

Google Cloud Next: agents, prompts, and the push to operational AI

April 15, 2025



At Google Cloud Next 2025, the focus wasn't just on new models — it was on structure. Across their announcements, Google made a clear statement: AI is maturing from a tool to a system.

For Flux, the biggest developments came in **agentic infrastructure**. Google's new **Vertex AI Agent Builder** and **Agentspace** platforms aim to support multi-agent workflows — where models talk to each other, make decisions, and complete tasks across tools and data sources. The new **Agent2Agent protocol** underpins this, enabling models to coordinate rather than simply respond. It's early, but it's a clear nod toward AI systems that resemble orchestration more than chat.

This theme carried through to **Gemini 2.5**, Google's latest flagship model which has been [well received by AI communities](#). The *Pro* version offers a 1 million-token context window and is tuned for complex reasoning and coding; *Flash* is its faster, lower-cost sibling for real-time applications. Other model updates — like **Veo 2** (video generation with inpainting and camera control) and **Lyria** (text-to-music) — push the boundaries of creative AI, though their enterprise use cases are still forming.

On the tooling side, **Gemini Code Assist** is now embedded in Android Studio, offering real-time coding help and documentation support. And the **Application Design Center** introduces a more visual approach to deploying apps, with templates and collaboration features baked in.

Also notable: Google dropped a practical [Gemini Prompting Guide](#) — a 45-page handbook designed for Workspace, but applicable well beyond. It reinforces prompting as a design discipline, breaking it into four elements: *persona*, *task*, *context*, *format*. It's a small but important step toward making prompting systematic and repeatable, especially inside teams.

Why it Matters

Taken together, the announcements point to a shift: less emphasis on novelty, more on making AI usable at scale — for coding, content, search, or support. For teams building systems like Flux, it's a nudge to think more in terms of *agents*, *interfaces*, and *workflows* — not just models.

Rabbit hole

[Watch everything released at Google Cloud in 12-minutes](#)

[Google's latest prompting guide](#)

[Do people rate Gemini 2.5 Pro? How does it compare to the recent Meta Llama 4 release?](#)