read Transaction file.
    Match fraud pattern
    Python logic
    Rule logic

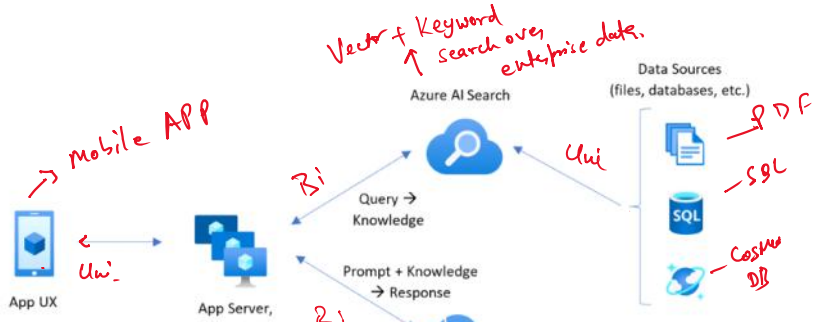
Match jaw
Pythonlogic
Coorelate logic
Generate Result.

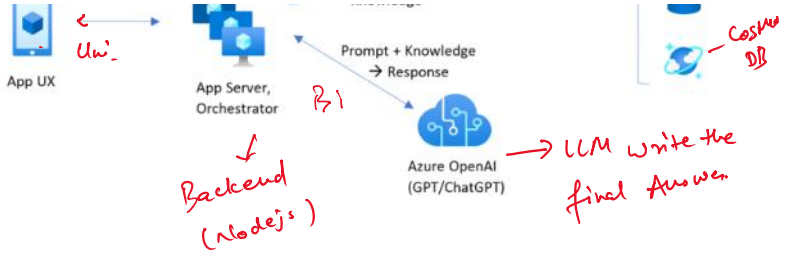
AI + Tool + logic of enterprise Workflow.

## RAG (Retrival Augmented Generation):-

RAG = LLM + External knowledge (Data).






Backend (Node.js)

→ LLM write the final Answer

## Deploy, Monitor & Secure:-

1. Managed Endpoint
2. Scaling
3. Key Vault
4. RBAC
5. Private endpoint
6. Telemetry
7. Drift Detection

## (1) Managed Endpoint:-

No Worry On:-

1. Port
2. Load Balancer
3. certificate
4. VM Scaling

## (2) Scaling:-

100 req — 10,000 req.

- Horizontal Scaling — 2CPU → 2CPU + 2CPU
- Vertical Scaling — 2CPU → 4CPU

Throughput Scaling:-

1. TPM – Token per Minute
2. RPM – Request per Minute

## (3) App Tier Scaling:-

## (3) Vault:-

(1) Azure Key Vault
(2) RBAC - Role Based Access Control
(3) Private Endpoint.

(1.) Azure Key Vault:- Store Certificates, Credentials, tokens, API Key.

(2.) RBAC:- User level Access.

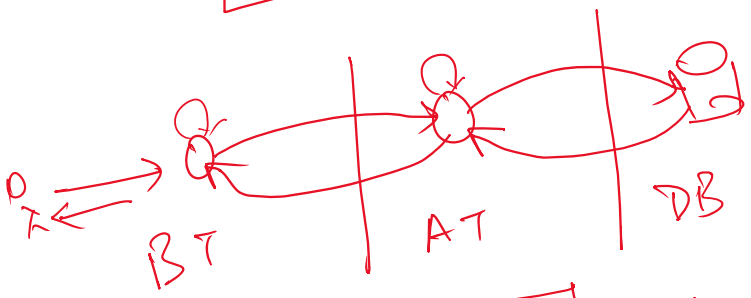
(3) Private Endpoint:- Give Access to the private Vnet.
Zero Internet Access.

## * Monitoring - Telemetry + Logging & Drift.

Telemetry = metrics + logs + Traces.
↳ Nodes.

CPU = 40%.

Timestamp




* Drift:-

- Data Drift — Data coming inside the model will change with time
- Model Drift — Model Behaviour degrades over the time.