Runtime Architects — Team Meeting Minutes

Date: 7 July 2025

Time: 6:40 PM – 7:25 PM

Location: Online Call

Attendees:

• Pablo Periañez Cabrero

• Firose Shafin

• Nithyakamal Ilamurugu (Kamal)

---

1. Backend Progress Update (Pablo)

- Pablo presented a modularized streaming class consolidating recent progress.
- Implemented an SQLite-based database using app.db to simplify initial integration.
- Currently includes models for user emails and credentials using passkey authentication (Face ID / Touch ID).
- Explained passkey mechanism and authentication flow, including challenge generation and verification.
- Registration works, but cookie/session management in frontend is still pending.
- Need to restrict routes based on login state.

---

2. Virtual Environments and Plotting (Kamal)

- Described a system with two virtual environments:
    - A base environment containing core packages.
    - A temporary working copy that is deleted after use.
- Justification: avoid persistent changes and maintain clean base.
- Showed progress on carbon data plots:
    - Monthly chart prototype completed.
    - Will adapt visuals to daily format for better insight.
- Challenges remain with performance/latency even in simplified setup.

---

3. Multi-Agent Coordination Issues

- Kamal highlighted issues with agents hallucinating tool access due to unclear prompts.
- Pablo and Kamal discussed prompt clarity and evaluation strategy.

- Team agreed prompts need to be streamlined before layering complexity.
- Pablo raised concerns about poor backend structure and suggested a refactor.

---

4. Evaluation and Prompt Flow Expectations (Firose)

- Firose shared confusion about the vague expectations for Wednesday's goal.
- Task: build an agent with 4 custom capabilities (reading, analyzing, plotting, summarizing).
- Aim: automatic evaluation and prompt adjustment based on agent output quality.
- Discussed risks of automatic prompt manipulation in production.
- Explored the use of LLM-as-a-judge and Prompt Flow; no conclusive implementation path yet.
- Agreed need to track and assess agent performance systematically.

---

5. General Concerns and Planning

- Team agreed on the need for:
  - Clearer requirements and scope.
  - Backend refactor to avoid bloated main.py files.
  - Proper prompt handling for stability across agents.
- Some frustration expressed over lack of structure and guidance from project leadership.
- Firose and Kamal will explore prompt standardization and evaluation via Azure/Prompt Flow.

---

6. Next Steps

- Pablo will improve backend structure and cookie/session handling.
- Kamal will adapt plotting for daily data and continue virtual env testing.
- Firose to prototype evaluation agent and explore prompt scoring approaches.
- Everyone to explore how to implement a "systematic evaluation" pipeline.

Meeting Adjourned

Prepared by:

Pablo Periañez Cabrero

7 July 2025