

The future of vibe-coding with Antigravity



Guillaume Vernade

Gemini Developer Advocate - Google Deepmind



Guillaume Vernade

Gemini Developer Advocate
Google Deepmind



Models & Research

- LearnLM
- AlphaFold 3
- Gemini 1.5 Flash
- Imagen 3
- Gemma 2
- Veo
- Project Astra

2024

May — June — July — Aug — Sept — Oct — Nov — Dec

- Gemini Live
- NotebookLM
Audio Overviews
- AI Overviews
- VideoFX

Shipping at Relentless Pace

- Genie 2
- Gencast
- Gemini 2.0 Flash
- Veo 2
- Gemini 2.0 Flash Thinking
- Project Mariner

Gemini 2.0
Flash-Lite

Gemini 2.0 Pro

Weather
Next

2025

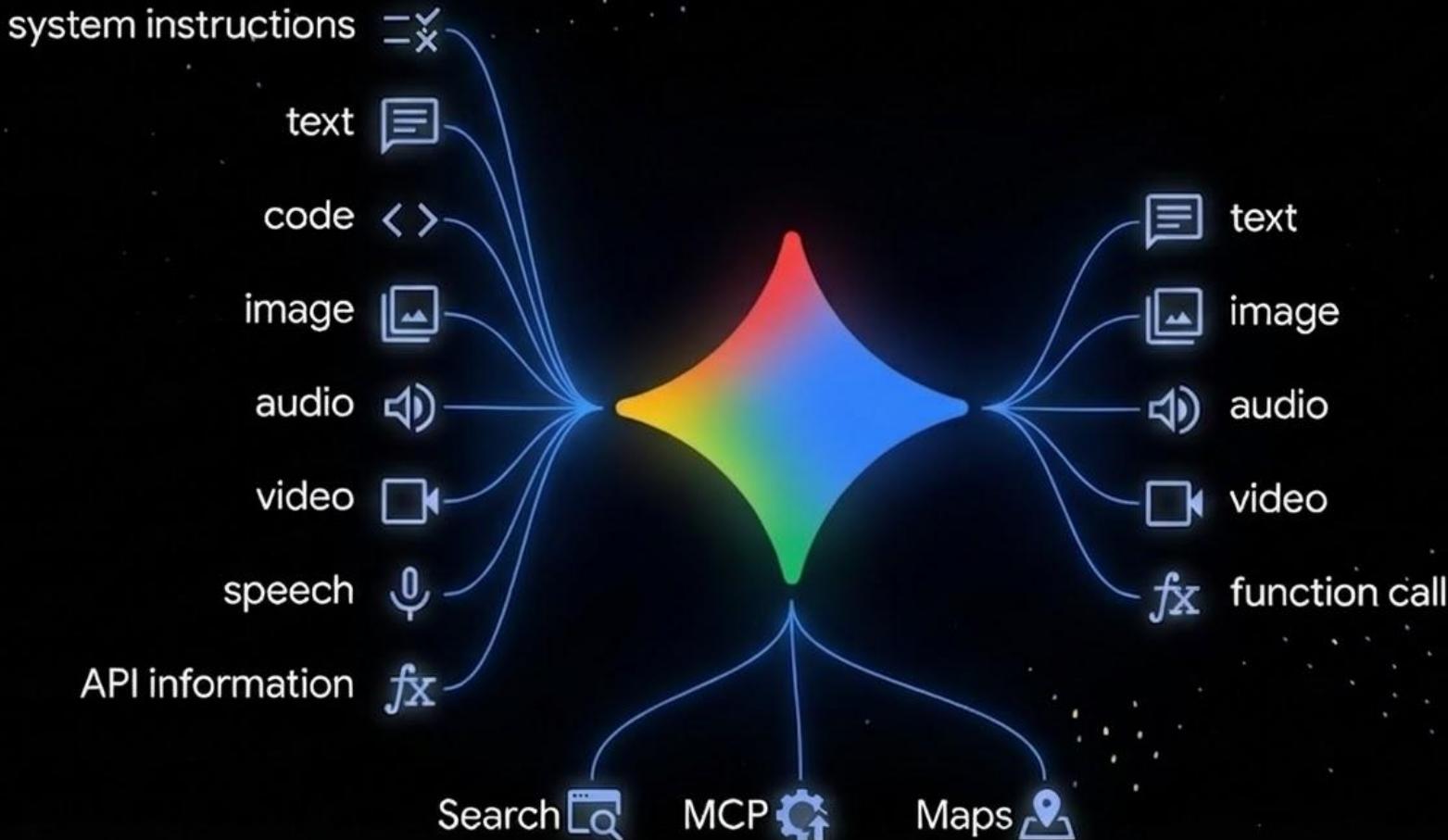
Jan — Feb — March — April — May

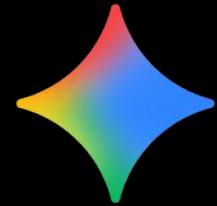
- Gemini Code Assist
- Quantum Willow Chip
- Deep Research
- Native Image Gen
in Gemini 2.0 Flash
- Jules
- Agentspace
- Career Dreamer
- Veo 2 in
Dream Screen
- AI Mode in
Search Labs
- Live Video &
Screenshare
in Gemini App
- Gemini Canvas
- Gemini with
Personalization
- Gemini on Wear,
Auto & TV

Products & Features

- Gemma 3
- Gemini Robotics
- Chirp 3
- Gemini 2.5 Pro (03-25)
- Gemini 2.5 Flash
- AlphaEvolve
- Gemini 2.5
(I/O Edition)

Gemini inputs and outputs





Gemini 3

SOTA Reasoning & Multimodality

Agentic Coding

Advanced Tool Use & Planning

Shifting to an era of action

UNDERSTANDING

Gemini 1

THINKING

Gemini 2

ACTION

Benchmarks

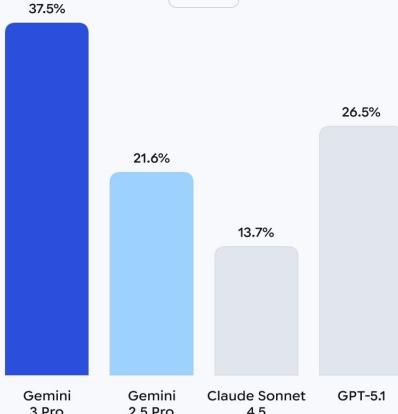
Gemini 3 Pro



Humanity's Last Exam

Reasoning & knowledge

Tools off



Methodology:

Gemini results: All Gemini 3 scores are pass @1 and were obtained by running with the Gemini API for the model-id gemini-3-pro with default sampling settings.

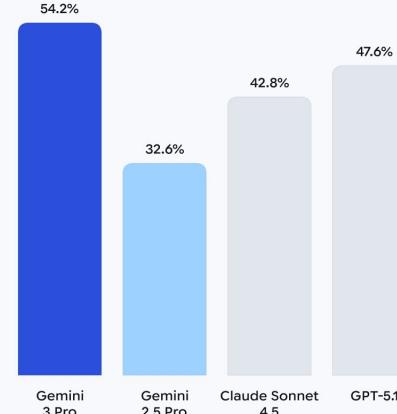
Humanity's Last Exam results for Gemini 2.5 Pro and Claude Sonnet 4.5 are from Scale AI leaderboard and GPT-5.1 from Artificial Analysis.



Terminal-Bench 2.0

Agentic coding

Tools off



Methodology:

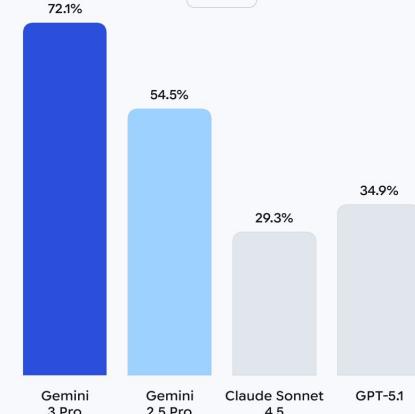
Gemini results were obtained by running with the Gemini API for the model-id gemini-3-pro and gemini-2.5-pro with default sampling settings. All of the results use the same Terminus+2 agent. Results come from <https://www.tberlin.ai/leaderboard/terminal-bench/2.0?agents=Terminus+2>



SimpleQA Verified

Parametric knowledge

Tools off



Methodology:

All Gemini, Anthropic, and OpenAI results are pass @1. Gemini results were obtained by running with the Gemini API for the model-id gemini-3-pro and gemini-2.5-pro with default sampling settings. Results come from <https://www.kaggle.com/benchmarks/simpleqa-verified>.

Benchmarks

Benchmark	Description	Gemini 3 Pro	Gemini 2.5 Pro	Claude Sonnet 4.5	GPT-5.1
Humanity's Last Exam	Academic reasoning	No tools 37.5%	21.6%	13.7%	26.5%
ARC-AGI-2	Visual reasoning puzzles	ARC Prize Verified 31.1%	4.9%	13.6%	17.6%
GPQA Diamond	Scientific knowledge	No tools 91.9%	86.4%	83.4%	88.1%
AIME 2025	Mathematics	No tools With code execution 95.0% 100%	88.0% —	87.0% 100%	94.0% —
MathArena Apex	Challenging Math Contest problems	23.4%	0.5%	1.6%	1.0%
MMMU-Pro	Multimodal understanding and reasoning	81.0%	68.0%	68.0%	TBC
ScreenSpot-Pro	Screen understanding	72.7%	11.4%	36.2%	3.5%
CharXiv Reasoning	Information synthesis from complex charts	81.4%	69.6%	68.5%	69.5%
OmniDocBench 1.5	OCR	Overall Edit Distance, lower is better 0.115	0.145	0.145	0.147
Video-MMMU	Knowledge acquisition from videos	87.6%	83.6%	77.8%	80.4%
LiveCodeBench Pro	Knowledge acquisition from videos	Elo Rating, higher is better 2,439	1,775	1,418	2,243
Terminal-bench 2.0	Agentic terminal coding	Terminus-2 agent 54.2%	32.6%	42.8%	47.6%
SWE-Bench Verified	Agentic coding	Single attempt 76.2%	59.6%	77.2%	76.3%
t2-bench	Agentic tool use	85.4%	54.9%	84.7%	80.2%
Vending-Bench 2	Agentic long term planning	Net worth (mean), higher is better \$5,478.16	\$573.64	\$3,838.74	\$1,473.43
FACTS Benchmark Suite	Held out internal grounding, parametric, MM, and search retrieval benchmarks	70.5%	63.4%	50.4%	50.8%
SimpleQA Verified	Parametric knowledge	72.1%	54.5%	29.3%	34.9%
MMMLU	Multilingual Q&A	91.8%	89.5%	89.1%	91.0%
Global PIQA	Commonsense reasoning across 100 Languages and Cultures	93.4%	91.5%	90.1%	90.9%
MRCR v2 (8-needle)	Long context performance	128k (average) 1M (pointwise) 77.0% 26.3%	58.0% 16.4%	47.1% not supported	61.6% not supported

Benchmarks

LMArena



1501

#1 on LMArena

Rank ↕	Rank Spread ⓘ (Upper-Lower)	Model ↗	Score ↓
1	1 ↔ 2	gemini-3-pro	1501
2	1 ↔ 3	grok-4.1-thinking	1484
3	2 ↔ 5	grok-4.1	1465
4	3 ↔ 8	gemini-2.5-pro	1451
5	3 ↔ 9	claude-sonnet-4-5-20250929-thinking-32k	1449
6	4 ↔ 8	claude-opus-4-1-20250805-thinking-16k	1449
7	4 ↔ 12	claude-sonnet-4-5-20250929	1445

18 Nov 2025

Results are from voting pre-release,
and tagged as preliminary

Benchmarks

WebDev Arena

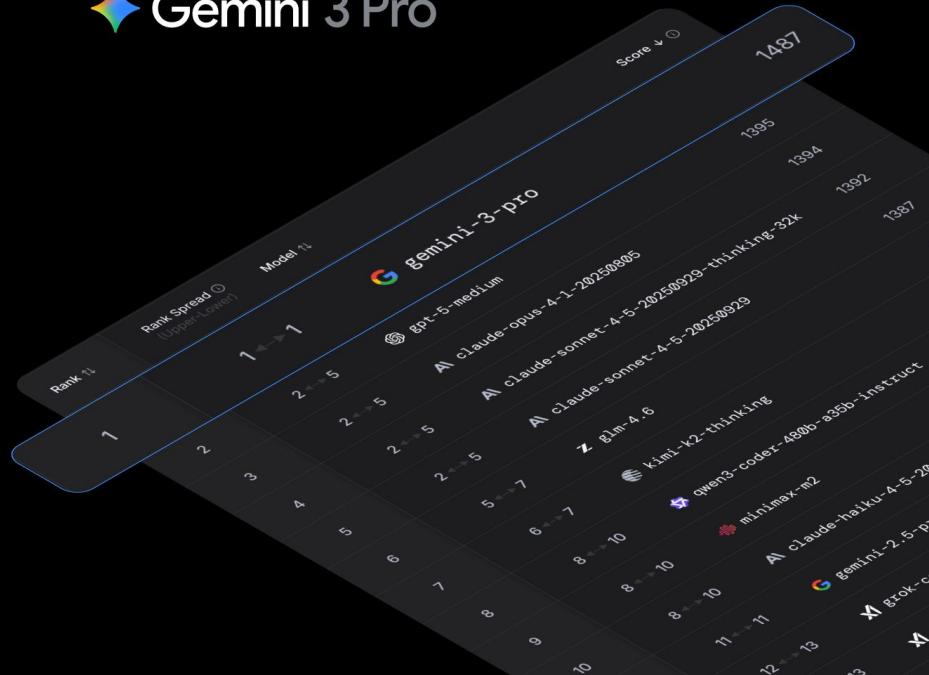
1487

#1 on WebDev Arena

18 Nov 2025

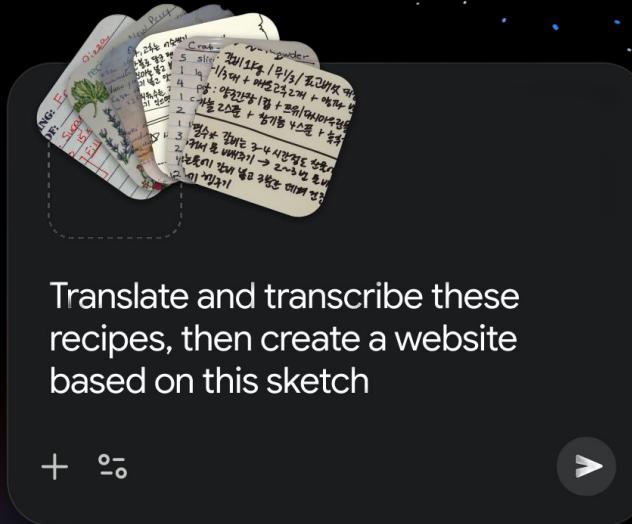
Results are from voting pre-release,
and tagged as preliminary

◆ Gemini 3 Pro



SOTA Reasoning & Multimodality

- Analyzes text, audio, video, and code within a single, industry-leading 1M context window
- Delivers highly factual reasoning for complex inputs (e.g., medical scans, logs)
- Handles difficult data like overlapping speakers or blurry text for maximum insight extraction

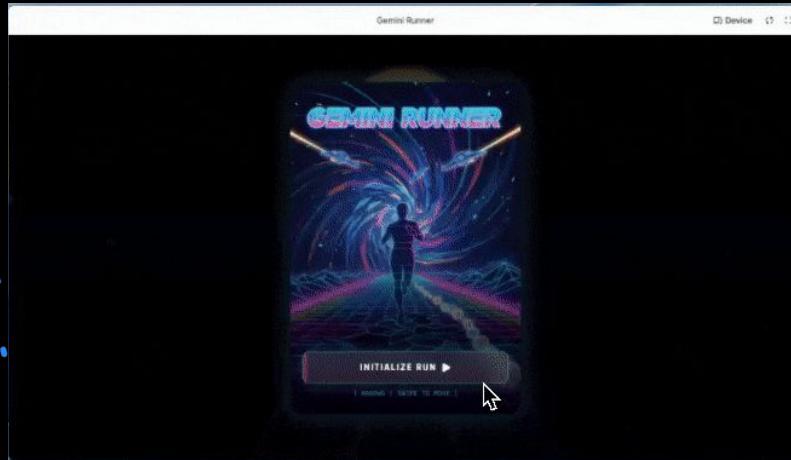


Translate and transcribe these recipes, then create a website based on this sketch

+ ⌂ ➤

Agentic Coding

- **Core Skill:** From code completion to system refactoring
- **Safety:** Comprehensive safety evaluations for code



Gemini 3



Advanced Tool Use & Planning

- Precise interaction with external APIs and systems with minimal prompt tuning
- Orchestrates complex, multi-step workflows, across disparate enterprise data
- Untangles ambiguous business problems to connect high-level strategy with autonomous execution

★ Gemini 3 Pro



Application



Gemini

Nano-banana

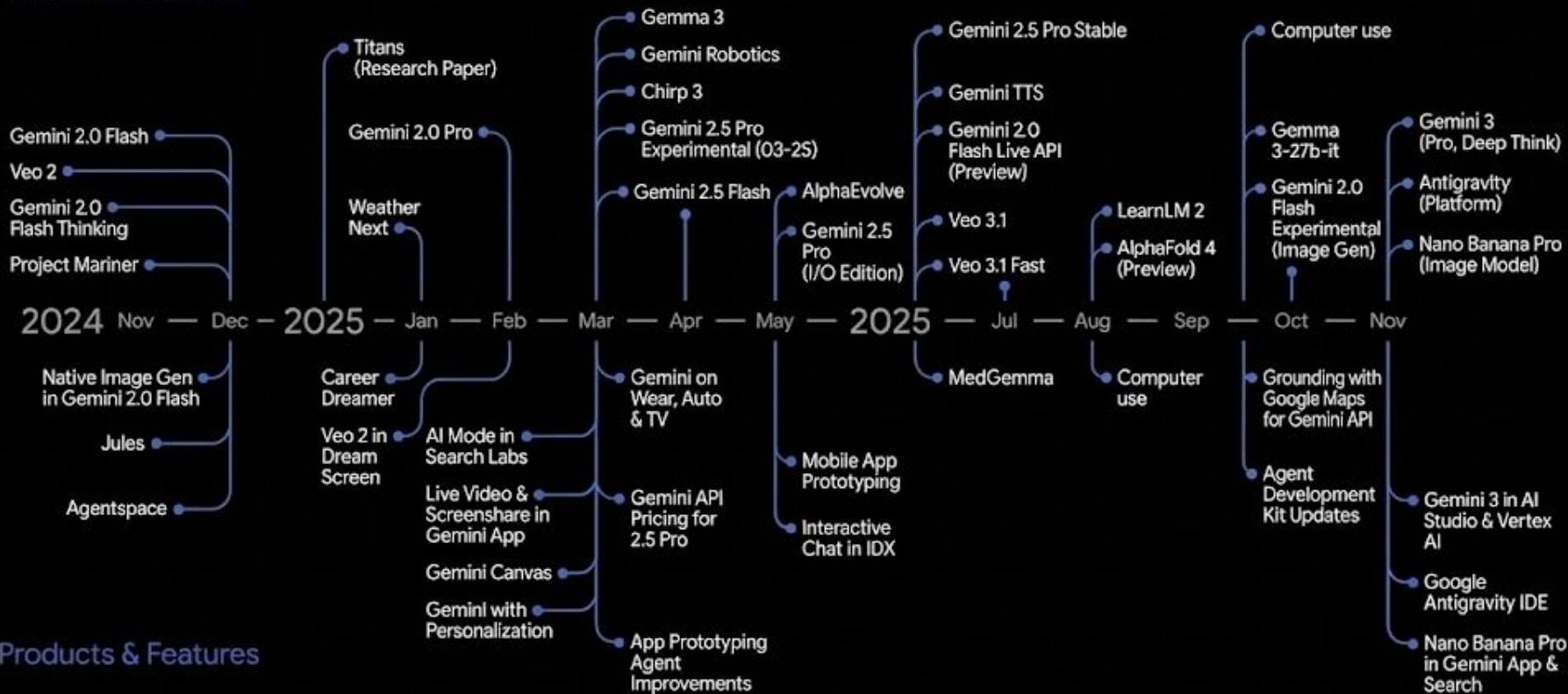
Search Grounding

Search the web then generate an image of isometric perspective, detailed pixel art that shows the career of Guillaume Vernade



Shipping at Relentless Pace

Models & Research



Products & Features

Beyond Gemini models

Image gen (Imagen & Nano-banana)



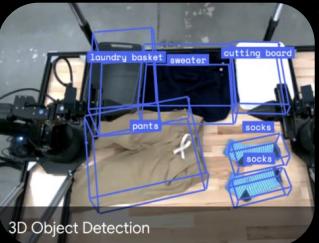
Video generation (Veo)



Music gen (Lyria)



Gemini Robotics



Agents (Mariner, Jules)



Open models

Gemma 3

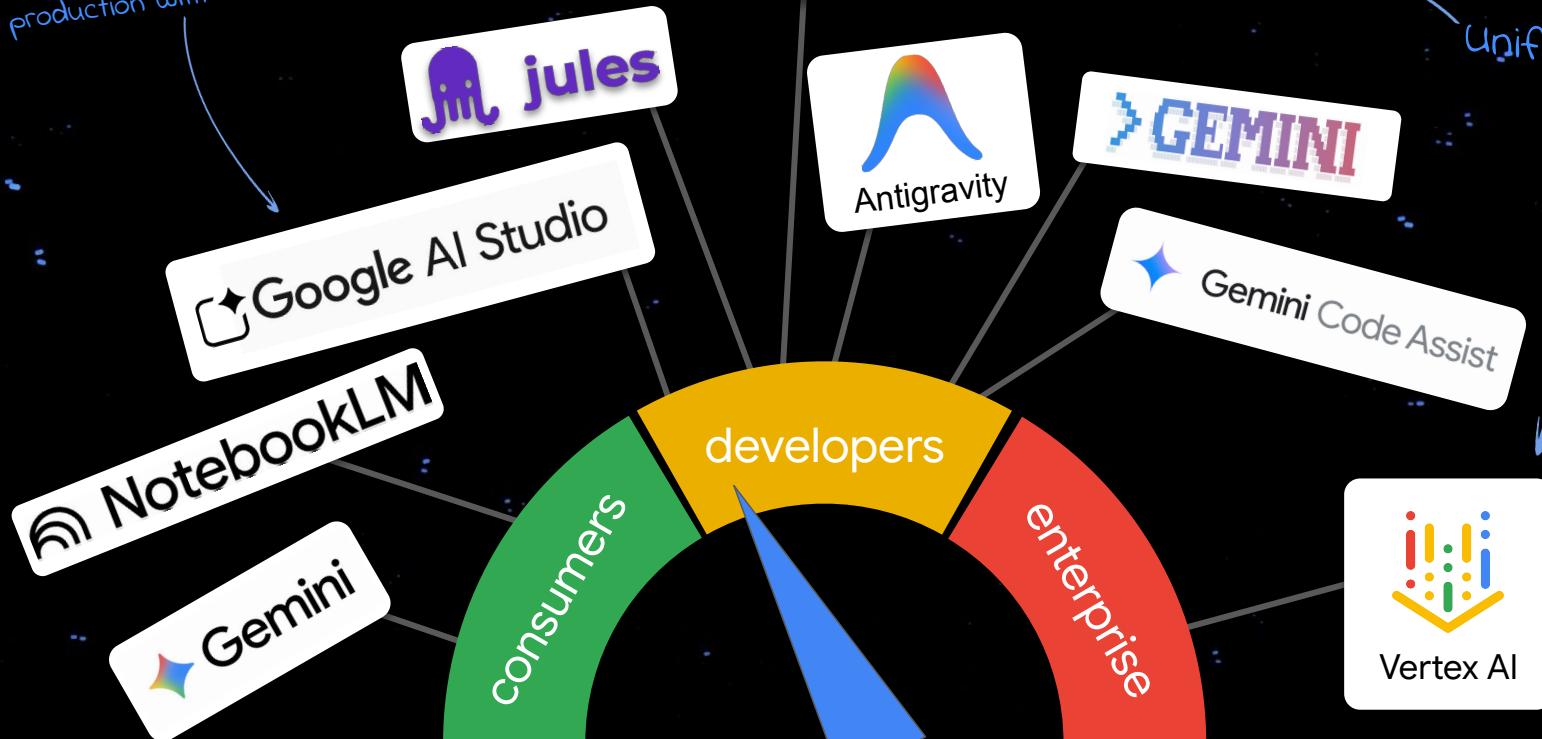
And more...



↳ The fastest path from prompt
to production with Gemini

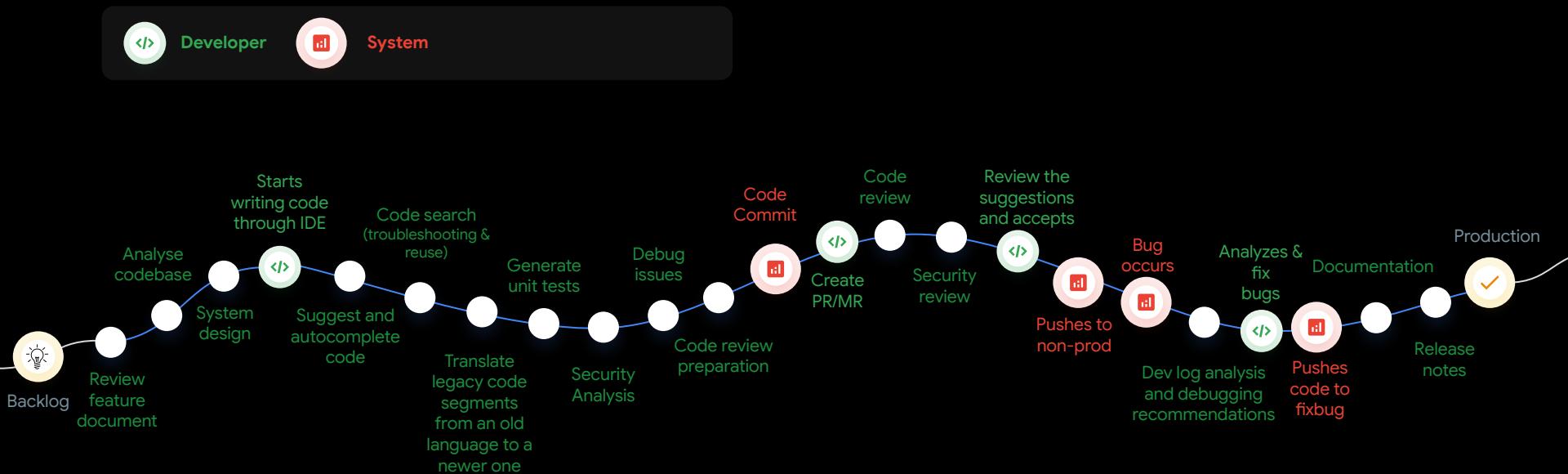
Gemini Developer API

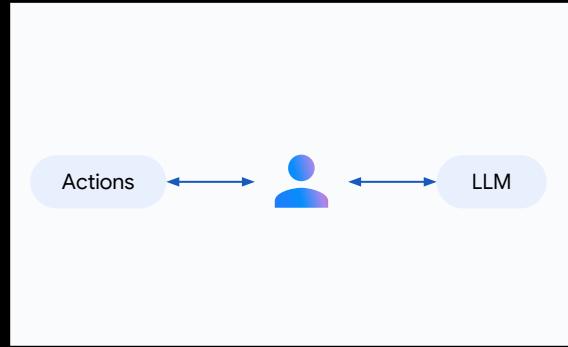
Unified SDK



and others!

Typical Workflow





Prompt and Response

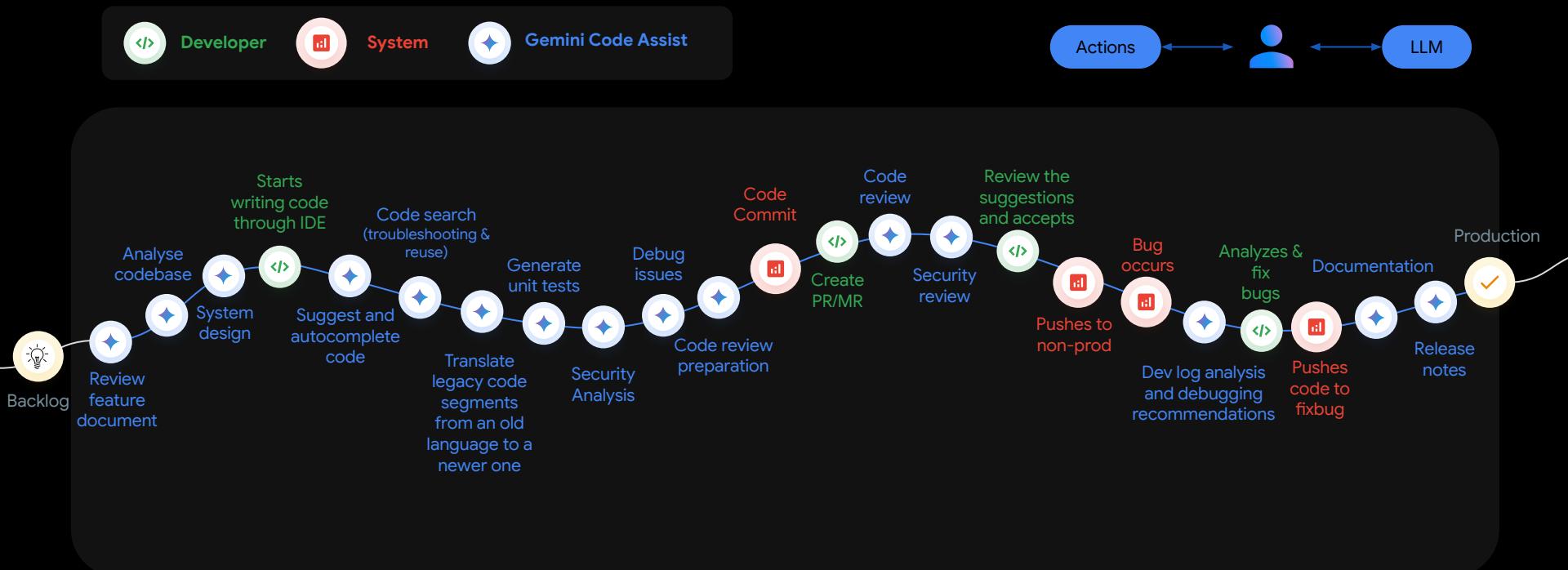
Single Step Inference Chat

Simple tasks

Humans perform actions

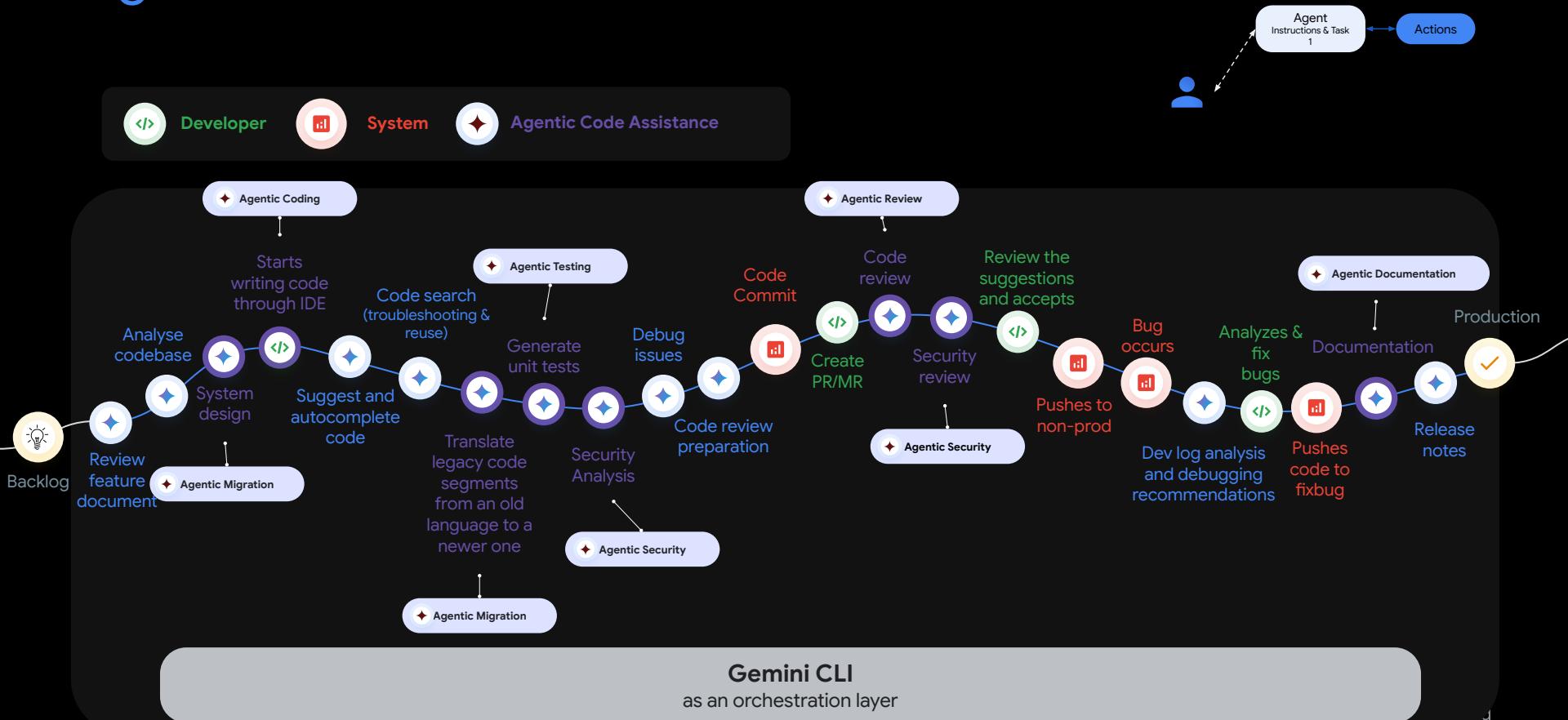
Humans in the loop

Typical Workflow with AI Assistance

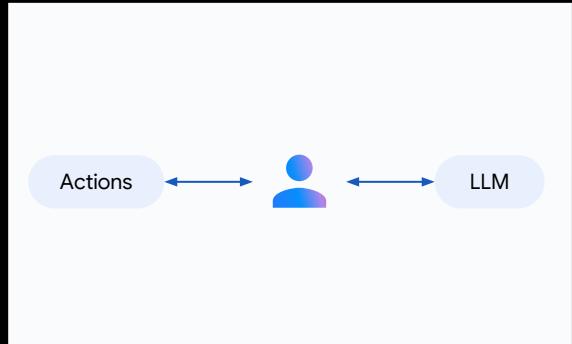


ai.studio

Agent Mode and SDLC



Evolution of AI and Agents



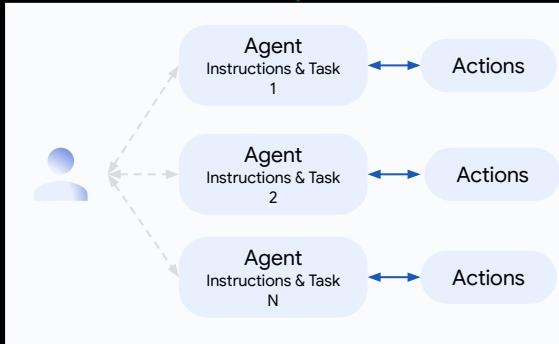
Prompt and Response

Single Step Inference Chat

Simple tasks

Humans perform actions

Humans in the loop



Interactive Agents

Multi step inference

Complex tasks

Agents perform actions

Humans less in the loop

Introducing



Google Antigravity

A new **Agentic Development Platform** from Google DeepMind.

Antigravity noun

an-tē-'gra-və-tē

AI development platform with:

- 1) Editor
- 2) Browser
- 3) Agent Manager

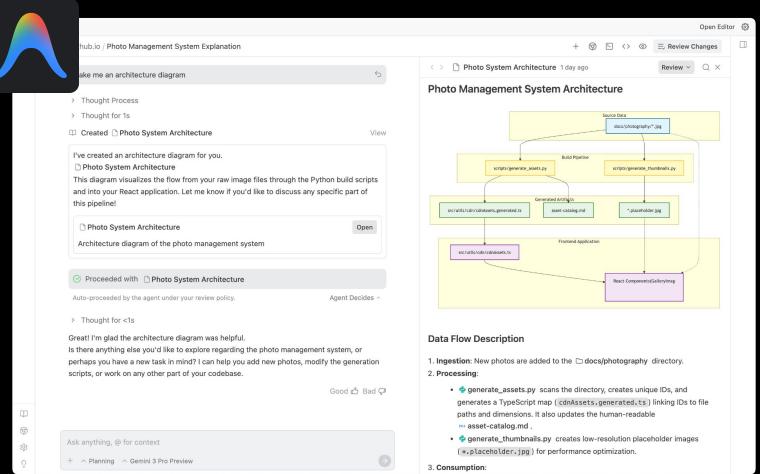
The screenshot shows the AI Editor interface with a dark theme. On the left is a sidebar with project files like 'index.js', 'App.js', 'index.html', 'index.css', and 'index.js.map'. The main area displays a code editor with the following snippet:

```

    const App = () => {
      return (
        <div>
          <h1>Welcome to React</h1>
          <p>This is a simple example to get you started with React.</p>
          <p>Edit <a href="https://reactjs.org/docs/getting-started.html#using-create-react-app">this file</a> and save to reload</p>
        </div>
      );
    }
  
```

A floating window titled 'Ask anything, @ for context' is open at the bottom, with suggestions for 'Planning' and 'Gemini 3 Pro Preview'.

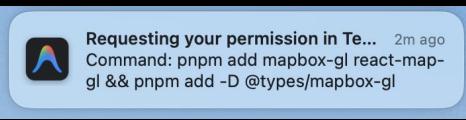
AI Editor



Agent Manager



Agent Controlled Browser



The screenshot shows a GitHub inbox interface with the following notifications:

- Install Mapbox with pnpm now** (Pending)
khou22.github.io
> git branch --show-current
Running...
Actions: Open, Run
- Structure Agent Manager Docs** (1 day ago)
gweb-jetski
I have fixed the sidebar redundancy issue. The 'Agent Manager' section header should now correctly link to the 'Agent Manager' page without showing a duplicate child item.
Idle
Actions: Open
- Awaiting User Task** (1 day ago)
flight-tracker
Idle
Actions: Open
- Architecture Diagram for Project** (2 days ago)
khou22.github.io
I've created an architecture diagram for your project. Please review it in [architecture.md](#).
Blocked
Actions: Open, Proceed

Age of Artifacts

Long live the agent manager.

What is an Artifact?

Dynamic representations of information specific to you and your use case.

They help the agent:

- Maintain focus
- Clearly communicate to the user, other agents, subagents

The screenshot shows a desktop application window titled "flight-tracker / Flight Lookup Next.js App". The main view displays a search interface with a search bar containing "AA123" and a "Search" button. Below the search bar, a message states: "1. Navigated to <http://localhost:3000>. 2. Verified the page loaded with the heading "Flight Tracker" and a search input. 3. Typed "AA123" into the search input and clicked search. 4. Verified flight details for AA123 (American Airlines, JFK to LAX, with times and time zones) were displayed. 5. Cleared the input, typed "INVALID", and clicked search. 6. Verified the "No flights found for that number." message was displayed." A note below says, "The application works as expected according to the test plan." Below this, there are two items in a list: "Playback available" and "Thought for 7s".

On the right side of the application window, there is a sidebar with sections for "Walkthrough now", "Review", and "Enter a flight number". The "Walkthrough" section contains a summary of the verification steps. The "Review" section has a "Review" button. The "Thought for 7s" section also has a "View" button.

Below the application window, the desktop environment shows a dock with icons for Finder, Home, and Mail. The status bar at the bottom indicates battery level, signal strength, and system status.

On the far right, a vertical sidebar titled "Flight Tracker Verification" lists various artifacts:

- Artifacts (1)
 - Flight Stack API Research
 - Implementation Plan
- Task
- Walkthrough
- Media Artifacts
 - Flight Lookup Verification (Nov 17 8:52...)
 - Invalid Flight Results (Nov 17 8:52 PM)
 - Valid Flight Results (Nov 17 8:52 PM)
- Files Changed
 - api.js src
 - main.js src

Multi-Modal Artifacts

Flight Lookup Web App

Implementing UI
Implemented the Mock API. Now building the user interface and cleaning up default Vite files.

Files Edited
(1) style.css (1) index.html Task

Progress Updates
 1 Updating HTML and CSS, cleaning up unused files
(1) Edited Task

```
~/flight-tracker $ rm src/counter.js src/javascript.svg
Ran terminal command Open in Terminal ⌘ - Exit code 0
```

View Auto ▾

Open diff

Open diff

Thought for 7s
 Reflecting on Implementation
 I've just finished updating `index.html` and `style.css`, and I'm ready to begin the core logic in `src/main.js`. My current plan involves adjusting the task boundary in `task.md` to indicate the UI is complete and logic is now underway. After that, I'll dive into the JavaScript coding.

Progressing with Code
 I'm now in the process of implementing the core logic. I've updated the task boundary to reflect this shift. In `src/main.js`, I've imported `searchFlight`s, and added a form event listener. I'm actively handling submissions and calling `searchFlights`. The next step involves rendering the results into the DOM. I'm looking forward to getting the dev server running for verification.

Implementing Logic
Implemented Mock API and UI (HTML/CSS). Now wiring up the logic in main.js.

Files Edited
(1) main.js Task

Progress Updates
 1 Writing main.js to handle form submission and render results
(1) Edited (1) main.js +77 -23

View Open diff

Thought for 7s
 Rethinking HTML injection
 I'm now reconsidering the way I'm handling the HTML structure in this app. I've realized my `main.js` is replacing the contents of `#app`, which makes the initial structure in `index.html` redundant. It's making me ponder whether I should keep the HTML in `index.html` and just attach event listeners in `main.js`. While the current approach of injecting HTML from JavaScript is common, especially with frameworks, I'm thinking a more pure Vanilla JS approach, with HTML in `index.html`, might be clearer. I need to make a final decision, especially if I can make the content appear on the page without replacement.

Verifying HTML Redundancy
 I'm checking the duplication of `main.js` and the `index.html` structure. It's not a dealbreaker yet; the current redundancy is manageable. Now I'm preparing to verify the system. I'll make a `task.md` entry. Let's make sure things are working.

Review Changes

Ask anything (M&L), @ to mention, / for workflows

+ ~ Planning ~ Gemini 3 Pro (High)

VS.

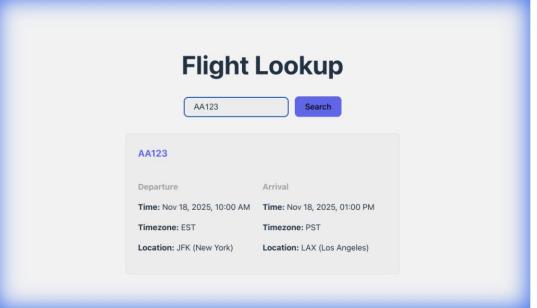
Walkthrough Final 55 secs ago Review

Walkthrough
 Built a Flight Lookup Web App with Vite + Vanilla JS.



Code Changes
 Initialized a Vite project, added `mockApi.js` for simulated flight data, created a responsive dark-themed UI in `index.html` and `style.css`, and implemented async search logic in `main.js`.

Verification
 Verified via automated browser test and manual checks.



Flight AA123 Results

AA123

Departure	Arrival
Time: Nov 18, 2025, 10:00 AM	Time: Nov 18, 2025, 01:00 PM
Timezone: EST	Timezone: PST
Location: JFK (New York)	Location: LAX (Los Angeles)

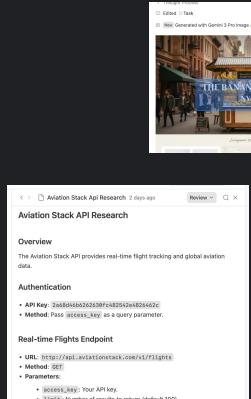
Try it

npm run dev

Search for AA123.

The *model* decides:

- **If** it should generate an artifact
- **What** that artifact should contain
- **Why** it is needed
- **Who** needs to see it (subagent, other agents, notify user, etc.)



Implement Plan Test

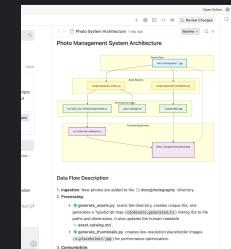
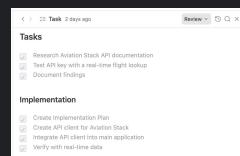
IMPORTANT
Open Questions

1. What specific features should be included in this test plan?
2. Are there any particular constraints or requirements to consider?
3. Should this plan link to any existing code or artifacts?

Please review the plan in [implementation_plan.md](#).

[Implementation Plan](#) [Open](#)

Plan to implement the Aviation Stack API for flight lookup, replacing the mock API.



Plan & Research

Feedback

Execution

Walkthrough

Memory

Expand all <
d TypeScript interfaces.

Overview

The Aviation Stack API provides real-time and historical flight data. The base URL is <http://api.aviationstack.com/v1/>. Authentication is done via an `access_key` query parameter.

Test Request

```
curl "http://api.aviationstack.com/v1/flights?  
access_key=YOUR_ACCESS_KEY&limit=1"
```

Use the key I gave you in the `.env.local`

Cancel Add Comment

```
"pagination": {  
    "limit": 1,  
    "offset": 0,  
    "count": 1,  
    "total": 75659  
},  
"data": [
```

Comments on Markdown Artifacts

> Thought Process

Edited Task

New Generated with Gemini 3 Pro Image 🎉

Review

Add a CTA to find my nearest banana stand

Cancel Add Comment

Visual comments on multi-modal

Implementation Plan • 2

LGTM. Send it!

+ ▲ Planning ▲ Gemini 3 Pro (High)

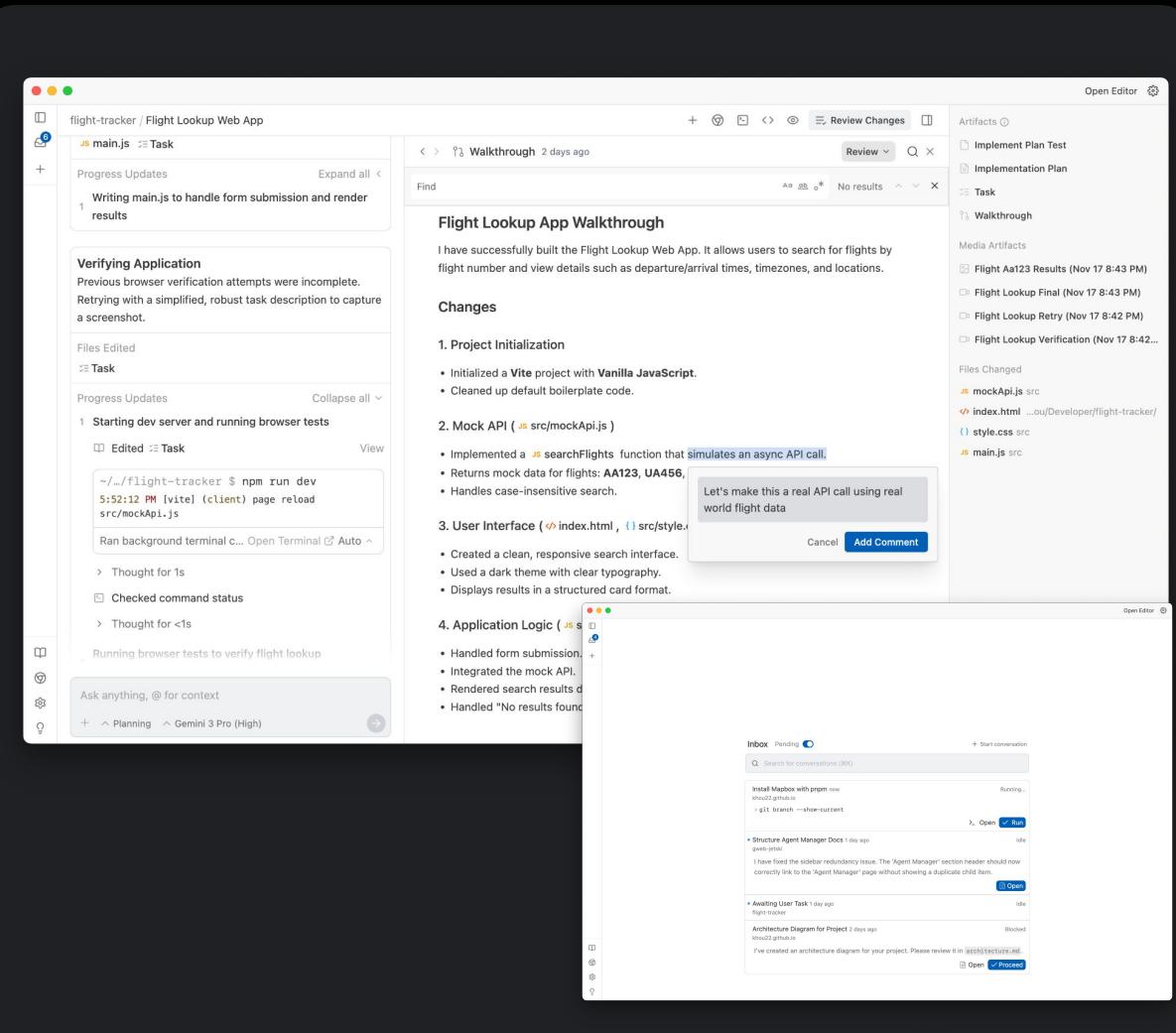
Submit →

Agent Manager

Form factor as models advance is going to look very different.

Must support:

- Artifacts
- Parallel subagents
- Background agents
- Code at a surgical level



Start new conversation in ▾ CollabWhiteboard

 View Inbox

|Ask anything, @ for context

+ ▾ Fast ▾ Gemini 3 Pro Preview



< > Open editor Use Playground

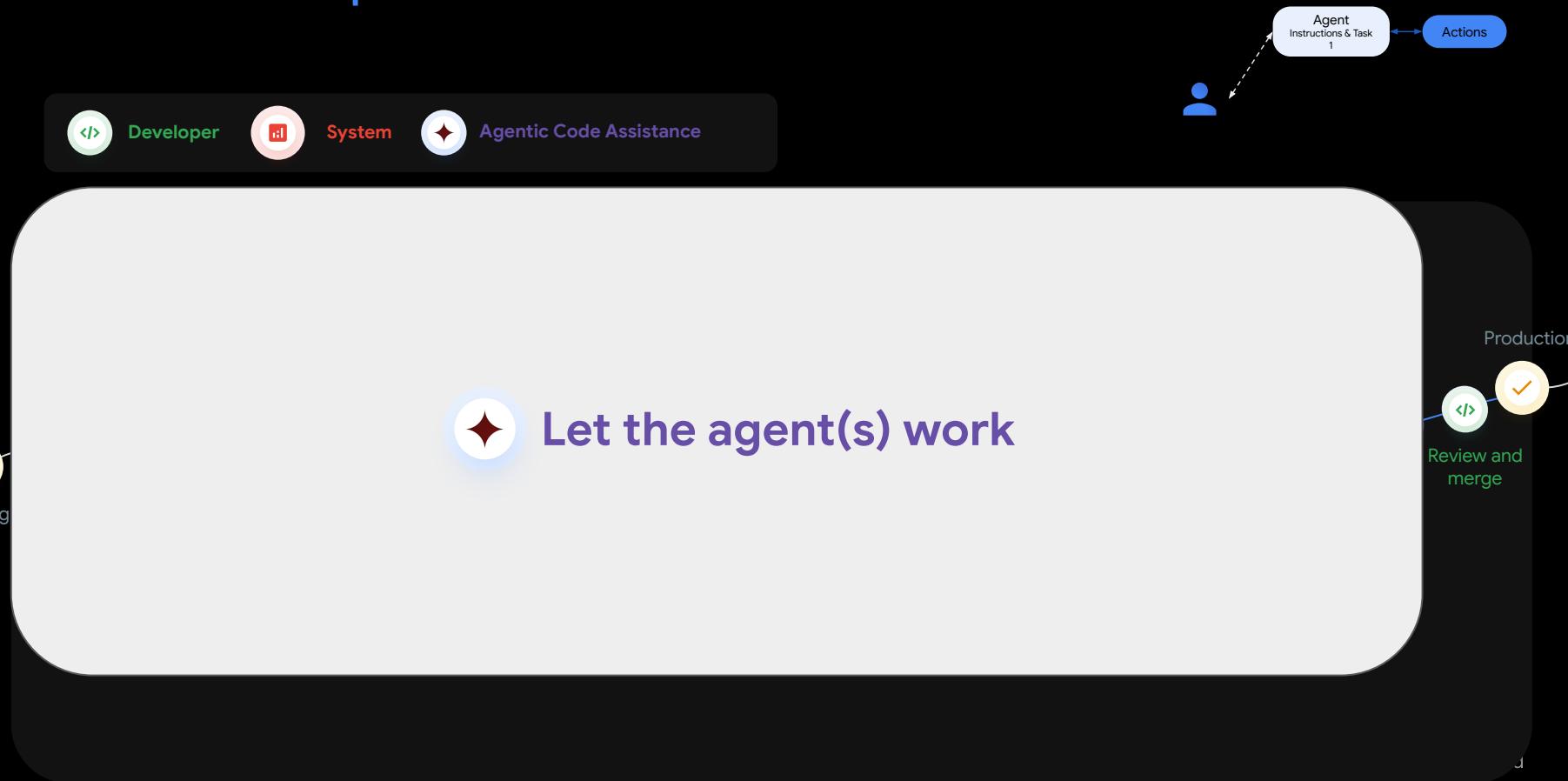




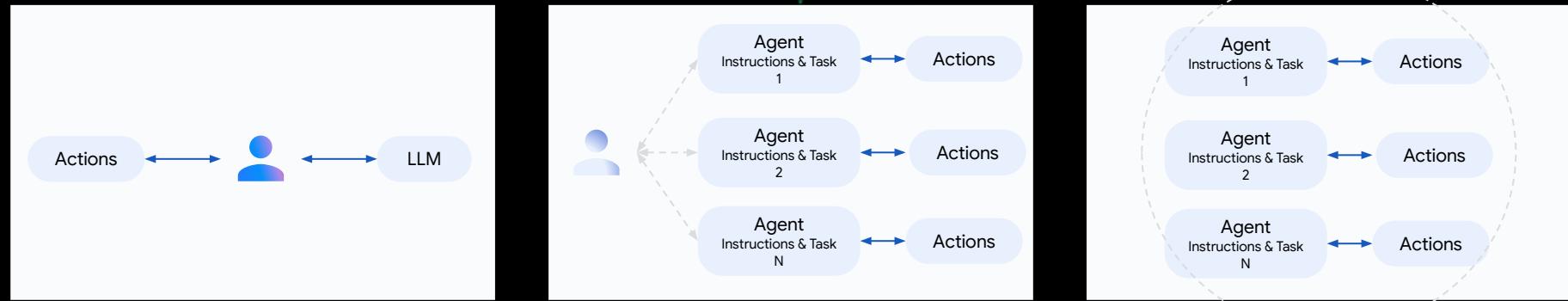
Gemini CLI



The next step



Evolution of AI and Agents



Prompt and Response

Single Step Inference Chat

Simple tasks

Humans perform actions

Humans in the loop

Interactive Agents

Multi step inference

Complex tasks

Agents perform actions

Humans less in the loop

Autonomous Agents

Automatically triggered

Work in background

Customizable

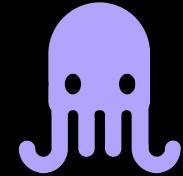
Humans not in the loop

Jules

Your experimental
AI-powered code
agent

Jules

An **asynchronous coding agent** that helps you fix bugs, add documentation, and build new features.



```
:xx:  
XMMO :MMW  
+ MMx ,MMM XMMc :MM .d.MMMMKL ,o MM.MMW,  
OMMx /MMM X.Mc :M.W ,WMN) .cWMX ,MM:  
OMMx .MMM XMMc :MMW OMMKKKKK.MM, LXMMWNKOo  
OMMx .MMMc :MMc :MMW LMMN, ... ,dMW  
OMMd LMM.W.WWWMMc :MMW :XMMKKNNN..NNBONRNWo  
KKXXMX .clc ,;::: ;::: .clc ,';:cl.'
```

It integrates with GitHub, understands your codebase, and **works autonomously** — so you can move on while it handles the task.

Spaces

Home Assistant



Summary

Timeline

Board

Calendar

List

Forms

Development

Code

Security

Releases

Deployments

More

Search board



J



Group

IDEAS 20

Sonar and co in home
assistant

NAS

HA-2



Cloudflare for NAS

NAS

HA-3



Bluetooth repeater?

Core

HA-4



Add bluetooth dongle

Core

HA-5



Automatically review entities
with thousands of entries and
send a notification on a
weekly basis

Core

TO DO 8

Fix Z2M

Zigbee

HA-12

IN PROGRESS 12

Replace PAC en
fonctionnement helper with a
template

Solisart

HA-8

REVIEW 1

switch off
light.lumiere_penderie
between midnight and 8am
during the week and 9am
during weekends and
wednesdays

Automations

HA-53

TESTI

Recover energy history

Energy

HA-13

Clean configuration.yaml

Yaml

HA-9

Fix efficiency history

Solisart

HA-26

Solisart integration create a
device and has UI set-up

Solisart

HA-16

Replace device_id/entities_id
with entities_ids in all scripts &
automations

Automations

Scripts

HA-34

Review what is recorded

HA-19

Review
integrations/statistics.yaml

Cleanup

Remove everything related to
ebox

Cleanup

HA-38

Giom-V/Home-assistant ▾

Configure repo ⚙

Remove beta label ...



Codebase overview

Suggested

Scheduled

⌚ Automatically find issues in your codebase ⓘ

2 / 5 repo max



⚡ Sessions

MISSION : Agis en tant qu'architecte. Implémente cette nouvelle fonctionnalité proprement....

⌚ Completed 16 days ago



MISSION : Agis en tant qu'architecte. Implémente cette nouvelle fonctionnalité proprement....

⌚ Completed 18 days ago



MISSION : Agis en tant qu'expert en maintenance. Aide-moi à améliorer la configuration...

⌚ Completed 22 days ago



MISSION : Agis en tant qu'expert Home Assistant...

⌚ Completed 19 days ago



MISSION : Agis en tant qu'expert Home Assistant...

⌚ Completed 23 days ago



View more

⌚ Schedule skill-based agents

⚡ Performance

🎨 Design

🛡️ Security

1 Open ✓ 157 Closed

Author ▾ Label ▾ Projects ▾

☒ [HA-54] feat[tracking]/ Add entity tracking sensor ✓

#158 by google-labs-jules (bot) was merged 2 weeks ago

☒ [HA-51] feat[database]/ Add database size sensor ✓

#156 by google-labs-jules (bot) was merged 2 weeks ago

☒ [HA-48] feat(basement): Basement lights failsafe ✓

#155 by google-labs-jules (bot) was merged 3 weeks ago

☒ [HA-41] cleanup/dafang-camera ✓

#154 by google-labs-jules (bot) was merged 3 weeks ago

☒ [HA-42] cleanup/remove-dark-sky ✓

#153 by google-labs-jules (bot) was merged 3 weeks ago

☒ [HA-39] cleanup[google-wifi]/ Remove Google Wifi configuration and integration ✓

#152 by google-labs-jules (bot) was merged 3 weeks ago

☒ [HA-33] refactor(scripts): move scripts from scripts.yaml to script/ folder ✓

#151 by google-labs-jules (bot) was merged 3 weeks ago

☒ [HA-7] improvement[automations]/turn off front yard lights when dining room cover closes ✓

#150 by google-labs-jules (bot) was merged 3 weeks ago

☒ [HA-20] Improve basement stair's light automation ✓

#149 by google-labs-jules (bot) was closed 3 weeks ago • Draft

☒ [HA-20] Improve basement stair's light automation ✓

#148 by google-labs-jules (bot) was merged 3 weeks ago

Agentic Coding Software Development Tips

1. Use Gemini 3 over Gemini 2.5: Both Gemini 3 Pro and Gemini 3 Flash have significant improvements that aid the agentic coding loop.
2. Plan first, code next: Specs, designs, interactive discussions with the models. More clarity helps models better understand and execute.
3. Invest in policy: Spend time setting agentic policy. This makes your agent more predictable, reliable and allows more autonomy.
4. Customize your agents: Skills, extensions, MCPs. Some, but not all. Tooling helps expand model capabilities. Too many can hinder progress.
5. Explore tackling problems with different tools: We have an incredible toolbox. Don't assume everything is a nail. Explore what each tool is good at and where they can be used together.

The Future of AI in Development

From Code Assistant to Team Member

- From simple code completion to AI agents that can work semi-autonomously
- Next breakthrough: AI QA/QC and AI mock-ups

From coding to directing & architecturing

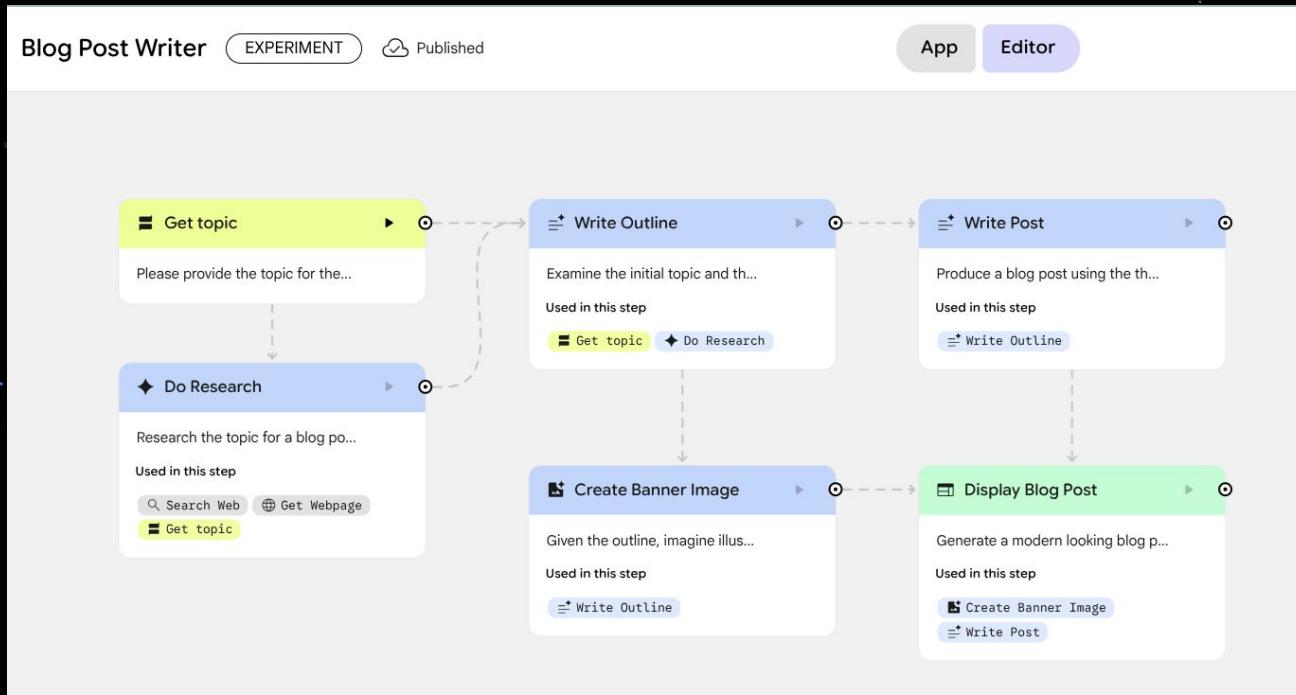
- Directing the AI agents
- Organizing projects for AI
- System instruction engineering

The Rise of No-Code and Low-Code

- Everybody can (no) code!
- Empowering people who are not developers to build applications.

Opal

No-code app and agents builder



Thanks!

► Google AI Studio
goo.gle/aistudio



★ Gemini Code Assist
goo.gle/codeassist



▲ Antigravity
goo.gle/antigravity



🐙 Jules
goo.gle/jules



> GEMINI
goo.gle/gemini-cli-landing

