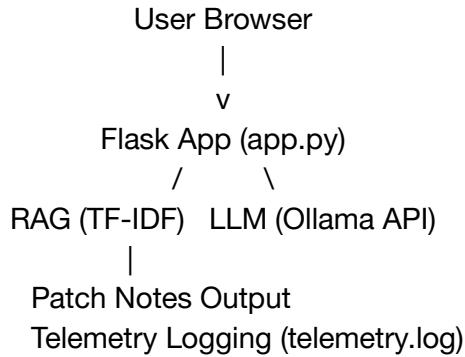


Tech Note

1. Architecture Diagram



2. Guardrails/Safety

- a. System prompt with explicit rules.
- b. Input length guard (max 2000 chars.)
- c. Prompt-injection detection: blocks “ignore previous instructions”.
- d. LLM error handling: fallback message if API fails.

3. Enhancement

- a. RAG: fetch previous patch notes for style guidance.
- b. Does not copy content, only guides LLM formatting and phrasing.

4. Evaluation

- a. Offline tests with tests.json (≥ 15 cases)
- b. Run_tests.py prints pass/fail rate and generates test_results.json.
- c. Patterns are regex/substring based to handle LLM phrasing.

5. Known Limits

- a. Multi-category bullet points may appear only in the primary category.
- b. LLM may rephrase, causing tests to fail minor patterns.
- c. RAG is limited to the seed JSON which doesn't guarantee style perfect matches.