Home Assignment - Agent Engineer

Context

1. **Overview**
   a. As an Agent Engineer in the Agent Engineering group, you will be part of a team of agent engineers who build the best enterprise-grade AI agents.
   b. The company is developing an **AI-powered pharmacist assistant** for a retail pharmacy chain. The agent will serve customers through chat, using data from the pharmacy's internal systems.
   c. Customers will ask about medications, stock availability, prescription requirements, and safe usage.
   d. The agent must follow strict policies to provide factual information only, while avoiding any form of medical advice or diagnosis.
2. **Assignment Objectives** - Build real-time conversational AI pharmacy agent, on OpenAI API. The agent expected to handle workflows through tools: prescription management, inventory control, and customer service.

Requirements

1. **AI Agent** - Implement real-time text streaming AI agent based on GPT-5 (key provided separately). The agent is state-less, and can talk both in Hebrew and English.
2. **Tool** - Design and implement agent's tools
3. **Database** - Create a synthetic database of 10 users and 5 medicals.
4. **UI** - Interact with the agent.
5. **Multi-Step Flow** - Design and implement at least three distinct multi-step flows the agent can execute
6. **Evaluation Plan** - Provide an evaluation plan for your agent.
7. **Docker** - Wrap your project in a Docker file.

Guidelines

1. **Backend Language -** Python, JavaScript, TypeScript or Go (no limitation for the frontend).
2. **AI assistant -** You may collaborate with any AI assistant (Claude Code, Codex, etc.) during development. However, you are expected to be prepared to explain the code and answer questions about implementation decisions.

Agent Requirements

These are the minimum requirements the agent must meet. You may propose and include additional requirements if you believe they are important for production readiness.

1. Provide factual information about medications.
2. Explain dosage and usage instructions.
3. Confirm prescription requirements.
4. Check availability in stock.
5. Identify active ingredients.
6. No medical advice, no encouragement to purchase, no diagnosis.
7. Redirect to a healthcare professional or general resources for advice requests.

8. Streaming

Flows to Design

The agent should be able to execute at least three distinct multi-step flows. A multi-step flow is the expected sequence of actions an agent follows in a real use case, covering all the steps a professional would take when assisting a customer from the start of their request to its resolution.
Define each flow, describe the expected sequence, and outline how the agent will use the functions and respond.

Function (Tool) Design Requirements

At minimum, design 3 different tools, for example:

1. **get_medication_by_name**: Inputs, outputs, example responses, error handling, fallback logic.
2. Additional tools may be added if justified.

For each function, document:

1. Name and purpose
2. Inputs (parameters, types)
3. Output schema (fields, types)
4. Error handling
5. Fallback behavior

Deliverables

1. **README.md** - Explanation of the project and its architecture, explanation of how to run the Docker.
2. **Multi-Step Flow** - Three multi-step workflow demonstrations.
3. **Evidence** - 2-3 screenshots of conversations.
4. **Evaluation Plan** - Evaluate the Agent flows.

Evaluation Criteria

1. Tool/API design clarity.
2. Prompt quality and integration of API usage.
3. Multi-step interaction handling.
4. Policy adherence.
5. Testing rigor (coverage in Hebrew, multiple variations per flow).
6. Quality and completeness of flow designs.

Good luck!