MagicAI Portal - Gemini 2.5 Flash - Based Conversational Chatbot

**1. Logical Solution**

The chatbot platform is built using a React JS frontend and a Python Flask backend. It integrates with Google's Gemini 2.5 Large Language Model to handle natural language conversations and is containerized using Docker. The infrastructure is hosted on Google Cloud Platform (GCP) and leverages Kubernetes (GKE) for scalable deployment. Jenkins is used for Continuous Integration/Deployment (CI/CD) automation.  
  
\*\*Logical Flow:\*\*  
1. The user interacts with the React-based web UI.  
2. User input is sent to the Flask backend via REST.  
3. The Flask app processes the input and calls the Gemini 2.5 Flash API for a response.  
4. The response is returned to the frontend for display.  
5. Optional logging, analytics, and data storage occur for further insights.  
  
This architecture ensures a scalable, maintainable, and efficient pipeline from user input to intelligent response.

**2. Infrastructure Requirements**

| **Component** | **Purpose** | **Suggested GCP Service** |
| --- | --- | --- |
| **Frontend Hosting** | Serve React SPA | Cloud Storage + CDN / GKE + NGINX |
| **Backend App** | Host Flask + Chat Logic | Google Kubernetes Engine (GKE) |
| **Container Registry** | Store Docker images | Artifact Registry |
| **CI/CD Pipeline** | Build & deploy on commit | Jenkins on Compute Engine |
| **Database (optional)** | User/chat session persistence | Cloud Spanner/ BigQuery |
| **LLM Integration** | Prompt/response from LLM | Gemini 2.5 / Vertex AI |
| **Logging & Monitoring** | System logs and debugging | Stackdriver / GCP Operations |
| **Secrets Management** | Store API keys securely | Secret Manager |
| **Domain & TLS** | SSL and domain routing | Cloud DNS + HTTPS Load Balancer |
| **IAM & Permissions** | Access control for services | GCP IAM |

Additional Optional Services:  
- Vertex AI: For custom model training or RAG pipelines

**3. Design Diagram (Logical + Infrastructure)**

The architecture is modular, leveraging GCP's managed services for security, reliability, and scalability:  
 