

Azure OpenAI Hackathon

Microsoft – Tiger Analytics Partnership

March 2024

Tiger Analytics' coaches introduction



Srikanth Sripada

Lead Data Scientist

5.5 years of experience in leading multiple teams and providing analytical solutions to business problems in QSR, Retail CPG, IoT, and Insurance industries.



Pushvinder Rohtagi

Senior Data Scientist

7.5 years of experience in designing and developing solutions using machine learning ,Deep Learning and NLP.



Abin Joseph

Senior Machine Learning
Engineer

ML/MLOps Engineer with 5 years of experience Cloud, ML Apps and ETL. Developed scalable cloud based solutions for advanced data analytics. Strong working knowledge in Python, Spark, Kubernetes.



Vinay Kumar Soni

Senior Analyst

Data Science specialist with 3 years of hands-on experience, I specialize in crafting end to end NLP solutions tailored to diverse business needs, and empowering businesses to enhance interactions, automate processes, and derive actionable insights.



Rahul Mehta

Tech Presales Manager

Data Engineering, Microsoft Infrastructure and Cloud professional with over 10+ years of experience. Passionate about delivering successful projects, fostering client relationships, and driving business growth.



We are a leading Pure Play Data & AI Services Company

Global Footprint:

USA, Singapore, Philippines,
Malaysia, Australia, Canada, UK,
Mexico, and India

Market Segments:



4,200+



1,000+

Cutting Edge Projects



10% Attrition



80% Repeat Revenue

Microsoft Partnership



Total ~500 certifications.

Innovations on Azure

- Azure Data Fabric
- Data Quality & Profiling Framework
- Data Acquisition Accelerator
- Azure Operations Monitoring
- Azure Cost Analysis & Optimization
- IoT & Streaming Analytics Solution
- Metadata Extractor – Power BI



2021 Partner of the Year Winner
Azure
Intelligent Data and AI



Leader in Analytics and AI
Services Specialists



Leader in the Data
Engineering and Data
Science Services



Our Generative AI Expertise Overview

Current Gen AI Capability Snapshot



30+ Gen AI projects executed



150+ Data scientists and MLEs with experience in Gen AI projects

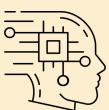


Dedicated Gen AI R&D Lab



60+ ongoing project discussions

Areas of Capability development



Query to insights generation

Custom LLM model development

Intelligent search and summarization

Conversational agent

Content Generation

Some of our current Gen AI client engagements



Notable Gen AI Industry Recognitions

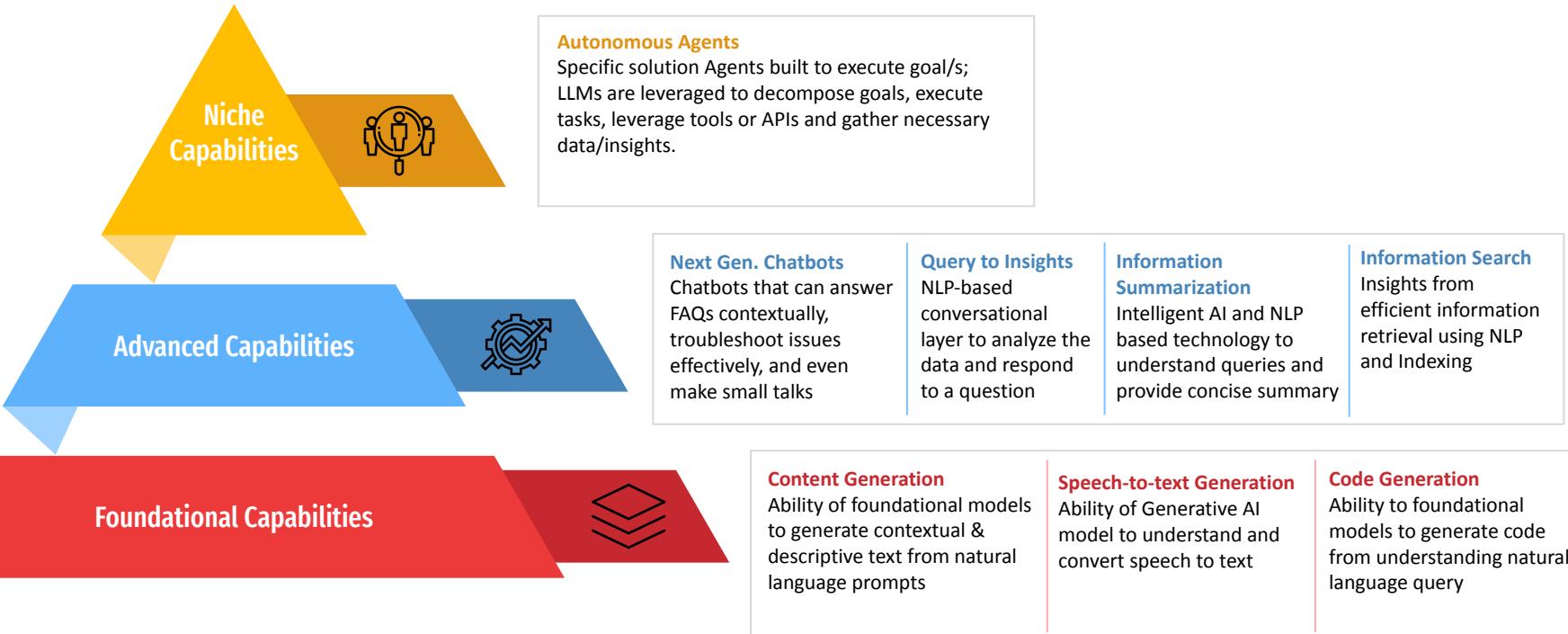


AIM's Top Gen AI Service Providers 2023



Minsky Award for Excellence in AI 2023
Best AI Implementation by a Service Provider

LLMs are enabling a variety of Novel Experiences through these Capabilities



Generative AI Use Cases – Opportunities for firms across various industries

	Retail & CPG	Banking & Insurance	Manufacturing	Health Care	Operations
 Text Summarization	Summarize customer reviews, Sales reports, Social media outlook	Summarize financial reports, Customer feedback, Regulatory compliance reports	Summarize production reports, Quality control data, Maintenance reports	Summarize patient records, Medical research, Health outcome data	Summarize incident reports, Performance reports, Meeting minutes
 Text Categorization	Categorize products, Customer inquiries, Customer reviews	Categorize transactions, Customer inquiries, Customer complaints	Categorize parts, Maintenance requests, Safety incidents	Categorize diseases, Patient inquiries, Medical research	Categorize tasks, Incident reports, Customer inquiries
 Generate New Content	Product descriptions, Marketing emails, Social media posts	Customer emails, Marketing materials and reports	Equipment manuals, Safety procedures, Product catalogs	Medical research articles, Patient education materials, Health newsletters	Process documentation, Training materials, Internal communications
 Language Translation	Multilingual customer service, Translate product descriptions, Customer reviews	Multilingual banking services, Translate financial reports, Customer communications	Multilingual product manuals, Translate safety procedures, Production reports	Multilingual patient communication, Translate medical records, Research articles	Multilingual operations management, Translate process documentation, Training materials
 Generate Code	Automate website updates, POS system updates, Inventory management system updates	Automate banking software updates, ATM software updates, Mobile banking app updates	Automate production line software, Quality control software updates, Inventory management software updates	Automate health record system updates, Hospital management software updates, Medical imaging software updates	Automate system updates, Workflow automation, IT infrastructure updates
 Sentiment Analysis	Analyze customer feedback, Social media sentiment, Product reviews	Analyze customer feedback, Social media sentiment, Customer service interactions	Analyze employee feedback, Customer feedback, Supplier feedback	Analyze patient feedback, Social media sentiment, Patient satisfaction Survey	Analyze employee feedback, Customer feedback, Supplier feedback
 Conversational Agent	Chatbots for Customer service, Sales, Feedback collection	Chatbots for Banking, Financial advice, Customer service	Chatbots for Production line troubleshooting, Maintenance scheduling, Supplier communication	Chatbots for Patient communication, Appointment scheduling, Medical advice	Chatbots for IT helpdesk, HR, Procurement

Gen AI Use Cases by Function

Function/ GenAI Solutions	Intelligent Search and Text Summarization	Business Insights	Conversational Agent	Description Generation
Supply Chain & Procurement	Support negotiations with suppliers by compiling research from publicly available information and historical contracts	Query information from structured databases using textual inputs and generating easy to consume insights	UI supported by natural language-based queries to provide information on the warehouse staff tasks	Summarization of existing documents including contracts; AI generated contract documents in preferred languages
Sales & Marketing	Insights from summarized market research and consumer survey reports	Intuitive descriptive insight generation as an additional layer of existing marketing & sales reports	Chatbot to address sales and customer conversations to improve customer experience	Generate product descriptions from meta-data, marketing content including creative taglines and relevant messaging
Warehouse Operations	Provide summary of the day's activities by reviewing order and delivery pipelines	Textual alerts to the site managers about any anomaly in operations, and labor utilization, fulfillment etc.	UI for site managers to ask questions in natural language to get real-time data on utility & material consumption	
HR	Faster candidate profile filtering through resume summarization	Create summaries from structured databases on workforce productivity at various levels	Chatbot to answer queries from employees	Hypothesis generation on attrition, satisfaction based on employee information
R&D	Enable search on Product information and summarization of R&D insights	Query information from structured databases and generate insights		

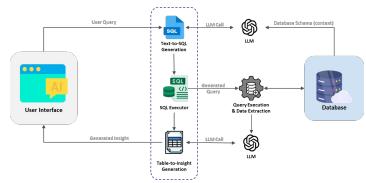
Demos



Gen AI Business Accelerators

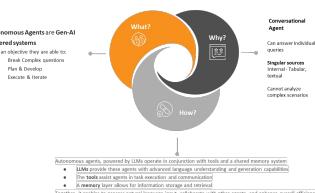
Insights Pro

Natural language interface to generate responses and insights based on user queries. [Demo video](#)



Autonomous Agents

Responds to complex business queries, provides actionable insights and recommends solutions, by analyzing various internal and external data sources. [Demo video](#)



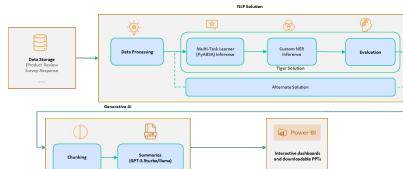
Search & Summarization

Efficient information retrieval solution with summarized insights, increasing productivity, enhancing accessibility, reducing costs and time. [Demo video](#)



Consumer Insights

Analyzes consumer feedback across diverse sources and translates it into actionable insights that aid in understanding consumer preferences and sentiment

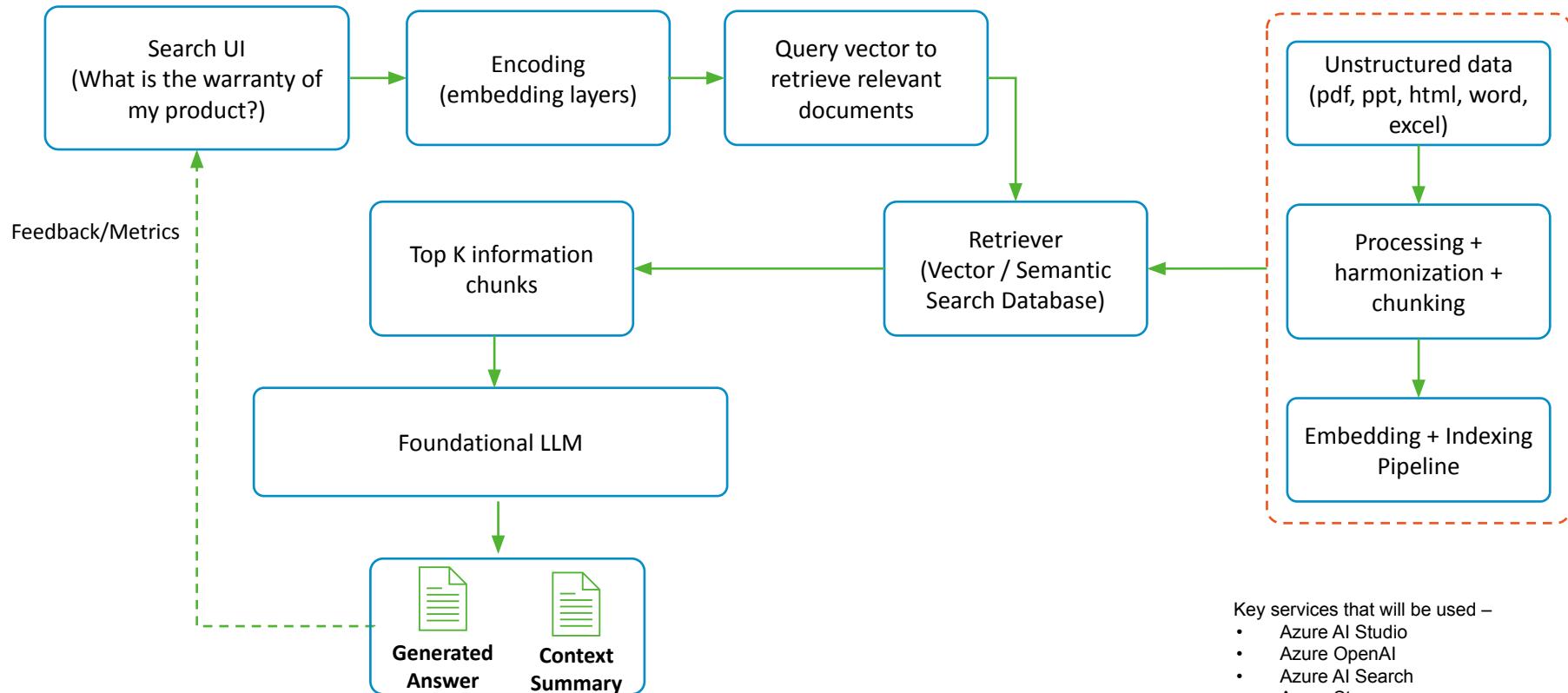


*Click images for more details

RAG



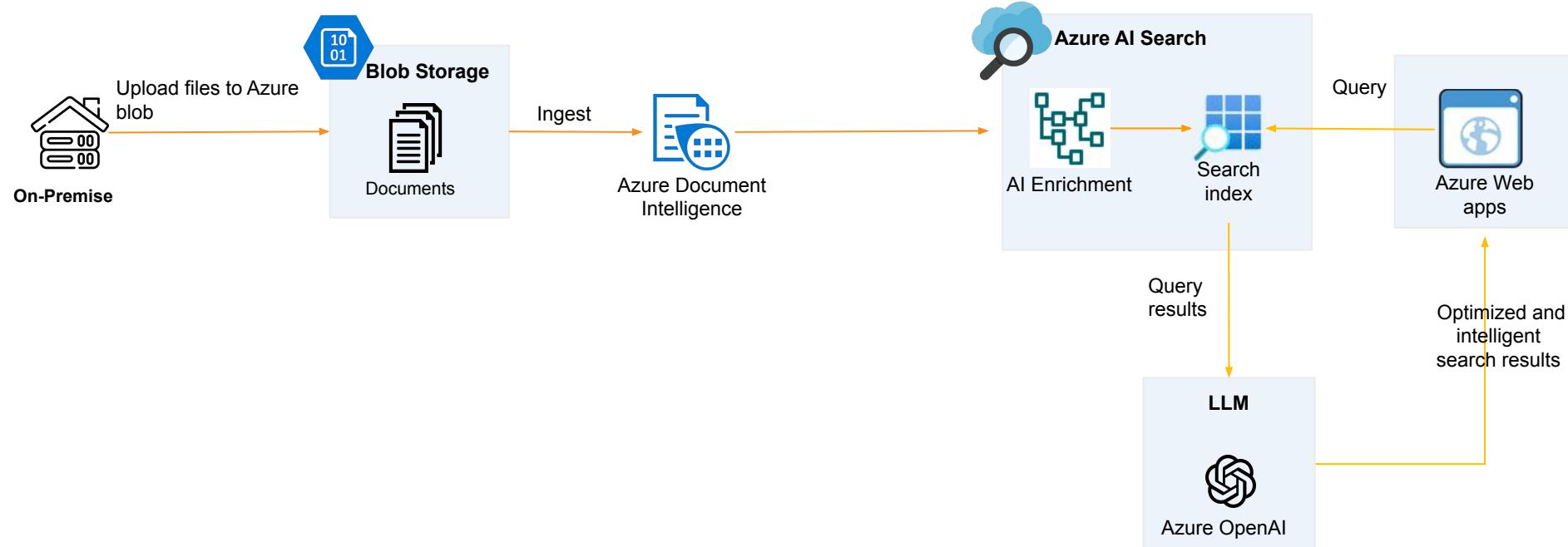
RAG – Solution Approach



Key services that will be used –

- Azure AI Studio
- Azure OpenAI
- Azure AI Search
- Azure Storage

Azure – Reference Architecture Diagram



Usecase-1

CX Copilot

Problem Statement



Objective: Build an AI-powered CX Co-Pilot

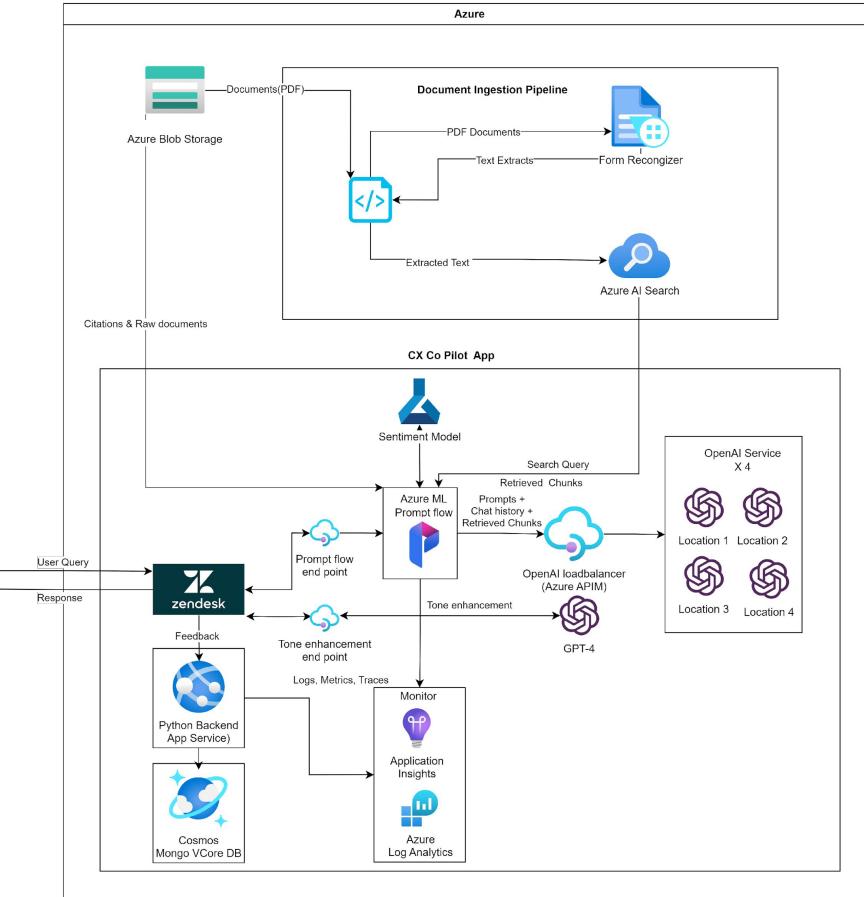


Purpose: Act as an interactive solution for contact center agents



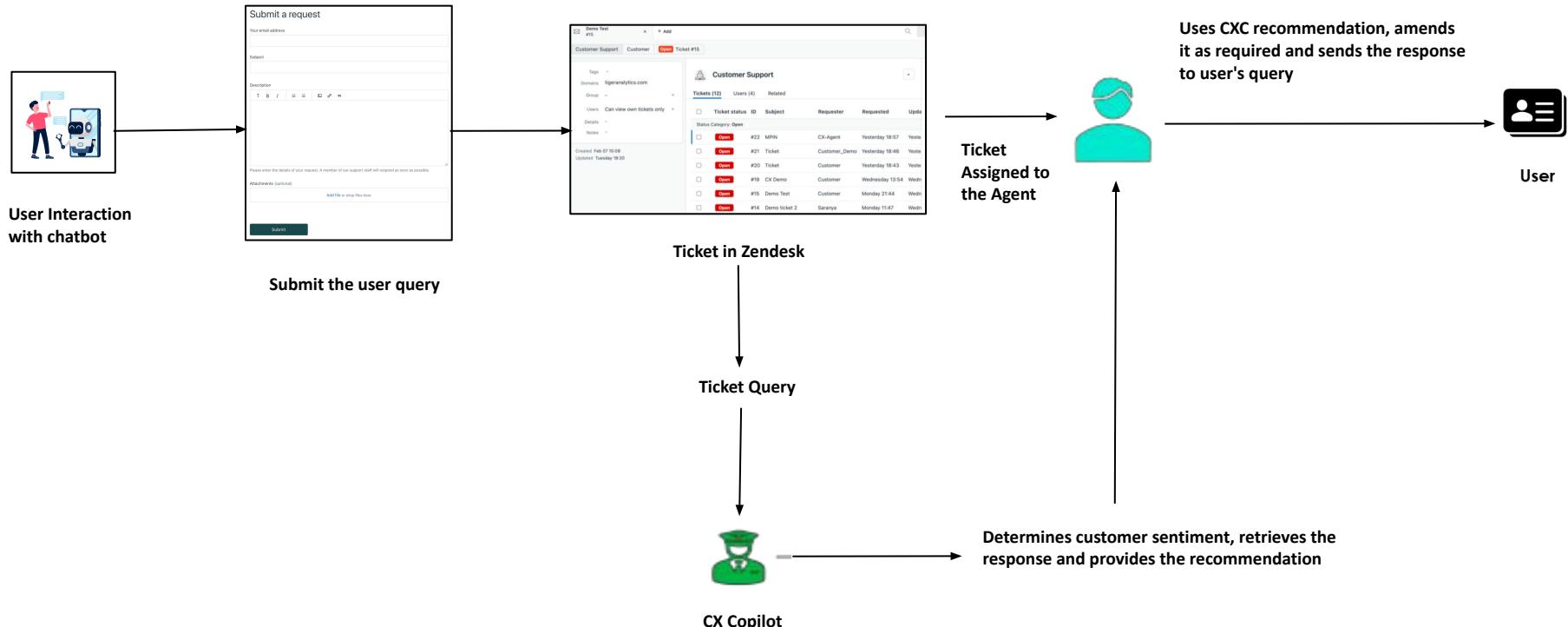
Goal: Reduce the average handling/query resolution time

Architecture



- **Document Ingestion and Analysis:** Utilizes Azure Blob Storage for storing PDF documents, which are then processed by a Document Ingestion Pipeline employing Form Recognizer for text extraction and Azure AI Search for indexing and retrieval.
- **Sentiment Analysis Integration:** Incorporates a Sentiment Model within the CX Co-Pilot App, built on Azure Machine Learning (ML), to assess customer feedback and enhance user experience.
- **Response and Interaction Layer:** Employs Zendesk as the customer interaction platform, backed by a Python backend, with tone enhancement capabilities, to provide informed and empathetic responses to user queries.
- **Database and Logging Mechanism:** Uses Cosmos Mongo DB for data storage, coupled with Application Insights and Azure Log Analytics for monitoring and logging, ensuring robust performance and data-driven insights.
- **Distributed OpenAI Service:** Leverages a load-balanced OpenAI Service with multiple locations to efficiently handle search queries, using GPT-4 for generating prompts, chat history, and tone-enhanced text responses.

Workflow with CX Copilot



Customer interacts with a rule based chatbot

Fills the request form

A ticket is created and assigned to the agent

Query reaches the CX Copilot- It searches for the relevant information in the knowledge base

Based on the sentiment of the user query , it formats the recommended response

The Agent can now amend the response and reply to the user

Components of User Interface

Intent Detection

Intent: I want to delete my Cash Account

Sentiment



Negative Neutral Positive

Conversation Summary

The user doesn't want to use his Cash account anymore. However, he wants to cash out his balance first before proceeding to delete his account.

The user also forgot his MPIN and needs assistance. He also wants to investigate the fraud incident happening in his account based on his other ticket under ticket #3990.

Support Procedure

The customer was able to provide the basic information and attached copy of his ID to verify his level 1 account. Upon checking:

- The provided name, home address, email address, date of birth, and the current balance matches with the KYC Portal using the provided mobile number
- The provided date of the last transaction also matches based on his transaction records in A+

Source: MPIN Reset Wiki

Recommended Response

This response is AI generated. Please make sure it is accurate and appropriate before sending

Hello Dan,

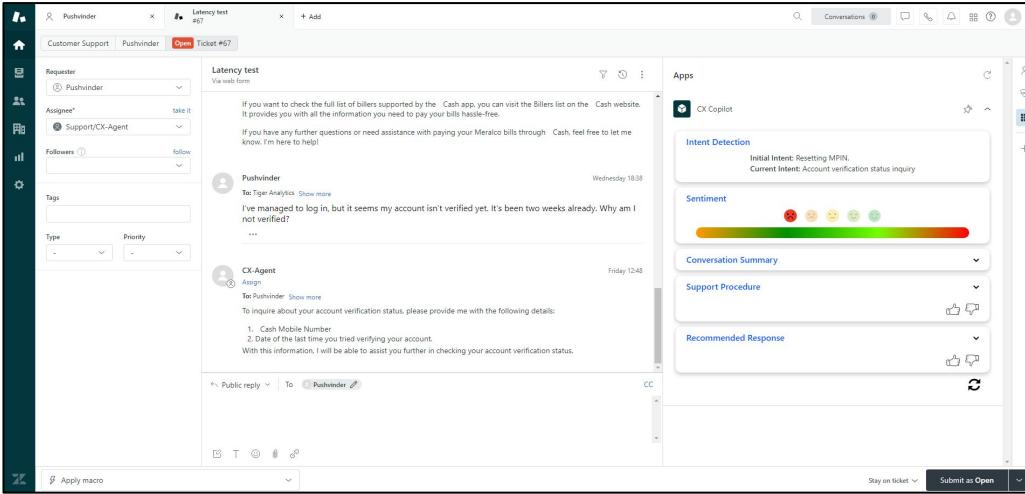
We have received your request for MPIN Reset and will now proceed to send the system-generated temporary MPIN. Upon receiving the temporary PIN, please log in and immediately change to your preferred MPIN details. It is important that you fully verify your account so that you can now do a more convenient and safe way of resetting your MPIN through our face verification

Send to Customer >

UI Features

- Detect conversation intent
- Point in time sentiment
- Conversational sentiment
- Conversation summary
- Support Procedures grounded on Knowledge base for the Agent
- Track similar tickets by the same customer
- Recommend response for the customer
- Fine-tuning responses based on customer sentiments.
- Edit response button
- Enhance response by formatting it in tones such as formal, friendly and neutral
- Ability to send the responses to customer
- Capturing response feedback
- Capturing procedure feedback

Sample UI



This screenshot shows the 'Recommended Response' section of the CX Copilot. It includes a 'Conversation Summary' and 'Support Procedure' summary. Below is a generated response: "Hello! Thank you for reaching out to inquire about your account verification status on GCash. I can help you check that right away. Please let me know if you have received any notifications regarding your verification on the app. Additionally, could you also tell me when you last tried verifying your account? If it has been more than 7 days, we will take further steps to see what's happening with your verification process. Rest assured, we strive to ensure a smooth experience for you." At the bottom are buttons for 'Send to Customer', 'Edit Message', and 'Enhance'.

This screenshot shows the 'Support Procedure' section of the CX Copilot. It provides instructions for checking account verification status: 1. Inquire if the user has already received a notification regarding their verification status. 2. Check the date when the user last tried verifying their account and obtain any relevant screenshots, such as the verification status prompt on the app. 3. Advise the user to check the GCash app for any updates on their verification status. If more than 7 days have passed since their last verification attempt, proceed to the next step. 4. Use the Jarvis portal to track the user's verification application status using the GCash mobile number provided by the user. 5. Depending on the status found in the tracker, follow the next action protocol: - If the status is "FOR APPROVAL" and there is no ongoing application found, escalate it to KYC Support - L3 for application follow-up. - If the status is "APPROVED", inform the user that their verification has been successful and they should be able to see their status updated on the GCash app.

Usecase-2

Wealth Coach

Problem Statement



Objective: Build an AI-powered Wealth Coach

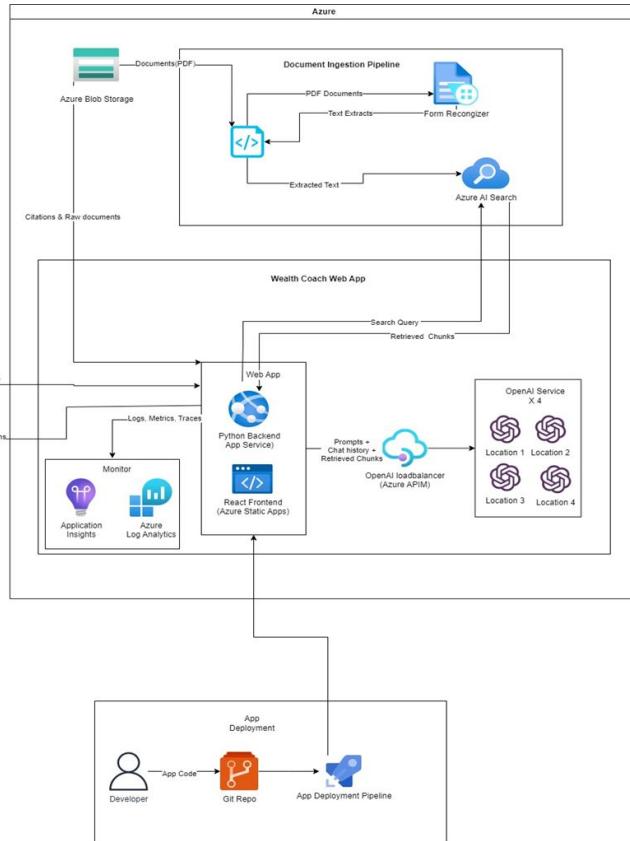


Purpose: Act as an interactive solution to inform customers about various wealth creation strategies and financial products.



Goal: Improve customers' financial literacy and decision-making by providing personalized, AI-driven advice

Architecture



- **Document Ingestion and Processing:** The architecture uses Azure Blob Storage to house PDF documents processed through a Document Ingestion Pipeline, with Form Recognizer for text extraction and Azure AI Search to facilitate quick retrieval of extracted text for citations and raw documents.
- **Wealth Coach Application Interface:** The Wealth Coach Web App acts as the user interface, receiving user queries and utilizing the backend services to fetch and deliver search query results and retrieved chunks of information.
- **Distributed AI Service:** Incorporates a distributed OpenAI Service configured across multiple locations, using GPT-4 to process prompts and chat history and delivers responses to the Web App through an Azure load balancer and API Management (APIM) service.
- **Monitoring and Analytics:** Continuous monitoring, capturing logs, metrics, and traces with Application Insights and Azure Log Analytics ensures high availability, performance, and proactive troubleshooting.
- **Continuous Integration and Deployment:** Utilizes a Git repository for version control of the application code and an automated App Deployment Pipeline, ensuring streamlined updates and releases of the Wealth Coach Web App.

Sample UI

The screenshot shows a user interface for a 'Wealth Coach' service. At the top, there's a blue header bar with the text 'Wealth Coach' on the left, 'Chat Ask a question' in the center, and 'Azure OpenAI' on the right. Below the header, there are two buttons: 'Clear chat' with a trash icon and 'Developer settings' with a gear icon. The main content area features a large title 'Wealth Coach' and a subtitle 'Ask anything or try an example'. Three cards are displayed below the subtitle, each containing a question: 'What are the limitations of technical analysis?', 'How is the interest in CIMB calculated?', and 'Can you work out only the interest that i will earn?'. At the bottom, there's a text input field with placeholder text 'Type a new question (e.g. What are dividends?)' and a blue arrow button to its right. A link 'Refer here for more use cases' is located at the bottom right.

Wealth Coach

Chat Ask a question Azure OpenAI

Clear chat Developer settings

Wealth Coach

Ask anything or try an example

What are the limitations of technical analysis?

How is the interest in CIMB calculated?

Can you work out only the interest that i will earn?

Type a new question (e.g. What are dividends?)

Refer here for more use cases

Prerequisites



Please refer to [this link](#) for setting up the prerequisites

Hackathon



- Please refer to [this link](#) for instructions on building your AI project in Azure AI studio
- Please refer [this link](#) for setting up the project using Azure AI SDK

Prompt Template

Instruction :

You are an AI assistant designed to help customer support agents. You help people find information about the various customer queries .You also generate citations for every fact you state from the retrieved documents in square brackets like [info1.pdf].

Guidelines :

- Identify the issue the user is facing based on the user question and conversation history.
- You should identify the steps to resolve the query.
- Classify the queries into two categories: "Customer Support Procedure", and "Recommended Response".

Question :

How do I rest my MPIN ?

Context :

Click on the entry to open the customer's profile.\nStep 3. Click on RESET MPIN on the bottom right corner.\nStep 4. Confirm to successfully reset the customer's MPIN. After resetting the MPIN, you will be seeing this in the audit log of the page. (see Notes > Approvals) Take note of the reference code generated here as you will be needing to input this in the Zendesk ticket.\nStep 5. In the Zendesk ticket, ensure to tick the [Agent] MPIN Reset in the mandatory Zendesk fields and provide the [Agent] MPIN Reset Code you got from Step 4.\nStep 6. The customer will receive a text message with a temporary MPIN.



- **Instructions** guide the LLM to perform a specific downstream task.
- We can add **guidelines** or rules to follow while generating a response.
- **User Query** for which the LLM should generate a response.
- **Context** acts as the reference from which the LLM should generate response.
- The context can be a user input or may be from a retrieval system (in case of RAG systems)
- If the answer is not present in the context the LLM should return a response containing "The answer is not present in the context". This can be included as part of instruction.

Thank You

Do you have any questions?

Tiger Analytics India Consulting Ltd

tigernlp@tigeranalytics.com

www.tigeranalytics.com



About Us

Tiger Analytics is pioneering what AI and analytics can do to solve some of the toughest problems faced by organizations globally. We develop bespoke solutions powered by data and technology for several Fortune 500 companies. We have offices in multiple cities across the US, UK, India, and Singapore, and a substantial remote global workforce.

We have received multiple awards ranging from being recognized as a Leader by Forrester Research to being ranked among the fastest-growing tech companies by Inc. and Financial Times. We consistently feature in prestigious 'Best Analytics Firms' lists.

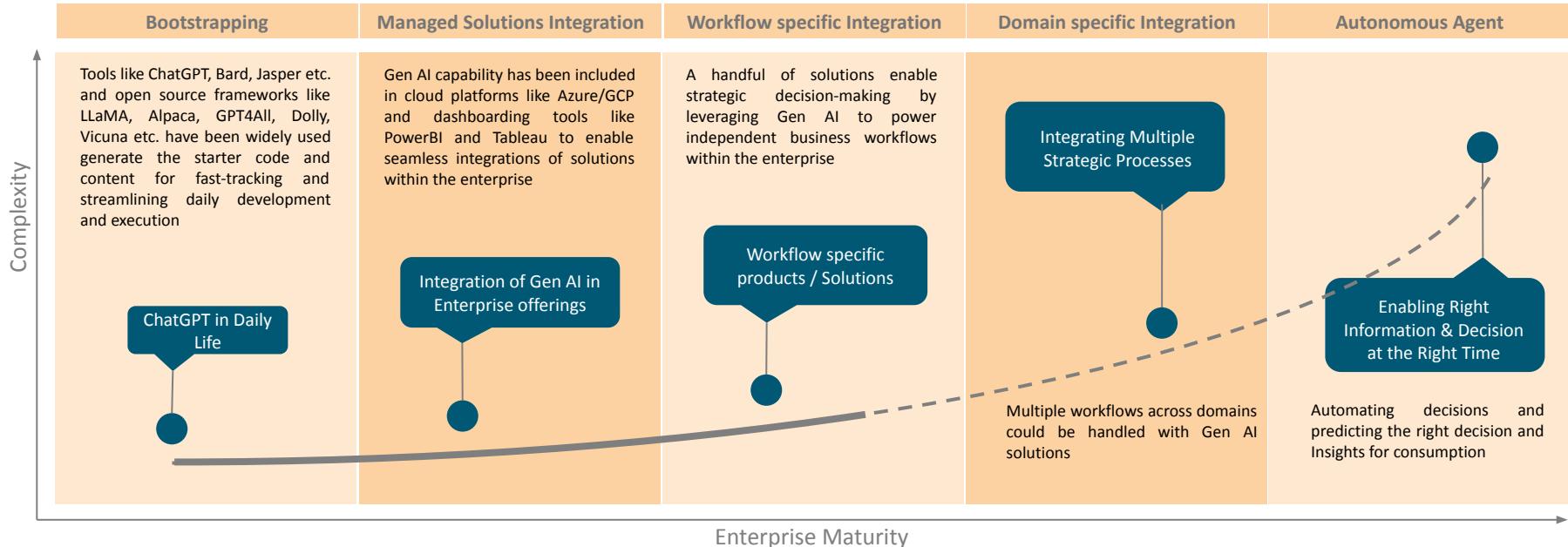
www.tigeranalytics.com



A

Appendix

Our Perspectives are based on our work with various clients

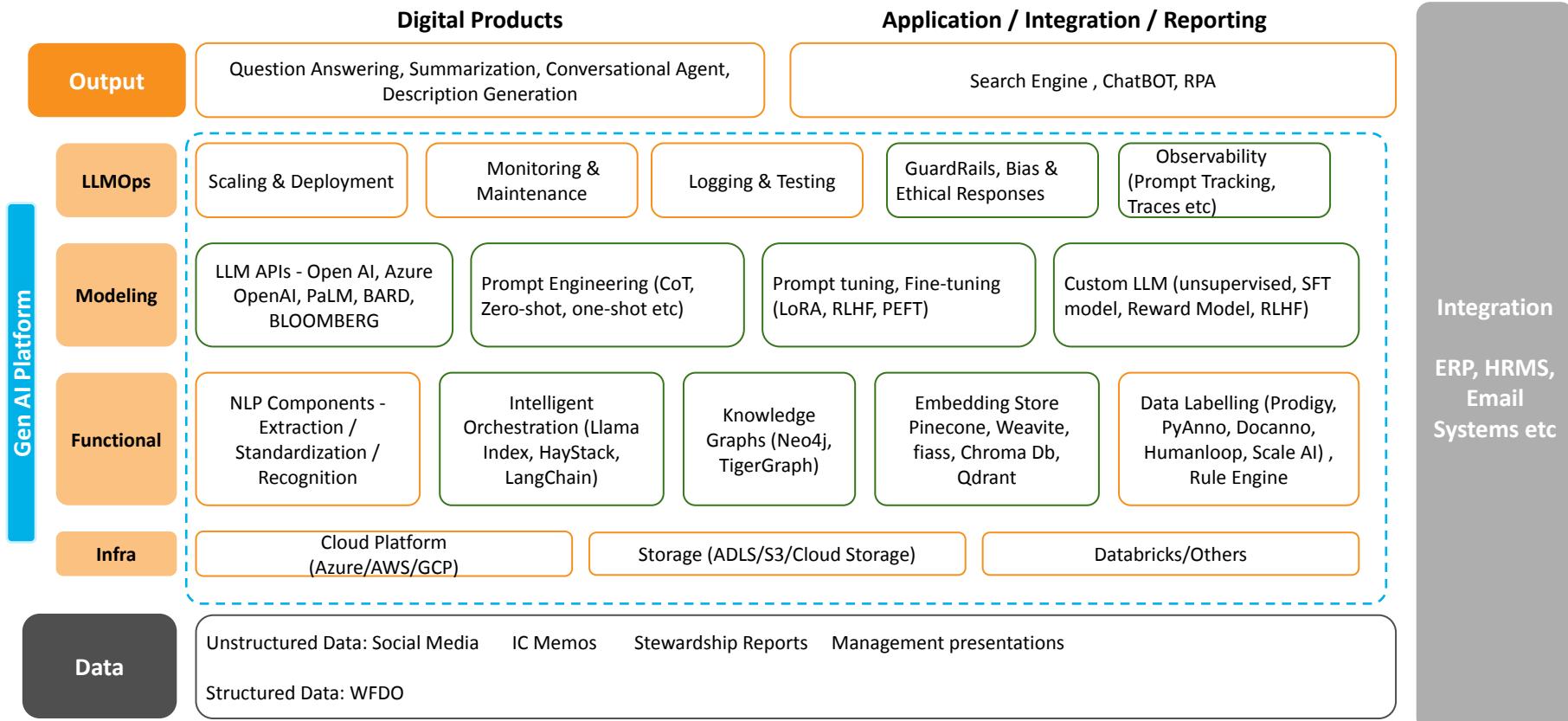


Eg: Code and Content Generation	Eg: Model Marketplace and LLMops, Tool Integrations	Eg: Pricing specific models to enable insights/access	Eg: Retail specific models leveraging cross-data access	Eg: Pricing Gen AI engine working with various Analytics tools to generate insights
---------------------------------	---	---	---	---

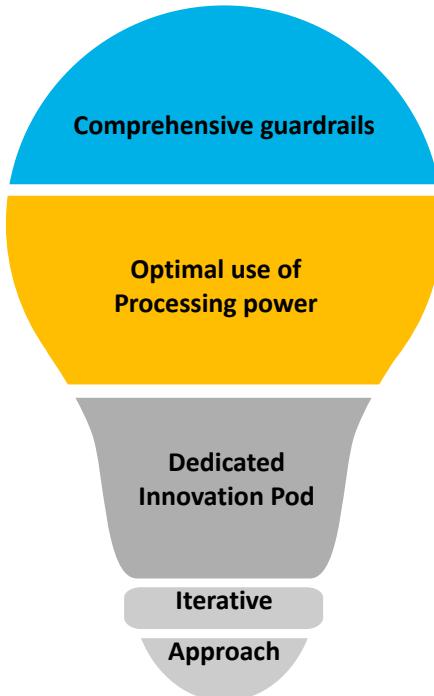


Frameworks will evolve around the Core principles of Responsible AI

Gen AI – Reference Functional Architecture



A robust approach is required to lay a solid foundation for Gen AI led innovation



Goals

To ensure the outputs provided by Gen AI overcome the known issues such as Silent Hallucinations

To invoke LLMs only when necessary and prioritize use of Frequently Cached Queries and Q&A Databases

To support development team with the latest research, innovations and releases in Gen AI space

To frequently take user feedback to evaluate and improve insights

Business Impact

Robust outputs to drive confidence among users thereby increasing adoption

Significant reduction in processing cost and time

Ensuring the most up-to-date solution is deployed without any impact on release schedule

Allowing to deliver a user centric solution that is easy to use and adapt

Gen AI Accelerators and Enablers

Retail & CPG

Customer Service –
Conversational agents for product recommendations, product and store queries

Marketing - Analyze customer feedback, Social media sentiment, Product reviews

Banking & Financial Services

Lending – Virtual assistant to address credit and risk related queries for CRO and generate insight

Accounting – Summarize financial and Regulatory compliance reports

Insurance

Pricing & Underwriting –
Patient admission history analysis and decisioning through natural language

Claims – Intelligent search & summarization reducing claims leakage

Healthcare & Pharma

Patient Journey –
Summarizes patient records, health outcome data to present accurate patient Journey assessment

Drug R&D – Analyzes scientific papers, research articles, and technical docs to provide insights and trends

Manufacturing

Supply Chain & Operations
– Summarized production reports, quality control data, maintenance reports for optimized planning & logistics

Product R&D – Faster access to insights on engineering design, tools and information from large corpus to documents

BUSINESS ACCELERATORS

Insights Pro

Natural language interface to generate responses and insights based on user queries

Autonomous Agent

Responds to complex business queries, provides actionable insights and recommends solutions, by analyzing various internal and external data sources

Search & Summarization

Efficient information retrieval solution with summarized insights, increasing productivity, enhancing accessibility, reducing costs and time

Consumer Insights

Analyzes consumer feedback across diverse sources and translates it into actionable insights that aid in understanding consumer preferences and sentiment

DATA ENGINEERING ENABLERS

Simplified Data Interaction

Chat2Viz & Chat2Data

Productivity

Data and Code Generation

Log Analytics

Anomalies and Trends

FinOps

Cloud Usage and Cost analysis

DataOps

Metadata Search, Security, Access and User Activity Insights

Accelerated SDLC using Gen AI

A1

Gen AI Use Cases

Gen AI Use Cases across Retail Value Chain

Marketing & Promotion

Improved sales agent effectiveness with intelligent search & summarization

Assortment

Product generation through text and image generation

Supply chain & Inventory

Designing product labels
Through text and image generation

Operations

Customer Assistant
through improved chatbot

Automated documentation
through description generation

Consistent product description
to maintain brand uniqueness
through text generation

Synthetic inventory data generation through text generation

Improved inquiry resolution by agents with intelligent search & summarization

Personalized promotions
through text and image generation

[Back](#)

Gen AI Use Cases across Insurance Value Chain

Marketing & Distribution

Improved sales agent effectiveness with intelligent search & summarization

Pricing & Underwriting

Automated information gathering and data entry through faster search

Claims

Automated information gathering and data entry through faster search

Operations

Self-serve customer queries through improved chatbot

Better explanation of policy/products through description generation

Enhanced risk assessment through intelligent search & summarization

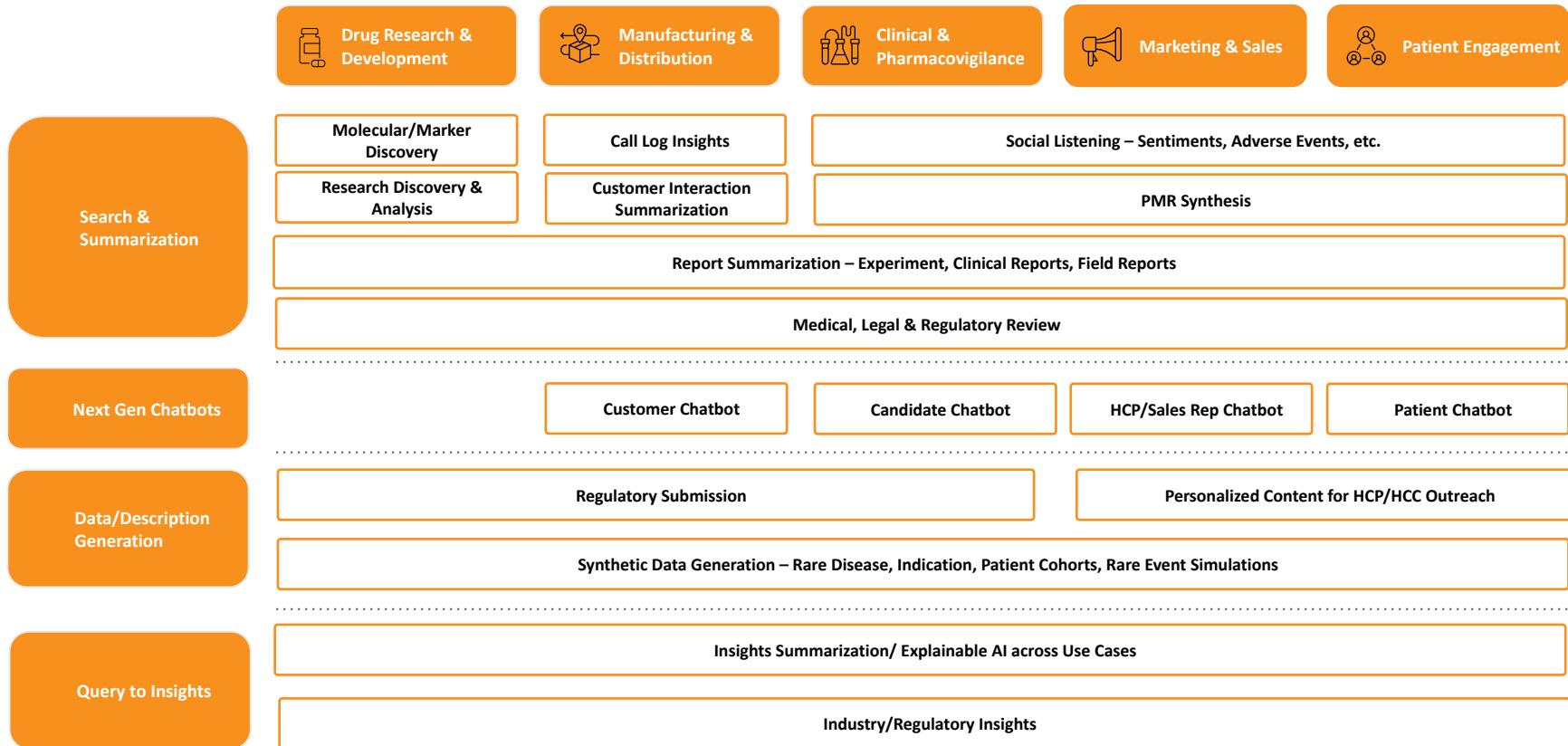
Reduction in claims leakage through intelligent search & summarization

Improved inquiry resolution by agents with intelligent search & summarization

Automated documentation through description generation

Automated documentation through description generation

Gen AI Use Cases across Pharma Value Chain



[Back](#)

Gen AI Use Cases across Pharma Commercial Ops

Use Case	Details	CRM (Veeva)	Syndicated	Models /Insights	Digital	Other Internal	External
		<ul style="list-style-type: none"> • HCP • Sales Rep • Interaction • Emails/Campaign • IC/Call Planning 	<ul style="list-style-type: none"> • Claims/RWE • Formulary • Distribution • EHR 	<ul style="list-style-type: none"> • Sales Forecasting • Promo Response • Segmentation & Targeting 	<ul style="list-style-type: none"> • Social Media • DCM • Paid Search • Organic Search 	<ul style="list-style-type: none"> • Marketing Content • Primary Research • Industry Report • Training • Speaker Program 	<ul style="list-style-type: none"> • Web Population • Medical Conditions • CMS
Social Listening	Real time summarized Insights from Web/Social media to understand brand sentiment, adverse events, quality deviations, competitor insights etc.				✓		✓
PMR Synthesis	Synthesis of Patient Medical Records to gain key patient insights like patient journey, efficacy, adverse reaction, etc.		✓	✓			
Report Summarization	Get answer to specific questions by summarizing Primary research, Industry reports, etc.					✓	✓
Medical, Legal, Regulatory Review	Instant review marketing content, messages, data for any potential medical , legal or regulatory risks like PHI, Competitor info	✓	✓			✓	
Chatbot	Next Gen Chatbot for HCP, HCC or Sales Rep to provide precise information or recommendations	✓	✓	✓	✓	✓	✓
Personalized content	Recommending personalized content for each HCP for outreach/campaigns based on HCP preference, interaction, etc.	✓	✓	✓	✓	✓	
Synthetic Data Generation	Overcome compliance, privacy challenges & improve patient outcomes by generating synthetic claims/RWE, distribution, etc. and using for ML model training		✓	✓	✓		
Insight Summarization/ Explainable AI	Summarize insights & provide explain recommendations based on outputs from other AI use cases like Sales forecasting, Promo Response models, segmentation, etc.			✓			
Industry/Regulatory Insights	Provide industry or regulatory insights based on internal/external reports, web, etc.					✓	✓

Back

A2

Gen AI Business Accelerators



Insights Pro – Insights Generation through Generative AI

OBJECTIVE

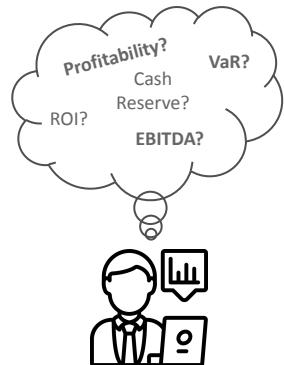
- Build Generative AI based solution to generate responses and insights based on user queries

FUNCTIONALITY

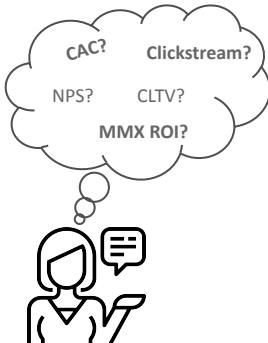
- Use data from internal documents to answer user queries through Gen AI
- To reduce the manual efforts to extract answers a plethora of dashboards/perform ad hoc analysis

On Demand Insights to Key Business Questions

Query To Insights



Finance



Marketing



Manufacturing



Human Resources



Enterprise Data Lake



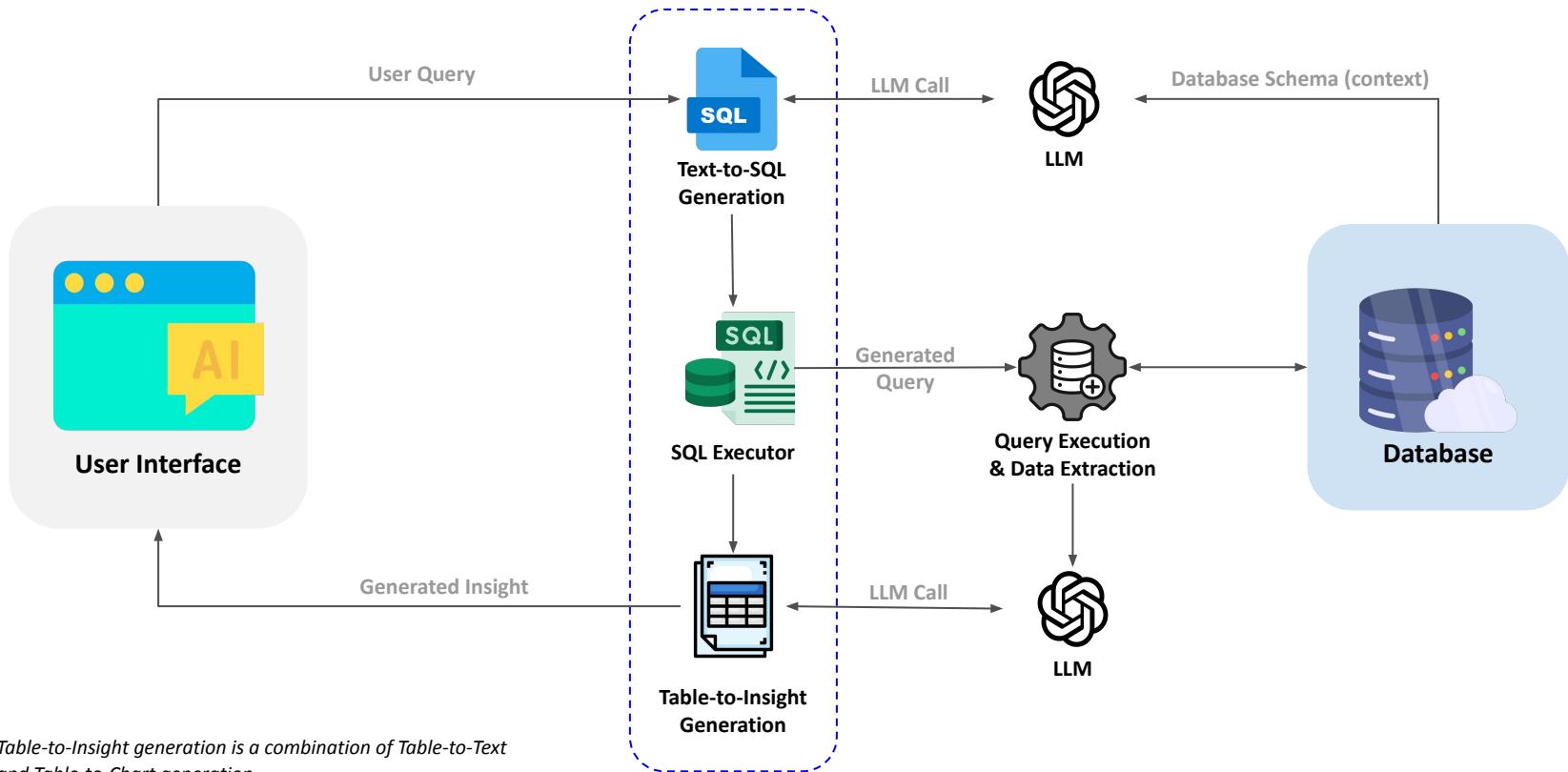
Third Party APIs



Text/Image/Voice

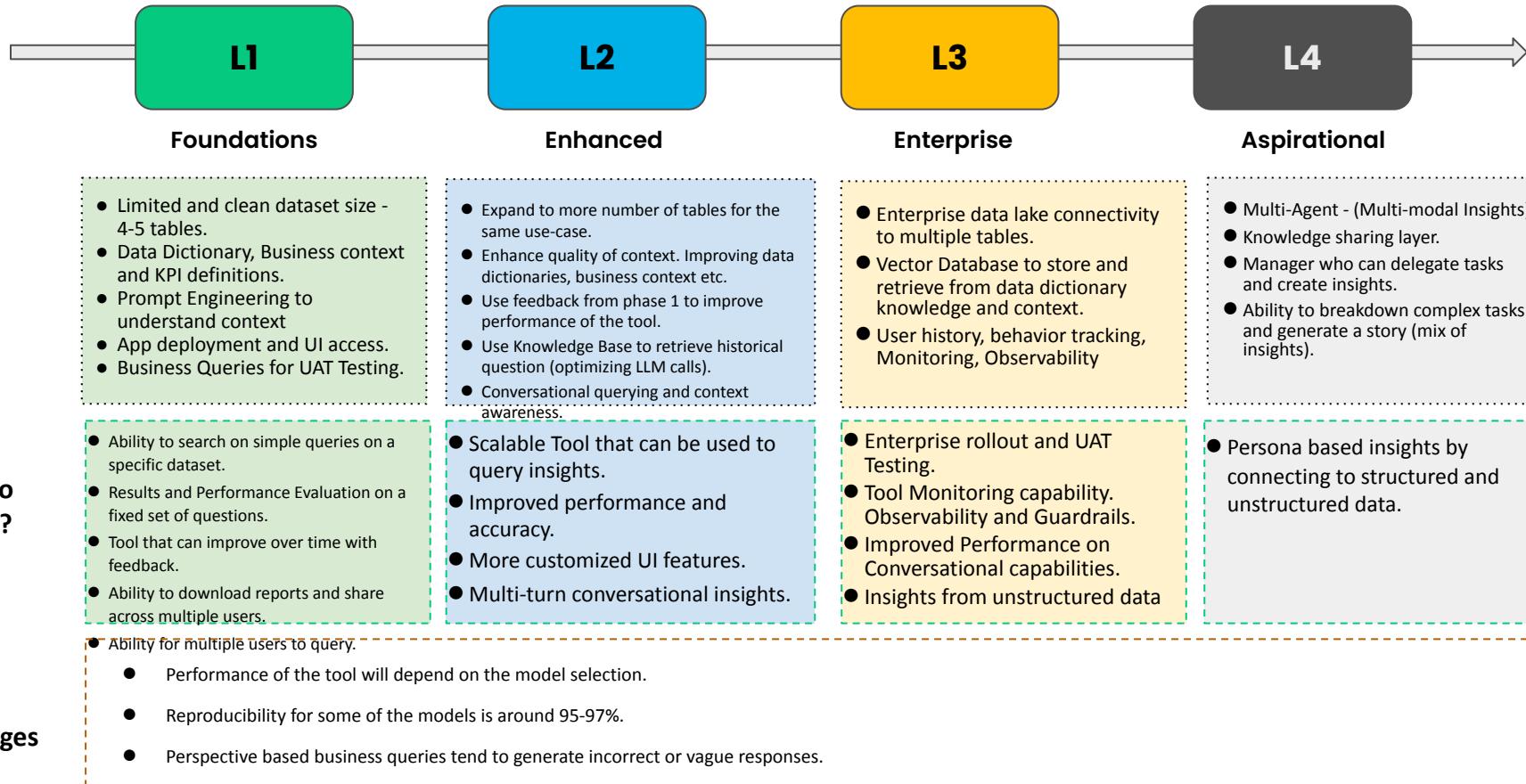
[Back](#)

Functional Architecture



- *Table-to-Insight generation is a combination of Table-to-Text and Table-to-Chart generation*
- *Table-to-Chart generation involves code generation which is executed on the table to generate the required chart*

Roadmap to Value Realization



Search & Summarization

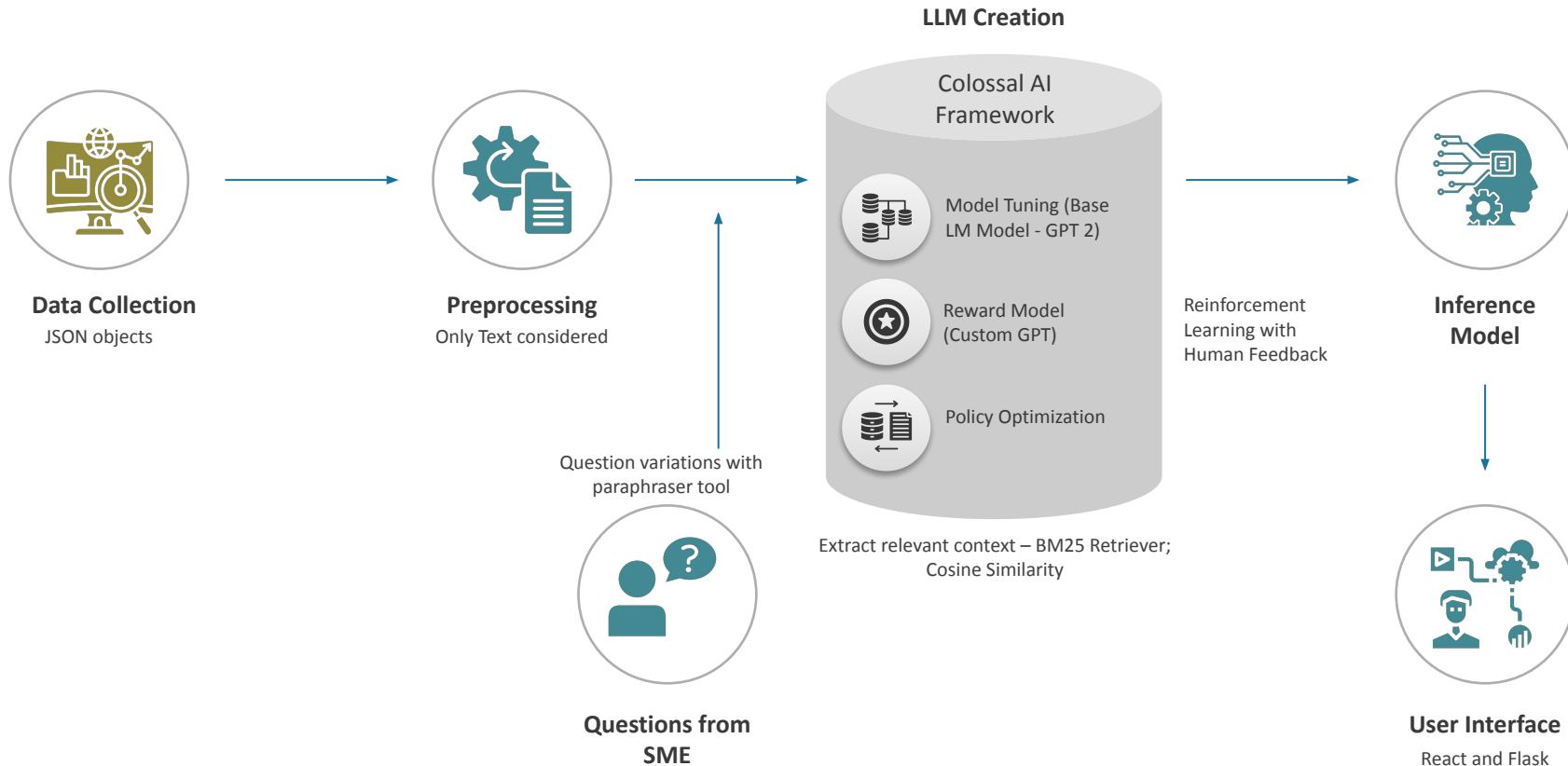
OBJECTIVE

- Instant accurate information retrieval from any input query anytime & anywhere
- To reduce the manual efforts, time and operational cost

FUNCTIONALITY

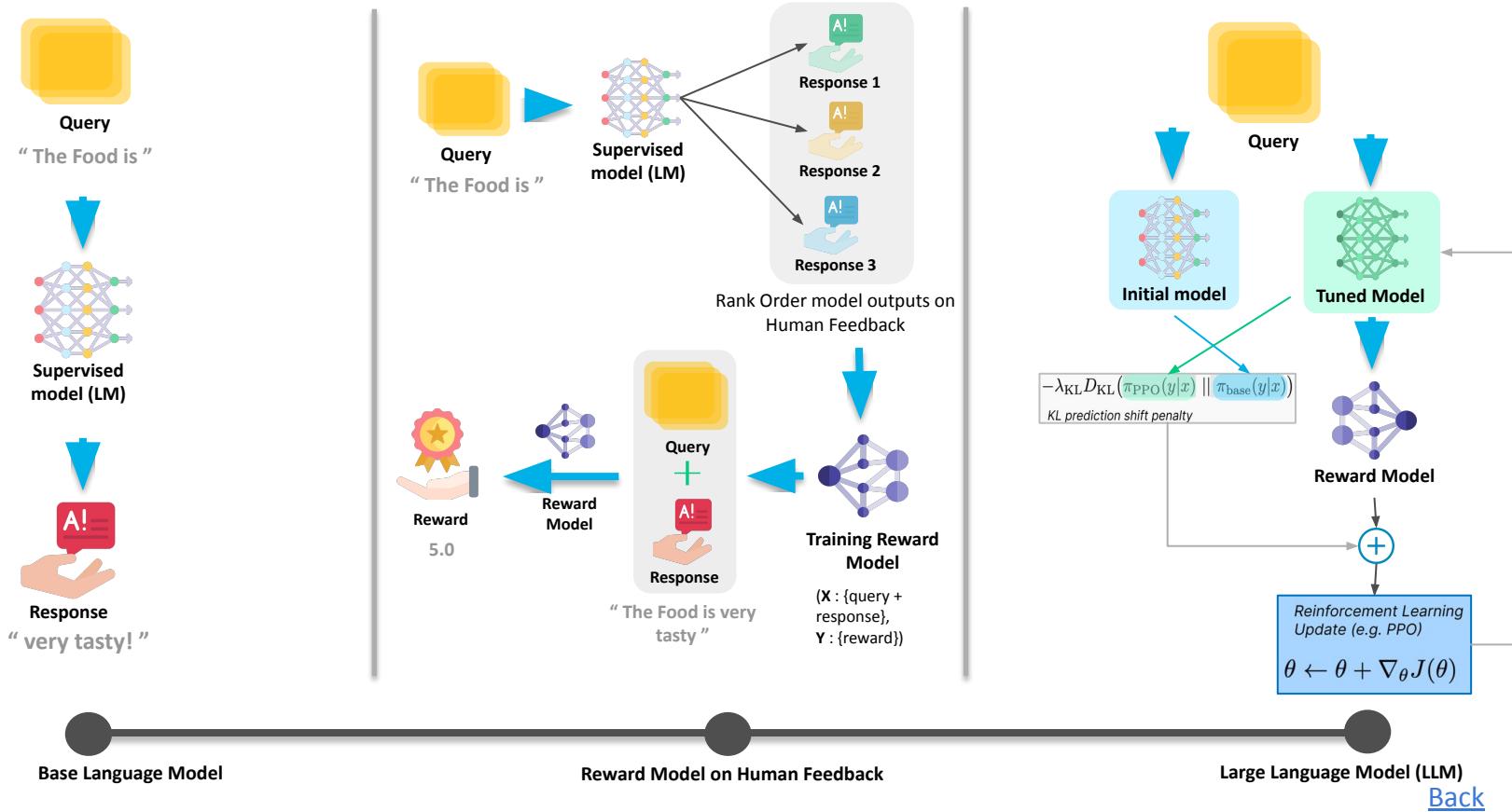
- Build a safe ChatGPT framework to automate and extract the relevant information instantly anytime
- To reduce the manual efforts to extract answers from plethora of documents

Custom LLM Building Process Overview

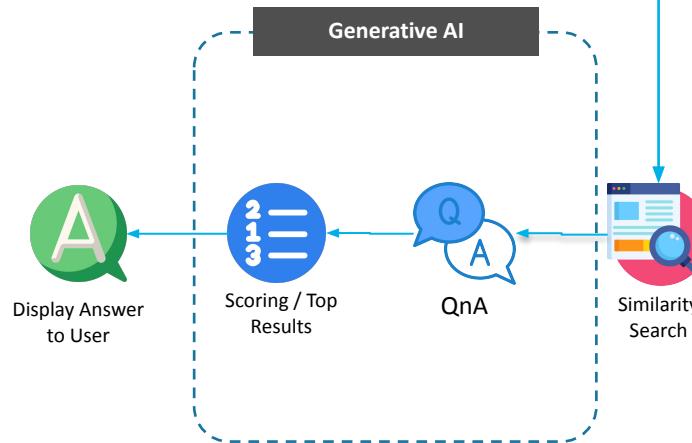
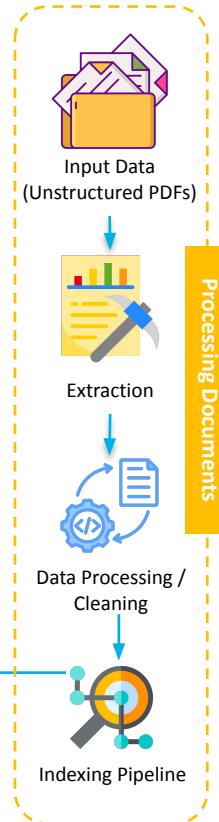
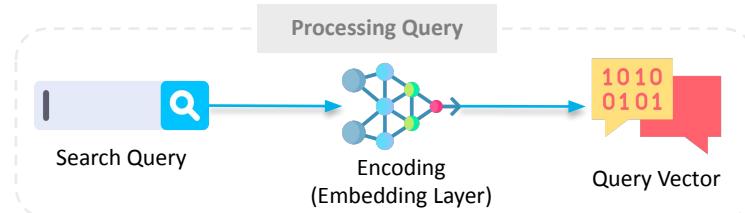
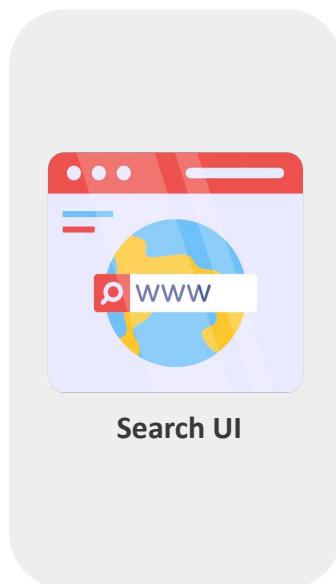


[Back](#)

Generative AI - Training



Solution Architecture



Autonomous Agents

OBJECTIVE

- Build a chatbot to have a human like conversation with the ability to answer complex business queries

FUNCTIONALITY

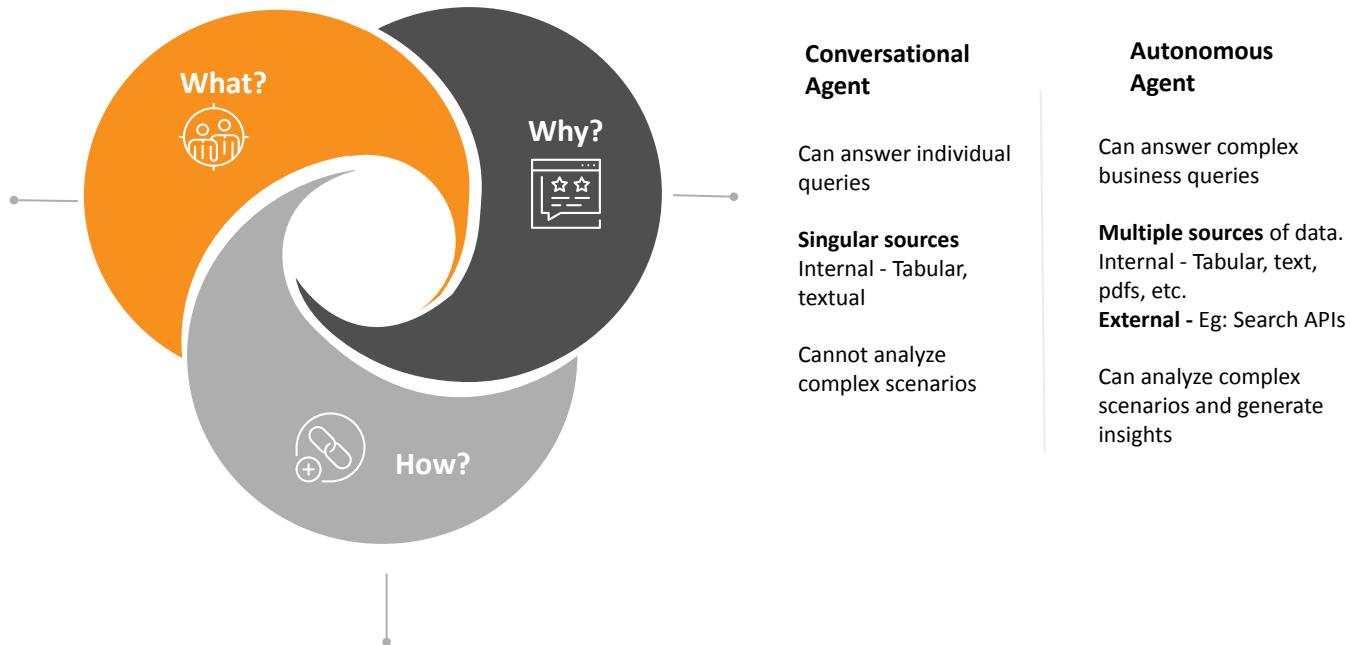
- The underlying corpus could structured or unstructured data
- Chatbot should remember the previous chat history and provide the relevant responses
- It should analyze complex scenarios and generate insights

Autonomous Agents – What, Why and How?

Autonomous Agents are Gen-AI powered systems

Given an objective they are able to:

- Break Complex questions
- Plan & Develop
- Execute & Iterate



Autonomous agents, powered by LLMs operate in conjunction with tools and a shared memory system

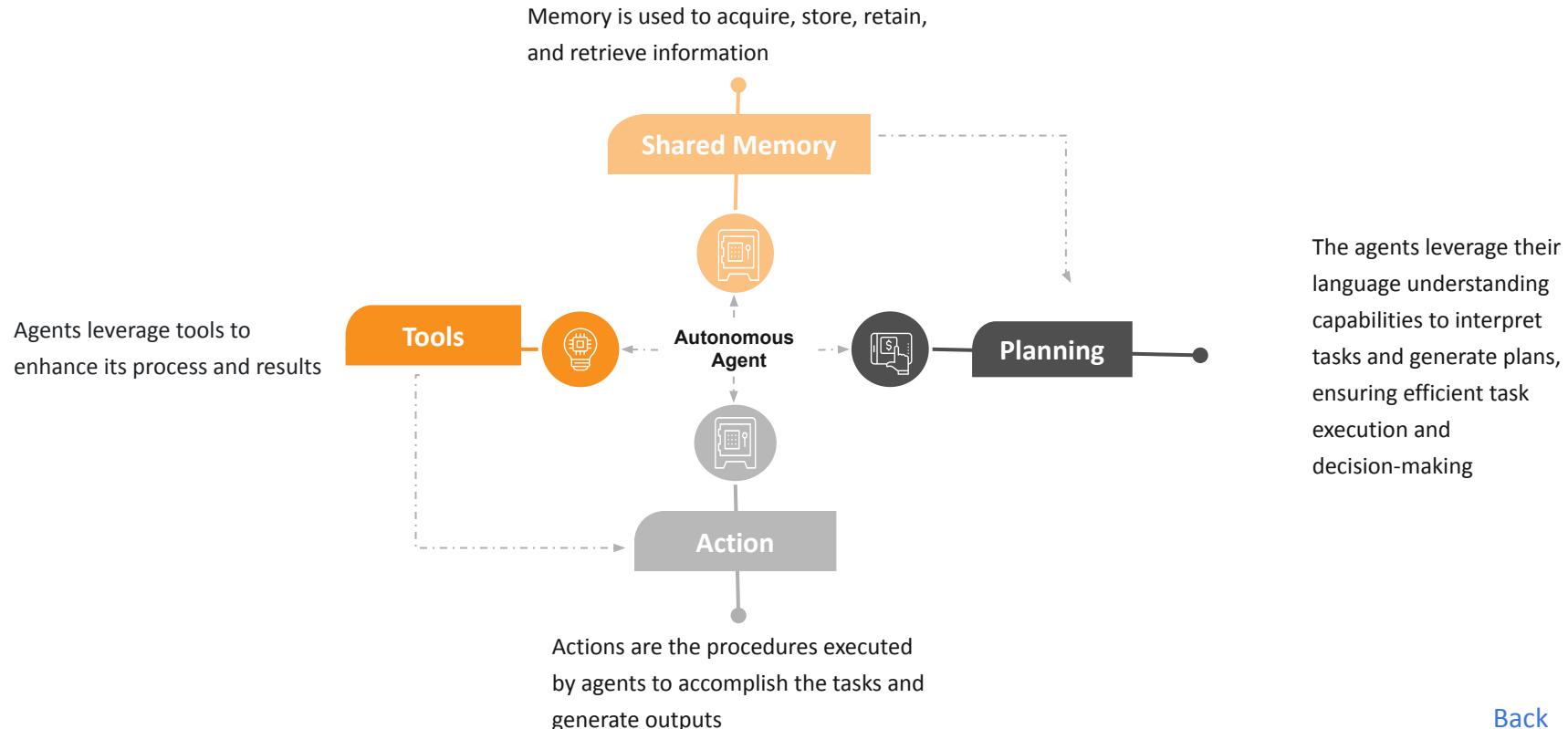
- LLMs provide these agents with advanced language understanding and generation capabilities
- The tools assist agents in task execution and communication
- A memory layer allows for information storage and retrieval

Together, it enables to process natural language input, collaborate with other agents, and enhance overall efficiency

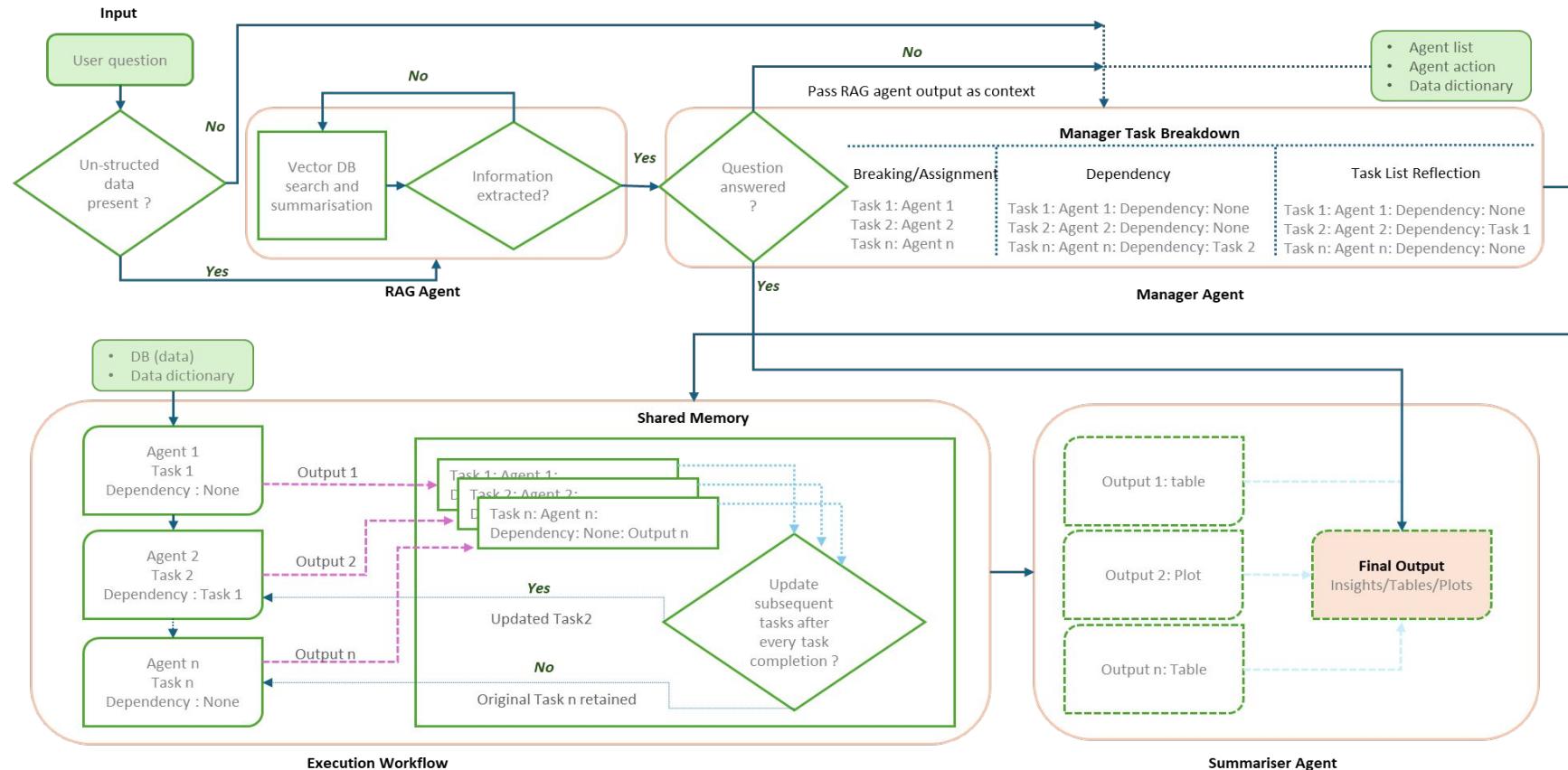
[Back](#)

Autonomous Agents framework

The LLM functions as the agent's brain, complemented by four main components: Planning, Memory, Tools and Action



Autonomous Agents Architecture



Consumer Insights

OBJECTIVE

- Develop a comprehensive and streamlined approach that can efficiently collect, analyze textual content and translate consumer feedback into actionable strategies using NLP and GenAI.

FUNCTIONALITY

- Use text analysis techniques to categorize data by feature and identify the sentiment and feature type attributed to each feature
- Group reviews at feature, feature level or as per business needs and use LLMs to generate summaries arriving at key positive/negative themes and consumer insights for businesses.

Consumer Insights - Context

Business Domains

Consumer insights are universally applicable, aiding in grasping customer behavior, preferences, and sentiment across all industries.

- CPG
- Insurance
- Healthcare
- E-commerce
- Finance
- Retail



Voice of Consumer

Consumer reviews inform business decisions, available from diverse platforms and sources via different **modes**.

Diverse **sources** are crucial for a comprehensive view of product/service quality.

[Back](#)

Key business challenges and the benefits of leveraging NLP and Gen AI solution

Business Challenges



Multi-Touch Points

- With consumers interacting across **various touchpoints** (websites, social media, in-store, customer service, etc.).
- Ensuring a complete customer journey view and capturing insights from every touchpoint is challenging for understanding consumer behavior.



Complex Feedback

Analyzing diverse consumer feedback, including text, ratings, reviews, and surveys, presents challenges due to its unstructured nature.



Interpretation and Consolidation of Feedback

- Interpreting** consumer feedback demands careful consideration, addressing conflicting or ambiguous responses.
- The challenge is converting this data into **actionable recommendations** for R&D, new products, customer satisfaction, and user experience.

Solution Benefits



Deeper Understanding of Customer Sentiment

Using **NLP** and **GenAI**, Businesses can identify evolving preferences and recurring issues, proactively staying ahead of market trends.



Enhanced Customer Satisfaction

Addressing **customer feedback** issues enhances products and services, fostering greater **satisfaction** and **loyalty**.



Enhanced Customer Service

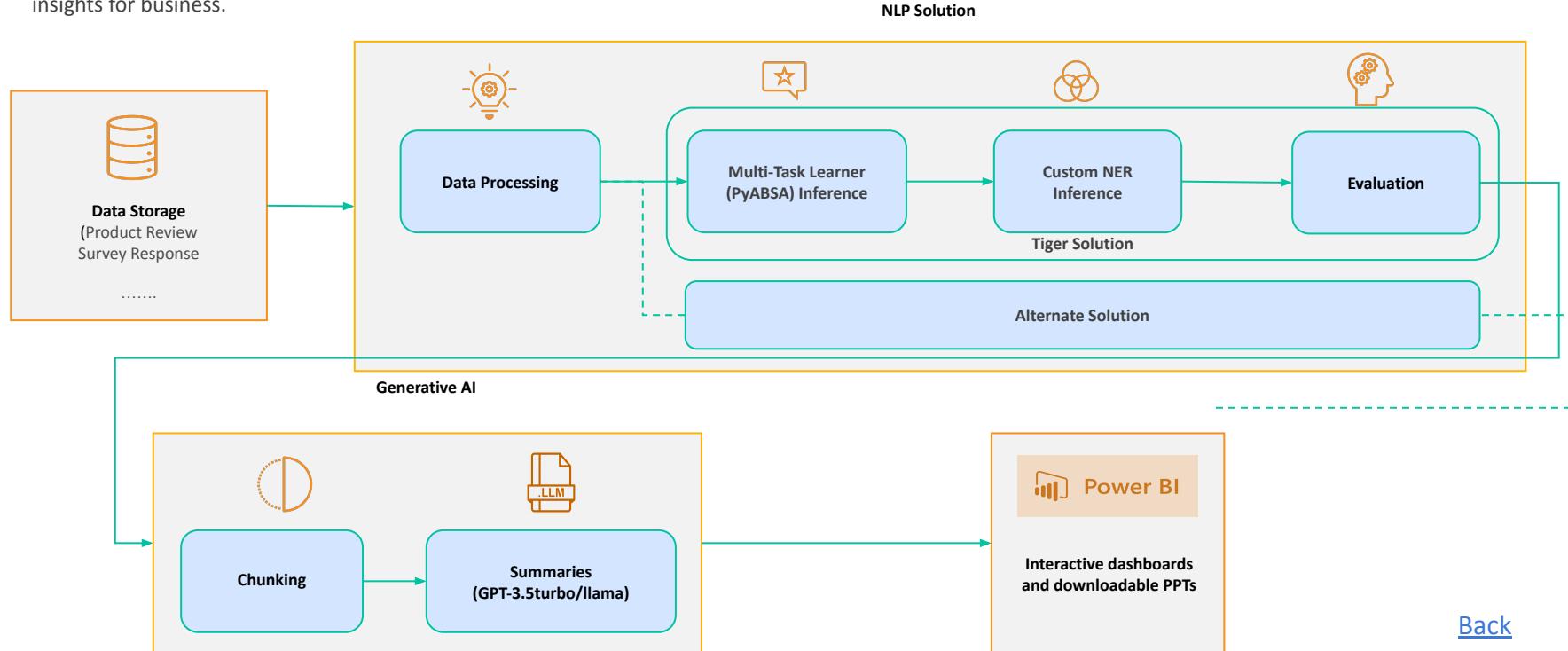
Understanding **customer sentiment** enhances tailored customer support solutions.

[Back](#)

Solution Architecture

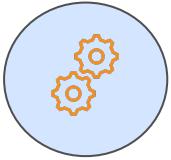
Consumer Insights solution can analyze customer feedback by associating specific sentiments with different features of a product or service using feature-based sentiment analysis (ABSA) and generate summarized insights using LLMs

- **ABSA** is one of the **text analysis technique** that categorizes data by feature and identifies the sentiment and feature type attributed to each feature.
- By grouping reviews at feature, feature level or as per business needs; using **LLMs** summaries can be generated to arrive at key positive/negative themes and insights for business.



[Back](#)

NLP Solution - ABSA

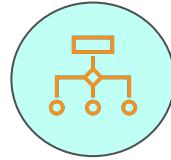


Seed Data Preparation

- Define Feature Types of interest - business decision - Tiger models has 12 aspect types Pricing, Package, Product, Format, Sensorial, Ingredient, Occasion, Product Type, Other**, Shipping, Quality, Size, Nutrient
- Utilize pre-trained model to tag features and polarity.
- Generate sample seed data for training.

Data Processing:

- Handling special characters
- Sentence segmentation
- Tokenization
- Semantic grouping of aspect terms to identify aspect types



Multi-task learning model:

The first step is to identify features and polarities for every review. BERT-based [multi-task learner](#) is used to identify the feature and polarity of the review.

Models used: BERT, Pre-trained large language models.

Custom NER model:

- Multi-task learner identifies overly granular features with multiple variants. Grouping similar features into types and aggregating sentiments at that level is suggested to find high-level drivers.
- Training a custom NER model to achieve this by classifying these feature into feature types.



Evaluation:

Multi-level evaluation of all models independently.

Overall evaluation for combined model performances - Precision, Recall, F1 scores.

Data Validation:

Validate models with SME / Domain experts using annotation tools.

Gen AI-Summarization

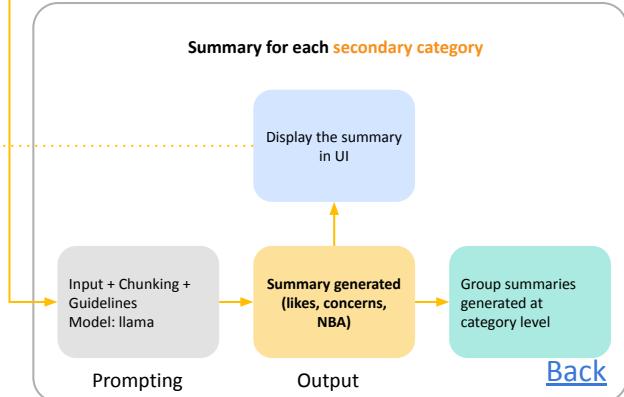
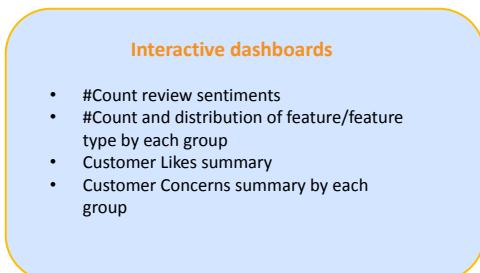
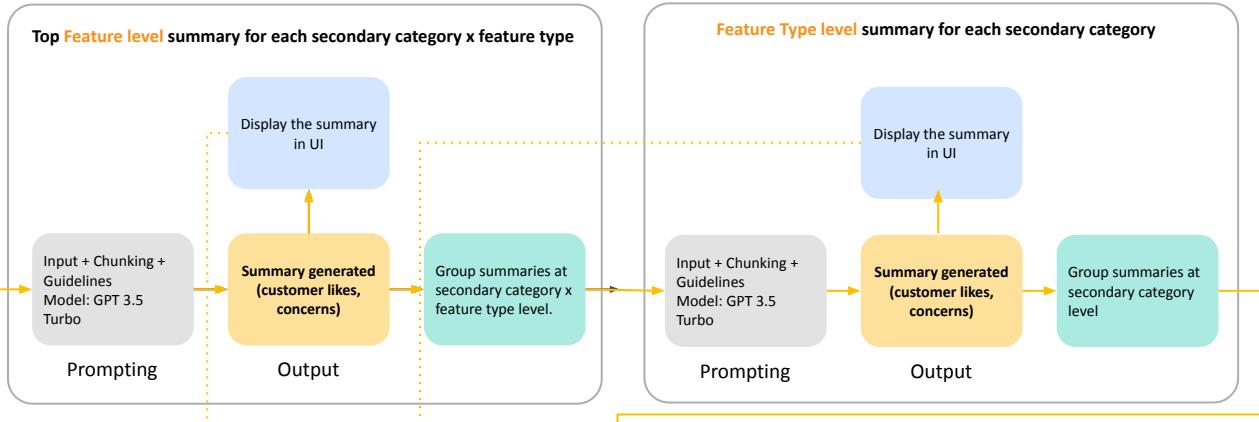
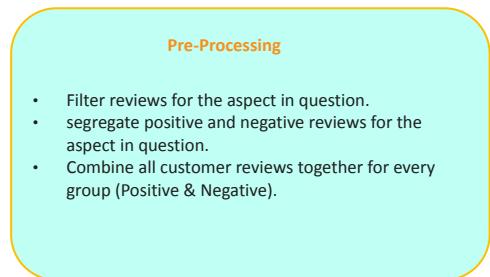
Legend

Data Preparation

Final output

LLM Prompt

Group level output





Dashboard - Demo

[Click here to go back](#)

Customer Insights Dashboard

Main Category	Subcategory	Primary Subcategory	Secondary Subcategory	Feature	Feature Type
Grocery	Beverages	All	All	All	All
Number of Reviews 14.97K		Number of Features 158		Number of Feature Types 15	
Email 5091	Website 4963	Chat 4917			Distribution of Sentiments 12253 (82%) 1708 (11%) 1010 (7%)
Average Sentiment Score 4.33 Positive					

Gen AI Generated Summary

Customer Likes	Customer Concerns	Next Best Action
Key themes/insights for beverages are: 1. Health Benefits: Customers prioritize beverages that provide health benefits such as low calorie, sugar-free, or probiotic contents. 2. Taste Preferences: Consumers prefer beverages with good taste, including flavor, aroma, and texture. 3. Convenience: Customers look for convenient packaging options, such as single-serve cups or portable bottles. 4. Environmental Concerns: Many consumers are interested in sustainability and eco-friendly packaging options. 5. Ingredient Quality: Consumers prioritize high-quality ingredients and natural ingredients over artificial ones. 6. Brand Loyalty: Customers tend to be loyal to specific brands and may choose one over another based on brand reputation and familiarity. 7. Unique Offerings: Some consumers seek out specialty beverages with unique flavors or ingredients that they cannot find elsewhere. 8. Price Points: Customers consider the affordability of beverages and may choose cheaper alternatives if they meet their other criteria. 9. Online Shopping: For some consumers, online shopping is preferred for	Key themes/insights for Beverages are: 1. Safety Concerns: Customers are worried about the potential health risks associated with improper or excessive consumption of certain beverages. 2. Unhealthy Ingredients: Consumers express concerns about the high sugar content in specific beverages, like sugar drinks, fearing an increased risk of obesity and related health issues. 3. Environmental Impact: Reviewers are increasingly concerned about the environmental impact of beverage production and disposal, including resource depletion and greenhouse gas emissions. 4. Social Issues: Customers are disturbed by the social problems linked to the consumption of specific beverages, such as binge drinking and alcoholism, affecting communities adversely. 5. Economic Impact: Consumers express concerns about the varied economic impact of the global beverage industry on local communities and economies, both positive and negative. 6. Limited Accessibility: Customers in certain areas are deeply concerned about limited access to clean water and sanitation, which hampers their ability to obtain safe and healthy beverages.	Next Best Actions: Safety concerns: - Provide clear warning labels on potentially harmful beverages. - Educate consumers on safe consumption practices. Unhealthy ingredients: - Encourage beverage companies to reduce sugar content and offer healthier alternatives. - Promote the consumption of low-sugar, nutrient-rich beverages. Environmental impact: - Encourage sustainable production and disposal practices. - Promote the use of reusable containers and encourage recycling. Social issues: - Support initiatives aimed at reducing binge drinking and alcoholism. - Encourage responsible drinking practices and promote non-alcoholic beverages.

Sentiment by Top Feature Types

Product	88.7%
Delivery	10.3%
Service	0.9%
Ingredient	88.3%
Pricing	88.0%
Portion	87.8%
Package	87.6%
Shipping	87.3%
On-time	87.0%
Quality	86.6%

Sentiment by Top Features

value	88.4%
ice	10.3%
flavor	1.3%
cost	88.4%
price	88.4%
water	78.6%
sugar	88.6%
size	88.6%
cup	88.6%
aldi	88.6%
package	88.6%

[Dashboard link](#)

Resume Matching

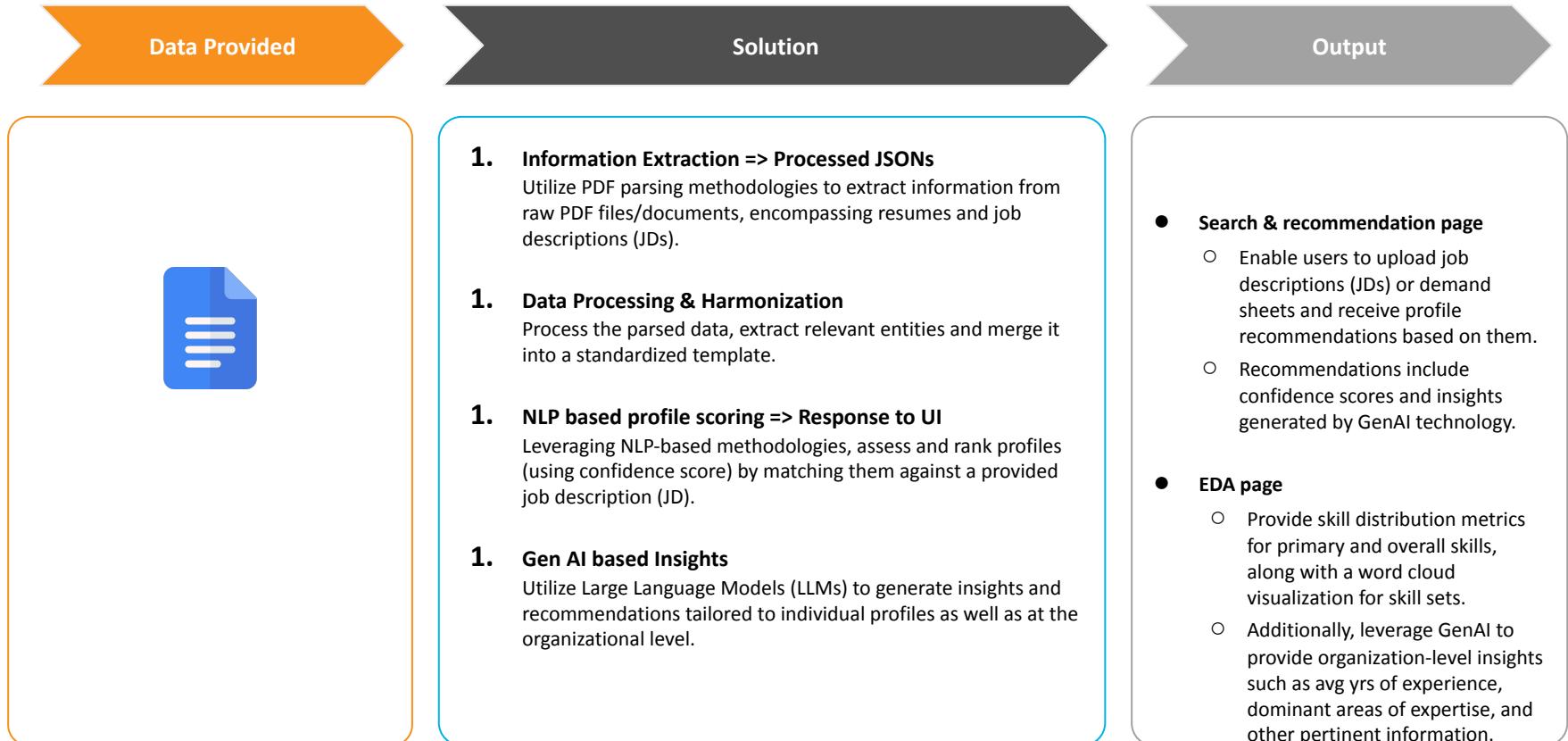
OBJECTIVE

- AI tool to provide candidate recommendations with explainable insights for a Job description

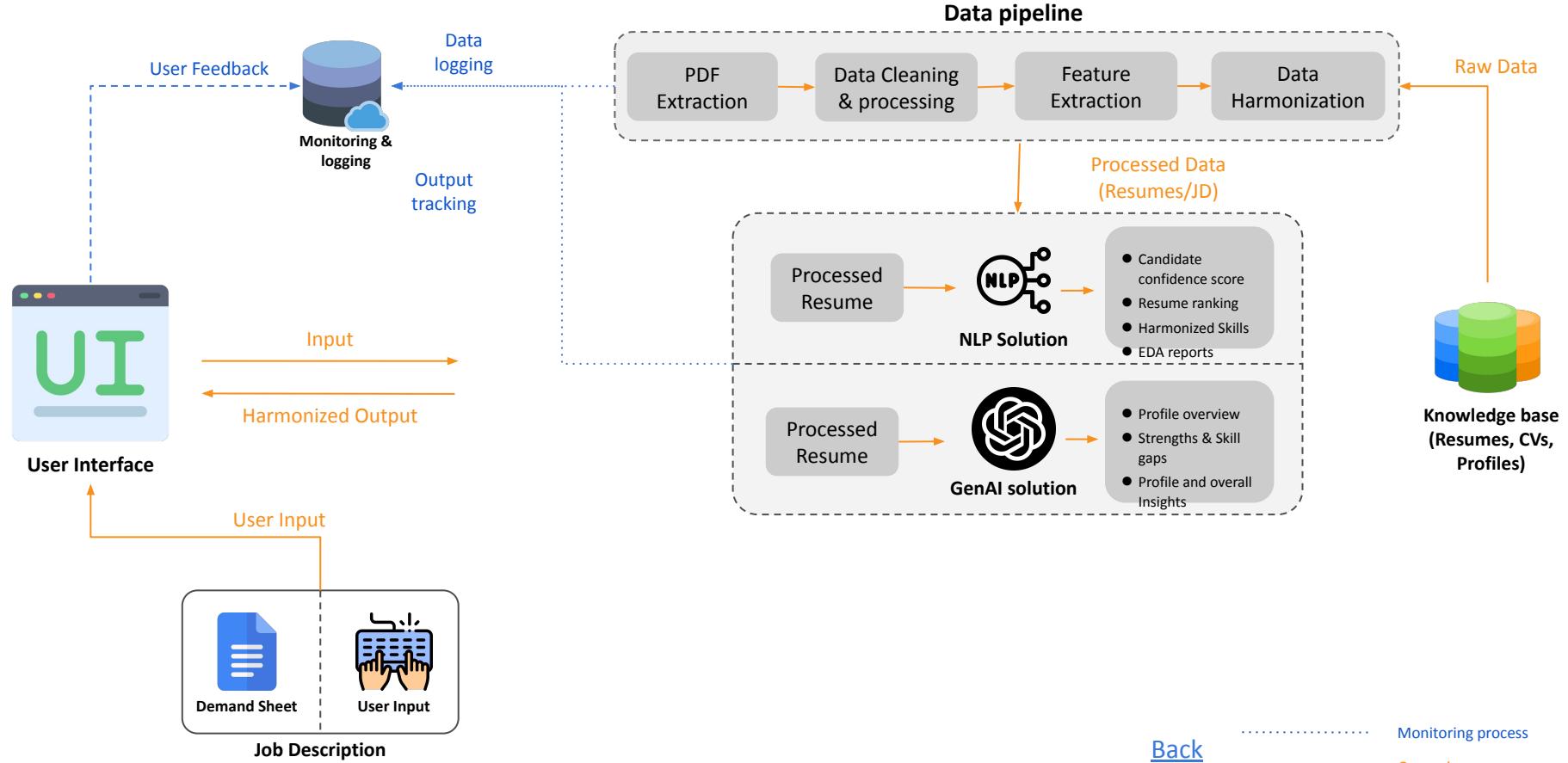
FUNCTIONALITY

- Use nlp techniques to drive relevant skills and attributes to resumes and job descriptions.
- Use similarity scores to rank the profiles.
- Use GenAI to create relevant insights by profiles to understand driving factors like skill gap, overall suggestion, strengths, wtc.

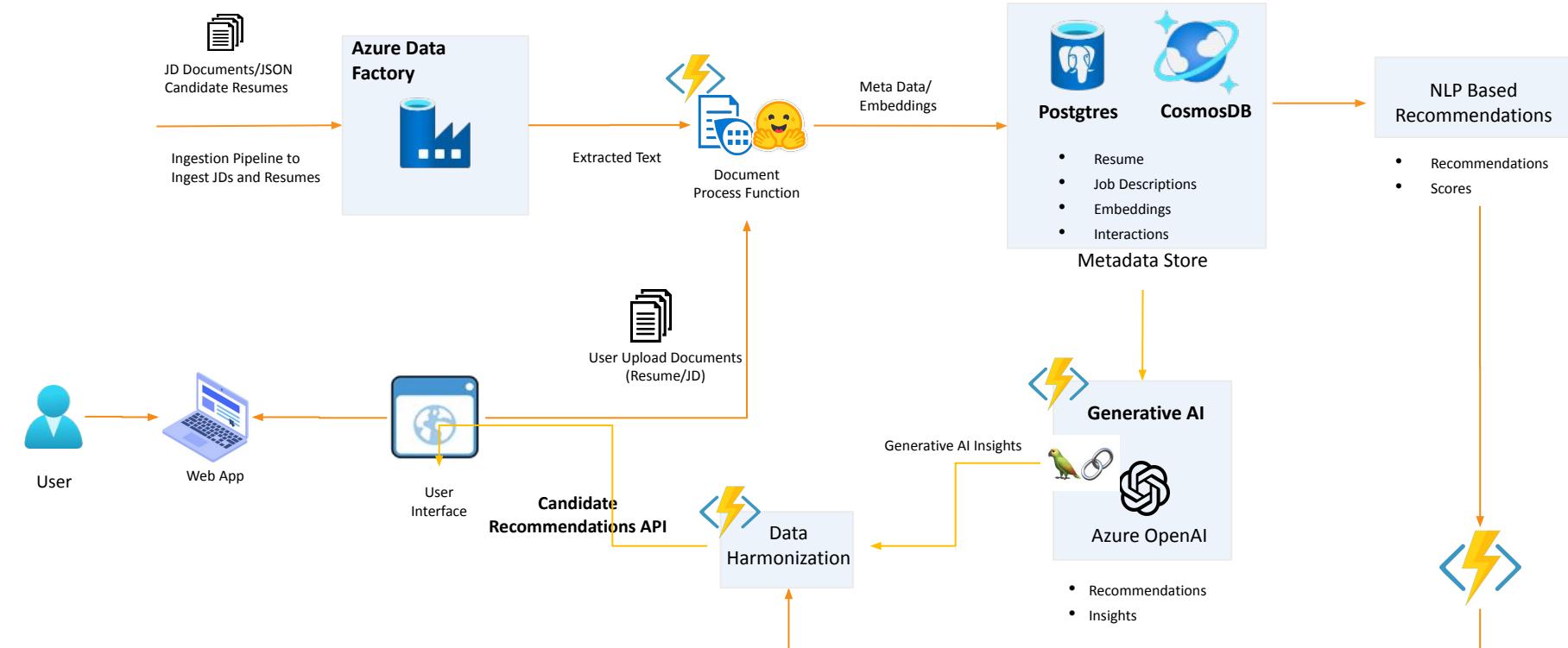
Solution Approach



Functional Architecture



Technical Architecture (Azure)



Dashboard - Demo

Tiger Analytics

Tiger Analytics

Home EDA Report

LLM Model

Azure Open AI - GPT 3.5

Stream

Data Science

No. of Profiles to show

1 4 10

Resume Matching Tool - Search

Select an option:

Upload JD

Click here to download [Sample JD Sheet](#)

Upload Excel file

Drag and drop file here
Limit 200MB per file • XLSX, XLS

Browse files

Demand Details Sheet.xlsx 31.8KB

Job Description

A strong foundation in natural language processing techniques, including but not limited to tokenisation, part-of-speech tagging, named entity recognition, sentiment analysis, and language modeling. Crafting sophisticated frameworks tailored for diverse NLP use cases, contributing to developing cutting-edge solutions. Building robust, reusable codes and libraries that empower the seamless deployment of NLP models across various applications. Experience in Text-based analytics using Deep learning or Language models. Collaborating closely with cross-functional project teams to analyze and address real-life business challenges, providing insightful solutions fuelled by NLP innovations. Keeping abreast of the latest advancements in NLP by exploring developments from leading research organizations, ensuring our methodologies remain at the forefront of industry trends.

Required Primary Skills

Required Secondary Skills

Dashboard - Demo

X

⋮

EDA Report

General Insights Skill Insights Additional Insights

Resource Mix

Category	Percentage
Data Science Core	66.7%
MLE	22.2%
NLP	11.1%

NOTE: Hover for additional insights

Resource Insights

Years of Experience

- The average years of experience among all the candidates is around 4 years.

Average work duration

- The average work duration among the candidates is around 3.5 years.

Domain Experience

- The average years of experience of the candidates in different domains are as follows:
- Data Science: 3.3 years
- Machine Learning: 3.5 years
- Natural Language Processing: 2.5 years
- Computer Vision: 2+ years
- Deep Learning: 3+ years

A3

Gen AI Implementation Approach

Implementation considerations and options

Data Privacy, domain contextualization, and implementation cost remain key implementation challenges...

Ethics & Data Privacy

- Need to expose data to external APIs
- Minimize unintended bias
- Ensure regulatory compliance

Reproducibility & Evaluation

- Ensure consistency of model response
- Define success criteria for evaluation

Domain Context

- Right enterprise data for different use case
- Leverage user feedback for reinforced learning

Implementation Challenges

- Size and infrastructure requirement
- Latency requirements and available compute

...which are being addressed through 3 deployment options

1. Enterprise API

Leverage external AI models trained on public data using API-connections to share data (e.g., GPT4)

Pros: Low implementation cost

Cons: Loss of Enterprise Data privacy, Limited domain context

2. Secure Zone

Deploy externally-trained AI models within a secure enterprise environment (e.g., Azure OpenAI)

Pros: Retain Enterprise Data privacy, Low implementation cost

Cons: Limited domain context, Limited ability to retain

3. Opensource finetune

Customize and fine tune Opensource AI models using enterprise data within secure environment

Pros: Retain Enterprise Data privacy, High domain context

Cons: High implementation cost

Identify Potential Use-cases

1. Discovering opportunity areas



Interview **business stakeholders** across functions and geographies within the enterprise



Identification of potential Gen AI use cases in line with key objectives



Interview takeaways (challenges, priorities & areas of interest, expectations)



Use Case Assessment report (Current state, gaps, analytics needs, next steps)

2. Use Cases Evaluation and supporting Commentary



Identify **opportunity areas/use cases** to drive Business value though Gen AI



Use Case Scoring report

Rated on various parameters around Complexity, Cost & Impact/value

Consolidate and summarize Gen AI Use cases identified across business functions

Compare and evaluate Gen AI use cases from a feasibility perspective

Analyze use cases on the Value Complexity framework

With **Commentary** on reason for scores, learnings and implications

3. Prioritization with Value-Complexity Framework

Data source	Description of Data	History (in Years)	Quality	Linking	Granularity	Availability	Accessibility
Salesforce	Customer, Lead, Opportunity Data	3	●	●	●	●	●
SAP HANA	Orders, Financial, Service (SG, HK) Data	5	●	●	●	●	●
ERA	ERP (AU Refill)	12	●	●	●	●	●
Autofitne	ERP (SG + HK)	8	●	●	●	●	●
Aufschick	Trade In Details	2	●	●	●	●	●
Autologs	Finance, HR, Payroll	1	●	●	●	●	●
F&I (Excel for SG + HK)	Customer spend data by Product category	4	●	●	●	●	●
Marketing Spend - AU	Customer spend data by Product category	4	●	●	●	●	●
Marketing Activity - AU	Campaign Impact on goals: Visits, Impressions, leads etc.	4	●	●	●	●	●
Marketing Spend - SG + HK	Marketing spend data by Product category	1	●	●	●	●	●
Marketing Activity - SG + HK	Campaign Impact on goals: Visits, Impressions, leads etc.	1	●	●	●	●	●
Google Analytics	CX, User website behaviour, UTM etc.	5	●	●	●	●	●

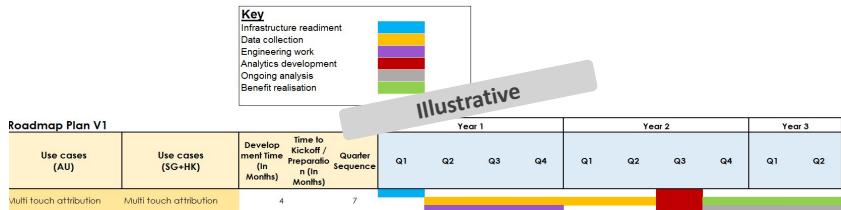


For genAI use-cases, along with **Evaluating enterprise value**, there would be considerations on **Model strategy** (Eg: Enterprise vs OS / single vs multiple models) and **Model Evaluation** (Technical considerations for the Use-case)

4. Defining Next Steps



Next steps for the identified Gen AI use cases for technical solutioning, feasibility analysis and value generation



Technical Landscape across Use-cases

1. Define Research Objectives



Define research objectives, e.g.



Applicable LLM & image models for the identified use cases



Resource requirement & complexity in using enterprise vs custom built LLMs



Cost implication of building custom LLMs vs using a SaaS platform/service

2. Collate Data



Identification of authentic sources for LLM and Image model details



Information about availability of domain agnostic + domain specific models for different use cases



Publicly available data on model performance & training requirements for customization

3. Compare and Evaluate



Define parameters for evaluating LLM and Image models and derive a scoring



Comparing models based on the scoring and associated pros & cons



Additional layer of intelligence based on the experimentation and research done at Tiger Analytics to compliment the external research data

4. Present findings



Knowledge exchange & brainstorming sessions with the PepsiCo data science team for their perspective on the research findings



Consolidate the research results



Enlist actionable recommendations for Gen AI roadmap creation

Dimensions for Use-case & Model Evaluation

Use-case & Value Assessment



Adoption

- Adoption Difficulty
- Adoption Readiness
- Specialized Skill requirement
- Business Complexity

Value

- Dollar impact
- Time to Value
- Ease of Measurement
- Risk

Use-case Req.

- Consumption
- Latency
- Accuracy
- Fluency

Data Req.

- Structured vs Unstructured
- Latency
- Accuracy
- Fluency

Technical Assessment

(Use-case, Data, Model, Technology)



Data

- Data availability
- Data preparation
- Data model/ metadata
- Integrations
- Data governance and security

Functional

- Embedding Store
- Vector DB
- Data Labelling
- Feedback loop
- Single vs Multi-objective models

Modeling

- LLM APIs/Hybrid/Custom
- Fine tune/Prompt tune/prompt engineer
- Latency
- Evaluations
- Token Limitations

Tooling & MLOps

- Model Management
- Model Monitoring
- Data / Model drift
- Labelling workflow
- Logging
- Deployment workflow

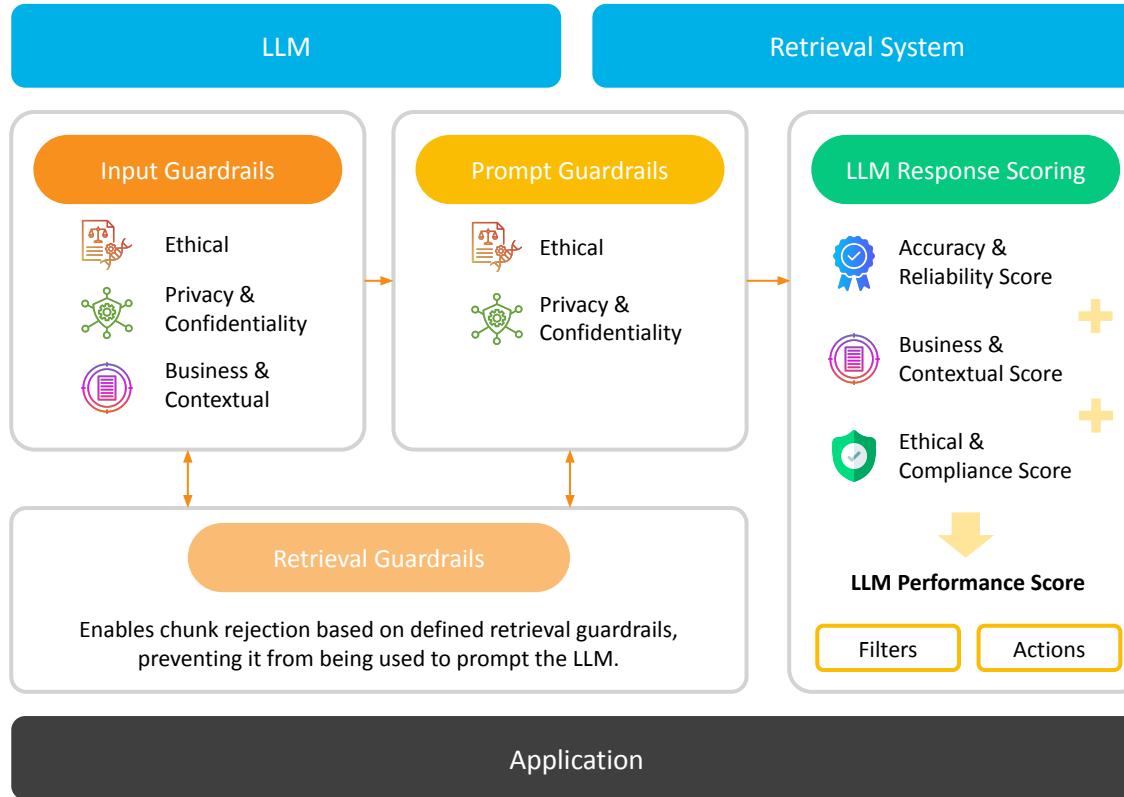
Infra

- Cloud / on-prem platform
- Orchestration tools
- Performance Tools
- Labelling Tools

Cost

- Training Cost
- Inference/Deployment Cost
- Scalability

LLM Governance and Evaluation Framework



- ❖ Input guardrails are applied to the input from the user – to enable input rejection, additional processing, or alteration
- ❖ Prompt guardrails are applied at a prompt level to influence how the LLM is prompted
- ❖ Retrieval guardrails are applied in RAG scenario
- ❖ Output evaluation consists of scoring LLM response against 3 key categories with relevant weights assigned to category level parameters
- ❖ Filters and Actions are applied to the output generated by LLMs to mask sensitive and confidential information or take required actions to refine the output before publishing
- ❖ Constant evaluation and filtering ensures preventing inaccurate, unreliable, or inconsistent, and unethical outputs being returned to the user

A framework for assessment of Gen AI solutions

Data Assessment



Data Type

- Sensitive Data or Non-sensitive data.
- Volumes, Format and location of data that can be leveraged.
- Multi-modality of data (Text, Video, Speech etc.)

Training Data

- Availability of Training data
- Availability of Business users to evaluate quality of results



Infrastructure Assessment

- Cloud Providers (AWS, Azure, GCP, Snowflake, Databricks etc.)
- Models and Approach (Open Source vs Enterprise)
- Vector Databases, Annotation Tools, Parsers, LLMOps etc.

Performance Assessment



Query Complexity

- Simple Questions: Direct queries - Ex: Search and Retrieve, Simple SQL queries
- Complex Questions: Analytical queries or nested queries. - Ex: Complex SQL queries, Reasoning based queries

Response Evaluation

- Summarization: Coherence, Fluency, Consistency, Relevance
- Code Generation: Syntax Correctness, Joins, Aggregations, Filters

Latency:

- Model Latency vs Solution Latency
- Open-Source Model vs Enterprise Model
- Vector Database Indexing and Querying Latency

Experience Assessment



- Feedback on Value Add | User acceptance
- **Impact Metrics:** Average Handling/Waiting Time, Error rates, CSAT scores etc.
- **Productivity Gains** - Time Savings in Coding, Manual Search
- **Human Evaluations:** Good vs Bad, Rating 1-5 (low - high), Comments etc.
- **Conversational Evaluation :** Evaluation of performance across multiple threads.

Cost



- Infrastructure & Compute
- Enterprise LLM v. Open Source LLMs
- Enterprise LLMs - Token Subscription costs
- Open Source LLMs - GPUs for customization
- Cloud Services: App Services, Operational Libraries (Parsers, Search | Vector Database Costs, Annotation Tools etc.)

A4

Addressing LLM concerns
through our learnings

Addressing some of the concerns with LLMs: Our learnings (1/2)

1



Data Privacy

- ✓ Custom Model training and fine tuning per objective within client environment preventing data leakage
- ✓ Azure Open AI - Data will not be used for training. 30-day retention before cache gets deleted
- ✓ Prompt engineer such that limited data can be passed to Enterprise models

2



LLM Results

- ✓ Evaluating results have two level of complexity:
 - Reproducibility
 - Hallucinations
- ✓ Hallucinations can be controlled using different prompting techniques.
- ✓ Self-consistency/Reproducibility can be controlled in some open-source models but not as much in Enterprise API (like GPT4)
- ✓ Specific evaluation metrics have to be designed to evaluate and control these results.

3



Bias

- ✓ Fine-tuning models using Rewards to ensure contextual bias is mitigated
- ✓ Identification of certain bias indicators such as gender, race, ethnicity and correcting them using prompts
- ✓ Ensuring human-in-the-loop for evaluating bias in the models
- ✓ Prompt engineering to correct bias with chain-of-thought prompting, one-shot examples etc.

Addressing some of the concerns with LLMs: Our learnings (2/2)

4



Hallucinations

LLMs could generate syntactically and semantically coherent text but far from reality and ground truth/facts. This is referred to as hallucination. Some ways to control hallucinations while creating a system with LLMs include:

- ✓ Retrieval Augmented Generation
- ✓ Regularization
- ✓ Chain of Thought prompting

5



Reproducibility

Levers available to address the reproducibility of the results, across multiple runs/users, include:

- ✓ Temperature - A variable that can be controlled and fixed to generate same response
- ✓ Ensemble Approach - For Enterprise LLMs, temperature variable doesn't fully control. We can use wisdom of crowds, majority votes, ensemble approaches to make responses more consistent
- ✓ Task Decomposition - Breaking down complex questions into further simpler tasks that can be controlled by Agents
- ✓ Human-in-the-loop - To validate response and pass it back to the reward model for further training

6

Other factors



Infrastructure



Deployment



Latency