

Agentic Voice-to-Voice AI Assistant for Product Discovery

Prepared By: Soham Mandal, Bruna Medeiros, John Melel,
Bradley Stoller, Aida Sarinzhipova, Jerry Gong

Date: December 12, 2025

Agenda

01

Architecture

02

Product Demo

03

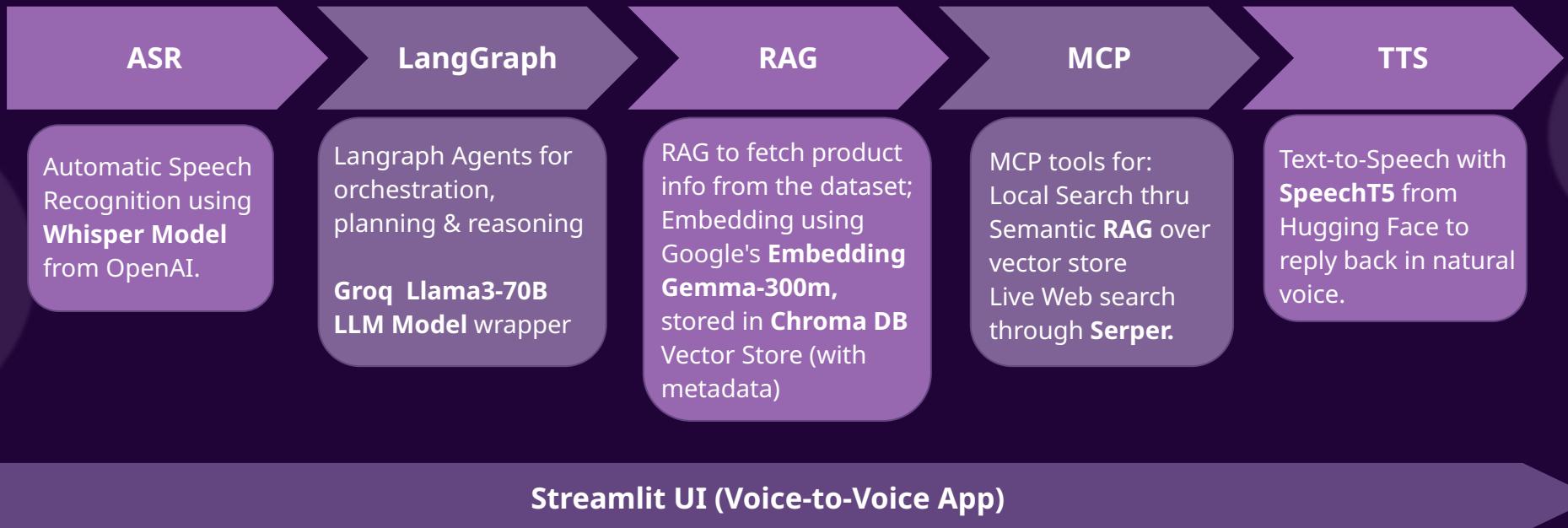
Limitations

The background features a dark purple gradient with a complex network of thin, light-colored lines forming a web-like pattern. Several bright, glowing points of varying sizes are scattered across the surface, some appearing as small dots and others as larger, more intense light sources.

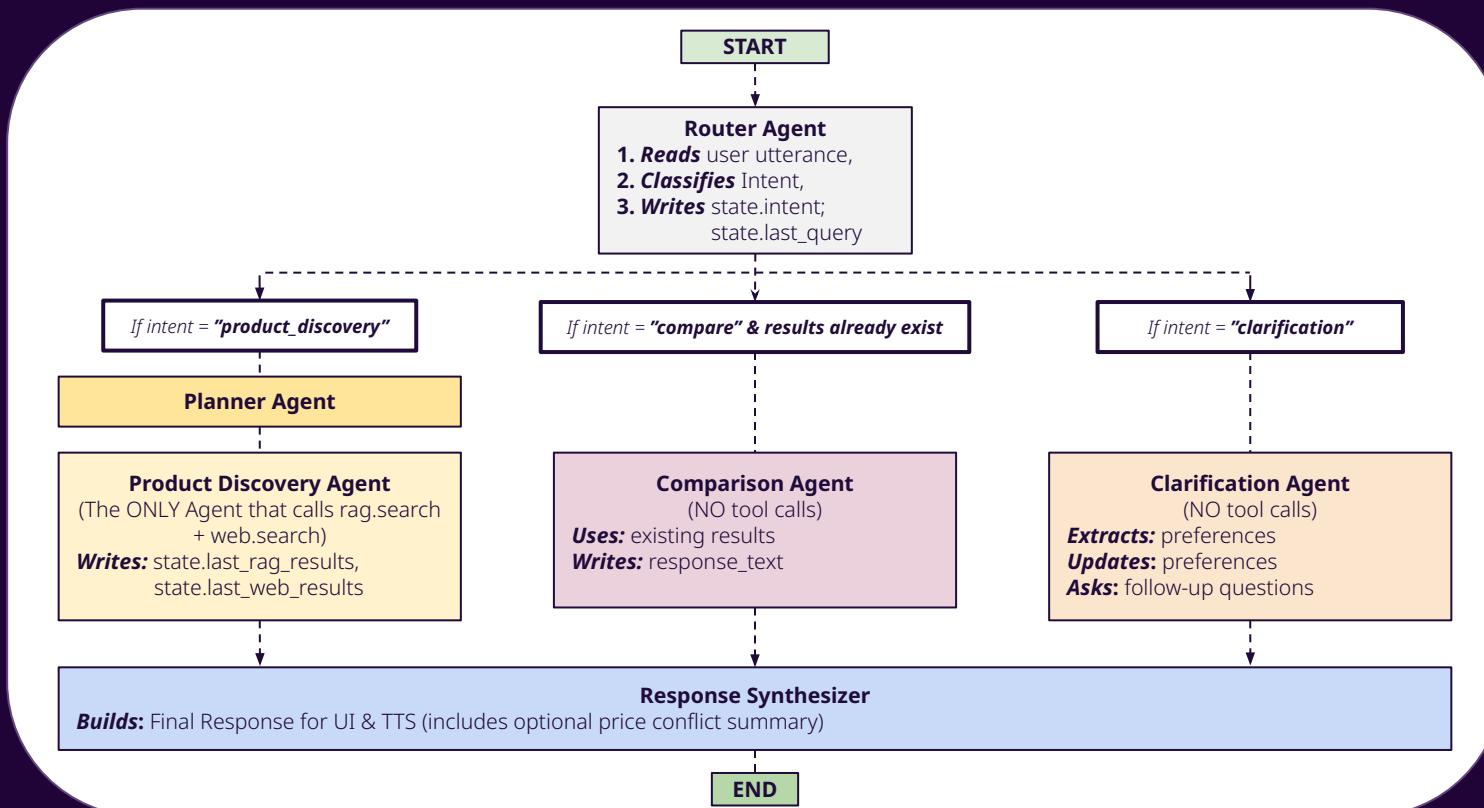
01

Architecture

Project Architecture



Project Architecture

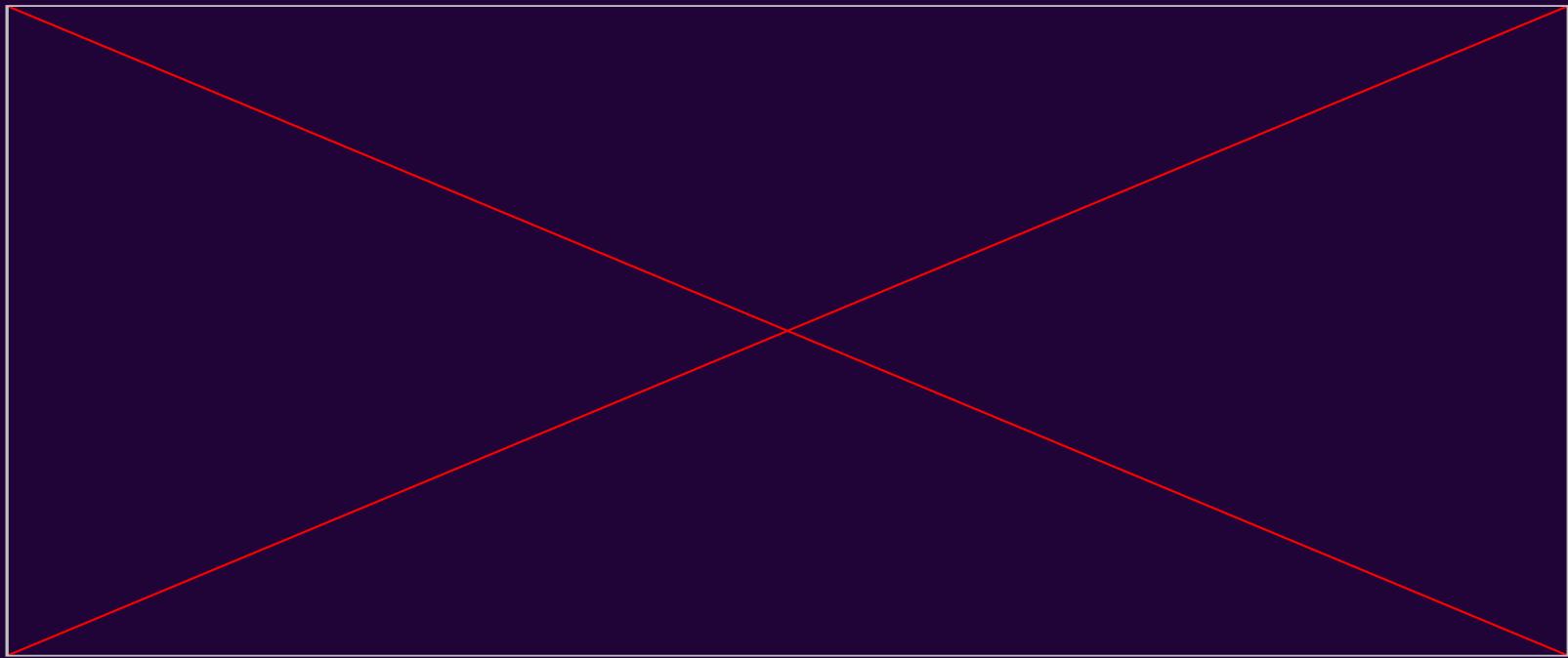


The background features a dark purple gradient with abstract white line art. It includes several glowing magenta dots of varying sizes, some with trails. In the center-left, there are two sets of vertical blue lines forming a shape reminiscent of a sound wave or a double-headed arrow, with small glowing blue dots along them.

02

Product Demo

Product Demonstration



The background features a dark purple gradient with abstract white line art. Several curved lines radiate from the left side, with glowing pink and blue particles at their ends. In the center, there's a cluster of vertical lines forming a shape reminiscent of a speaker or a waveform, with blue and white glowing highlights.

03

Limitations

Limitations (Known Tradeoffs)

Web Search

Live web data adds real-world market context **without** overwriting catalog prices, as real-time price matching is intentionally excluded to maintain modularity.

Embeddings

Fixed embeddings may lose semantic alignment over time, requiring re-embedding to stay accurate in production systems.

Intent classification relies on LLM reasoning combined with heuristic rules rather than supervised models, introducing occasional misclassification for short or ambiguous queries.

Intent

Latency increases due to multi-agent orchestration, speech processing, and MCP tool calls, which would require optimization for real-world deployment.

Latency



Thank You!

Questions?

[GitHub Link](#)