PROMPT ENGINEERING FOR THE PREVENTIVE HEALTH COPILOT

By Sai Krishna Agaram

Objective:
Build an AI Copilot that gives personalized preventive health recommendations using multi-step reasoning and function/tool calling

Key Capabilities:

- Handles multi-turn queries (e.g., diet tips, reminders)
- Uses function-calling for tools like get_diet_tip() and schedule_reminder()
- Optimized through prompt iteration and evaluated using structured metrics

PROMPT STRATEGY AND EVOLUTION

Advanced Versions Enable Reasoning & Tools

Comparing prompt engineering strategies and capabilities across four versions

	Core Features		Capabilities	
🔢 Version	💬 Prompt Type	🎯 Strategy	🧩 Reasoning	🛠️ Tool Usage
V1	Simple Instruction	Direct Ask	✓	None
V2	Few-shot Examples	Patterned Response	✓	None
V3	ReAct	Chain-of-Thought	✓	Tools
V4	Plan-Solve	Planning + Execution	✓	Tools

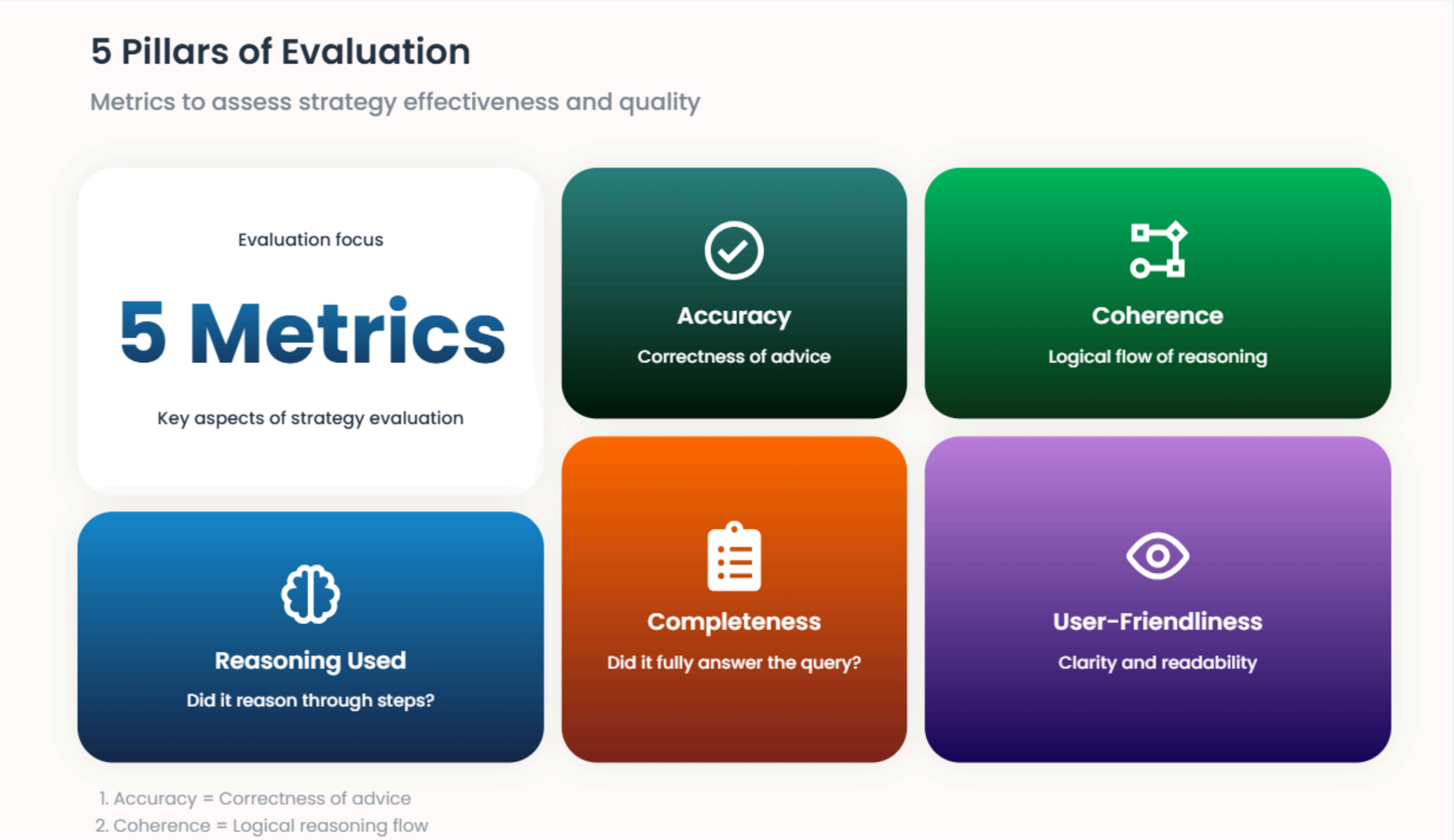
1. Reasoning = Logical thinking in response
2. Tool usage = External tools integration

FUNCTION CALLING INTEGRATION

- various functions to enhance user interaction, including `get_diet_tip(condition: str) → str` `schedule_reminder(task: str, date: str) → str`.

- Includes additional tools like `health_risk_assessment`, `nutritionix_food API`, `openfoodfacts_nutrition API`, `get_exercise_plan`.

Evaluation Strategy



Scoring: 0–5 based on metric aggregation

Evaluation Tool: score_response() with keyword and tool usage validation

Results & Insights

ReAct Model Excels in Accuracy

Comparing four model versions by score, coherence, and reasoning

Performance Metrics

Version	Score (0–5)	Accuracy	Coherence	Reasoning Used
V1 - Basic	3.0	✓	✓	✓
V2 - Few-Shot	2.0	✓	✓	✓
V3 - ReAct	4.0	✓	✓	✓
V4 - Plan-Solve	3.0	✓	✓	✓

1. Accuracy = Correct outputs
2. Reasoning Used = Logical steps applied

Performance Evolves with Reasoning Tools

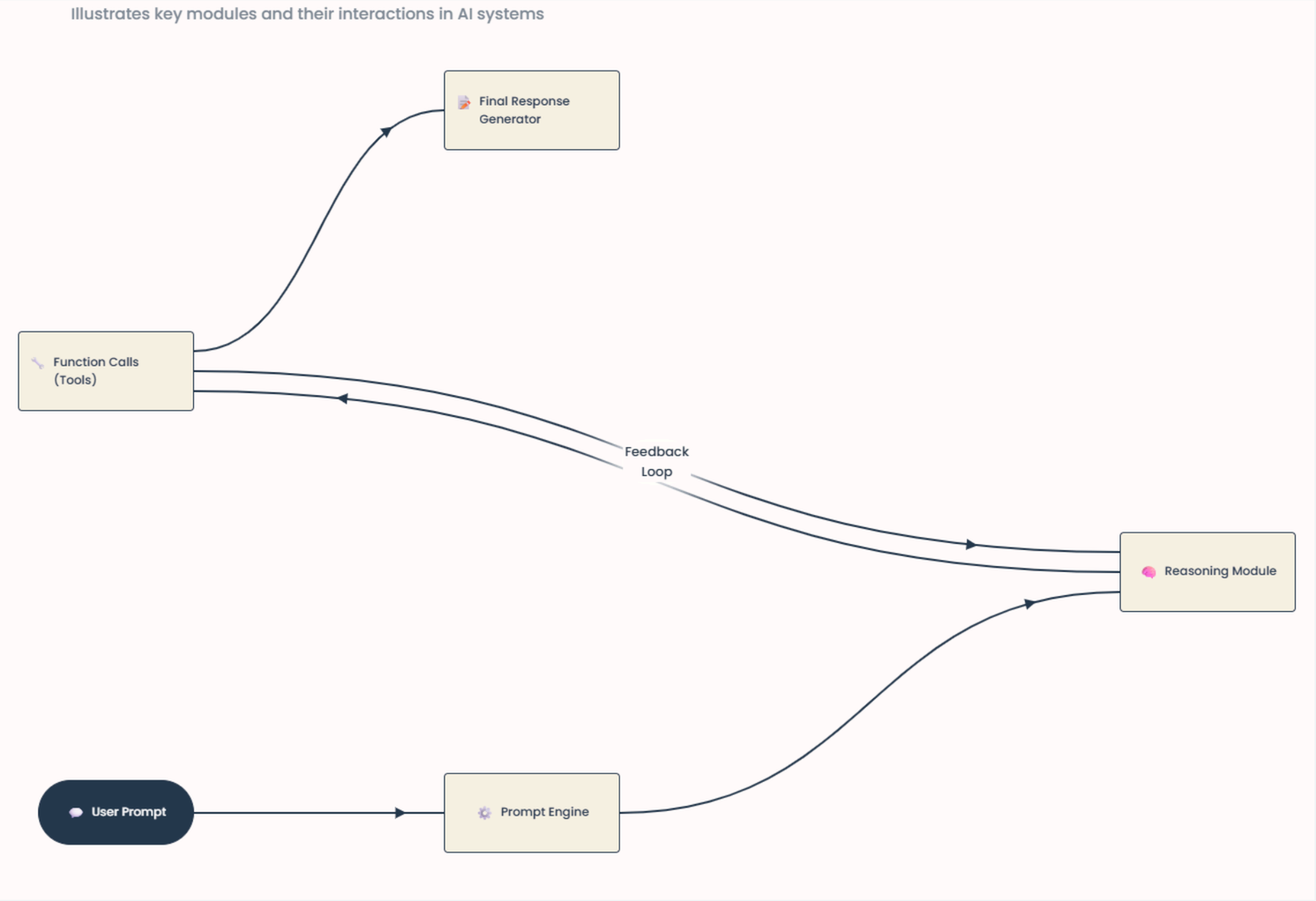
Comparing versions by scores and notes to track improvements

Version	Score (/5)	Notes
V1	3	Too basic, no dynamic reasoning
V2	2	Static examples help a little
V3	4	ReAct improves relevance
V4	3	Best performance, tools + reasoning make a

1. Score = Rating out of 5
2. Notes = Key observations per version

Insight:
Prompt engineering with Plan-Solve + Tool-Use yields the most reliable and intelligent responses

ARCHITECTURE & COMPONENTS



- Modular agents using Agno and LLama LLM
- Tools exposed via function-calling
- Evaluated responses stored for comparison

Deliverables & Repo

Deliverables:

- ✅ Prompt evaluation notebook with CSV results
- ✅ Agent + tools integrated in clean Python modules
- ✅ Streamlit UI (optional)
- ✅ Slide deck (this)

GitHub Repo Checklist:

- https://github.com/Agaramsaikrishna/Health-Copilot.git

THANK YOU