

Gen AI Academy

APAC Edition

Build in APAC. Build for the world.



Participant Details

- a. Participant name: Vaibhav .
- b. Problem Statement: Build and Deploy AI Agents using Gemini, ADK and Cloud Run

Idea Overview – AI Research Assistant

Title: AI-Powered Research Assistant using ADK & Gemini

Built a lightweight AI agent that answers user questions in real-time

Uses **Google ADK framework** for agent structure and orchestration

Integrates **Gemini model** for intelligent response generation

Enhances answers using a **custom Wikipedia search tool** for factual accuracy

Deployed as a **scalable web service on Cloud Run**

Accessible via:

- HTTP API
- Chat-style web interface

Meeting the Build Criteria

1. Single AI Agent (ADK-based)

- Designed one focused agent using ADK
- Handles end-to-end flow: input → reasoning → tool use → response

2. Uses Gemini for Inference

- Integrated Gemini as the core LLM
- Generates clear, structured, and human-like answers

3. Performs One Defined Task

- Task: **Question Answering**
- Takes user query and returns a precise answer

4. Tool Usage (Reasoning Capability)

- Agent decides when to use Wikipedia tool
- Demonstrates **agentic behavior (tool + LLM integration)**

5. HTTP Endpoint + Deployment

- Exposed via REST API
- Fully deployed on **Cloud Run**
- Publicly accessible and scalable

Opportunities & Differentiation

01

How is it Different?

Combines:

- LLM reasoning
- External knowledge retrieval (Wikipedia tool)

Fetches real-time contextual data Build using ADK

How Does it Solve the Problem?

1. Agent dynamically retrieves verified information
2. Uses Gemini to **refine + structure the answer**
3. Ensures **better accuracy and relevance**

02

03

USP (Unique Selling Proposition)

“Hybrid Intelligence: Reasoning + Retrieval + Generation”

List of features offered by the solution

Core Features

- Accepts user queries via **API or chat interface**
- Generates **clear, structured answers** using Gemini
- Integrates **Wikipedia search tool** for factual grounding
- Real-time response generation

Intelligent Capabilities

- **Agent-based reasoning** (decides tool usage)
- Context-aware answers
- Reduces hallucination using external data

Deployment & Accessibility

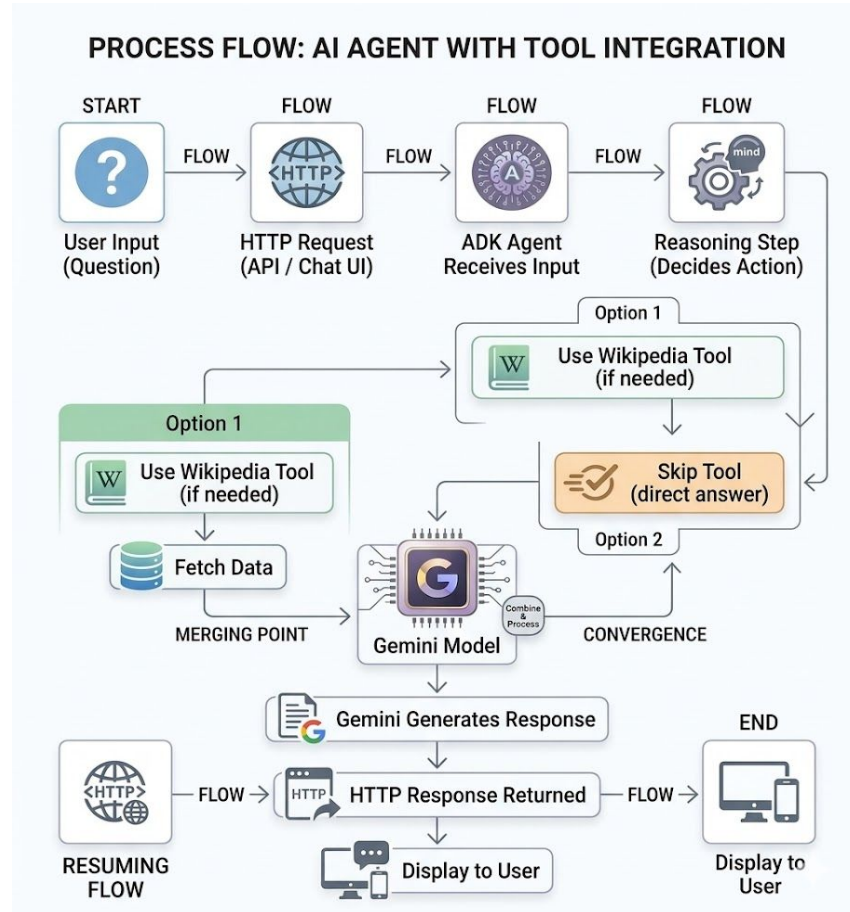
- Hosted on **Cloud Run**
- Public HTTP endpoint
- Scalable and serverless

Extensibility

- Easily add new tools (Google Search, PDFs, databases)
- Can evolve into:
 - Research assistant
 - Customer support bot
 - Knowledge engine

Process flow diagram or Use-case diagram

The agent first reasons whether external data is required, then uses a tool if needed, and finally generates a structured response using Gemini



Technologies Used

AI & Agent Framework

- **Google ADK** → Agent orchestration & tool handling
- **Gemini Model** → Natural language understanding & generation

Backend & API

- Python (FastAPI / Flask) → API handling
- REST API → Communication layer

Tool Integration

- Wikipedia API (custom tool) → External knowledge retrieval

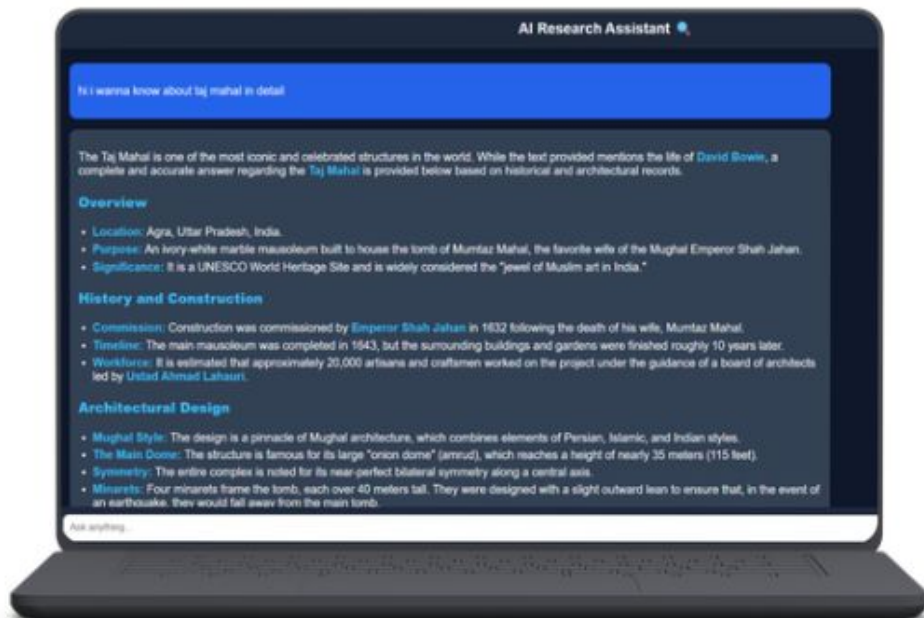
Deployment & Cloud

- Google Cloud Run → Serverless deployment
- Docker → Containerization

Frontend (if included)

- HTML, CSS, JavaScript → Chat interface

Snapshots of the prototype



Google Cloud 

Gen AI Academy

Build in APAC. Build for the world.

Thank you

