

# Agents & MCP @ Google Scale

Google-Scale MCP: Building, Scaling, and Governing Your Agents

Google Cloud





**Alan Blount**

Senior Product Manager  
Google Cloud



**Vaibhav Katkade**

Senior Product Manager  
Google Cloud

## Special thanks

**Boteng Yao**

Senior Software Engineer  
Google Cloud

**Yubin Gong**

Senior Software Engineer  
Google Cloud

Scan, register and  
stop by our booth  
for Google swag

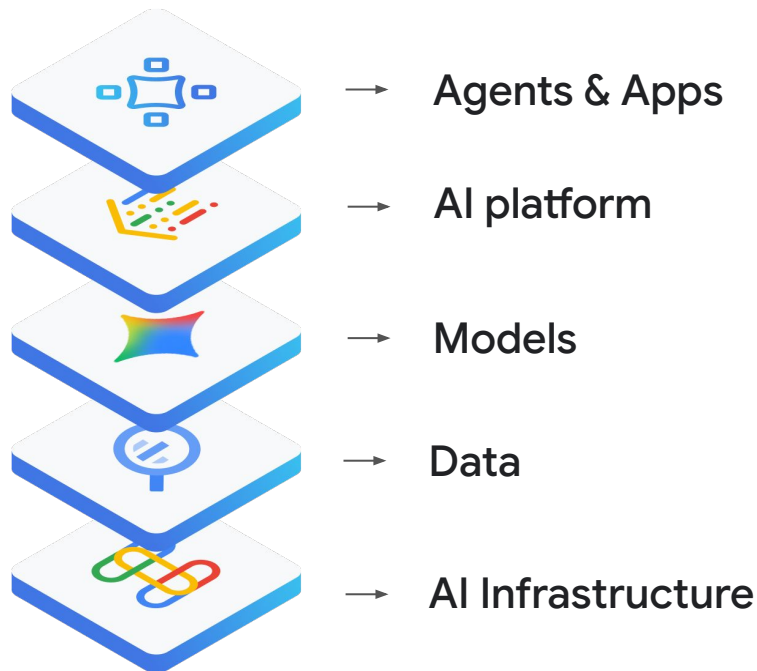


“Does Google use MCP”

# Google Cloud's AI-optimized stack



Secure, enterprise-ready foundation



# Vertex AI Agent Builder

The best platform to build, scale, optimize & govern custom agents

## Build

Open-source toolkit for building, evaluating, and deploying

## Scale

Deploy and manage agents in production with managed services

## Govern

Govern agent sprawl; manage agent risk and compliance

## Optimize

Measure accuracy, visualize logic, and ensure enterprise-grade quality and safety throughout the agent lifecycle

Models/Gemini API

Tools, data, and other agents

# Vertex AI Agent Builder

The best platform to build, scale, optimize & govern custom agents

## Build

Agent Development Kit

Agent Designer

Agent Garden

Marketplace

## Scale

Runtime

Sessions

Code exec

Example store

Memory Bank

Computer use

UI to manage agents

Framework agnostic

Provisioned Throughput

## Govern

Identity

Registry

Model Armor

Secure Web Proxy

Security command center

Compliance

VCP-SC, CMEK, DRZ, AXT, HIPAA

Policy & Access Controls

Govern A2A

Govern MCP

## Optimize

Evaluations

Traces

Observability

Simulation

Optimization

## Models/Gemini API

**Gemini Models**  
Optimized for agentic reasoning

**Model agnostic**  
Choose the model for your needs

**Model Garden**  
Hundreds of curated LLMs

**Audio/Video streaming**

## Tools, data, and other agents

A2A

Search grounding

Ecosystem tools

MCP

RAG

Open platforms

Enterprise context

Data for agents

100s of connectors

# Vertex AI Agent Builder

The best platform to build, scale, optimize & govern custom agents

MCP is *actually* an integral part of our platform...

## Build

- Agent Development Kit
- Agent Designer
- Agent Garden
- Marketplace

## Scale

- Runtime
- Sessions
- Code exec
- Example store
- Memory Bank
- Computer use
- UI to manage agents
- Framework agnostic
- Provisioned Throughput

## Govern

- Identity
- Registry
- Model Armor
- Secure Web Proxy
- Security command center
- Compliance  
VCP-SC, CMEK, DRZ, AXT, HIPAA
- Policy & Access Controls
- Govern A2A
- Govern MCP**

## Optimize

- Evaluations
- Traces
- Observability
- Simulation
- Optimization

## Models/Gemini API

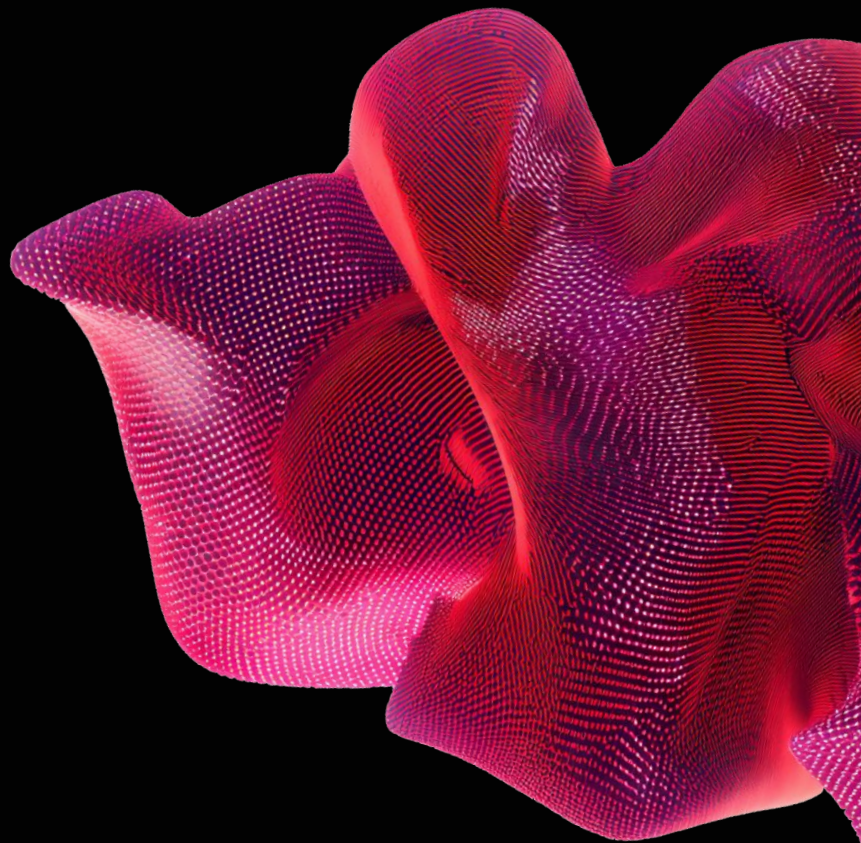
- Gemini Models  
Optimized for agentic reasoning
- Model agnostic  
Choose the model for your needs
- Model Garden  
Hundreds of curated LLMs
- Audio/Video streaming

## Tools, data, and other agents

- A2A
- Search grounding
- Ecosystem tools
- MCP**
- RAG
- Open platforms
- Enterprise context
- Data for agents
- 100s of connectors

01

# Build



# Kaggle 5 day intensive course

## Learn more

Every ~6 months we run a “5 day intensive course”, and we dedicated the most recent one entirely to Agent building & Agent Ops.

1.5M learners

2+M whitepaper views

3+M notebook views

[www.kaggle.com/learn-guide/5-day-agents](https://www.kaggle.com/learn-guide/5-day-agents)

## What is the 5-Day AI Agents Intensive?

The 5-Day AI Agents Intensive course with Google is a hands-on program originally held live from November 10 - 14, 2025. It is now available as a self-paced Kaggle Learn guide so anyone can explore the foundations, architecture and practical development of AI agents.

This course was crafted by Google's ML researchers and engineers to help developers explore the foundations and practical applications of AI agents. You'll learn the core components – models, tools, orchestration, memory and evaluation. Finally, you'll discover how agents move beyond LLM prototypes to become production-ready systems.

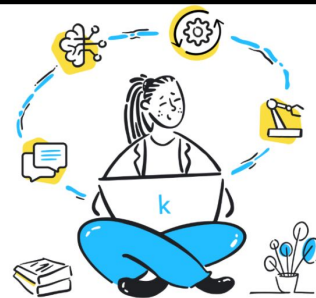
Each day blends conceptual deep dives with hands-on examples, code labs, and live discussions. By the end, you'll be ready to build, evaluate, and deploy agents that solve real-world problems.

## Welcome to our Agents Intensive Learn Guide!



### What's being covered?

- **Day 1 - Introduction to Agents:** Explore the foundational concepts of AI agents, their defining characteristics, and how agentic architectures differ from traditional LLM applications, laying the groundwork for building intelligent, autonomous systems.
- **Day 2 - Agent Tools & Interoperability with Model Context Protocol (MCP):** Dive into the world of tools, understanding how AI agents can "take action" by leveraging external functionalities and APIs, and explore the ease of discovering and using tools offered by the MCP.
- **Day 3 - Context Engineering: Sessions & Memory:** Explore how to build AI agents that can remember past interactions and maintain context. Learn how to implement short-term and long-term memory to create more robust agents capable of handling complex, multi-turn tasks.
- **Day 4 - Agent Quality:** Learn to build robust and reliable AI agents by mastering the critical disciplines of evaluating and improving agents. This session will cover observability, logging, and tracing to provide visibility, alongside key metrics and evaluation strategies to optimize agent performance.
- **Day 5 - Prototype to Production:** Go beyond local testing and learn to deploy and scale AI agents for real-world use. This session will cover the best practices for deploying your agents so that others can use them, including how to create a truly multi-agent system with the Agent2Agent (A2A) Protocol.



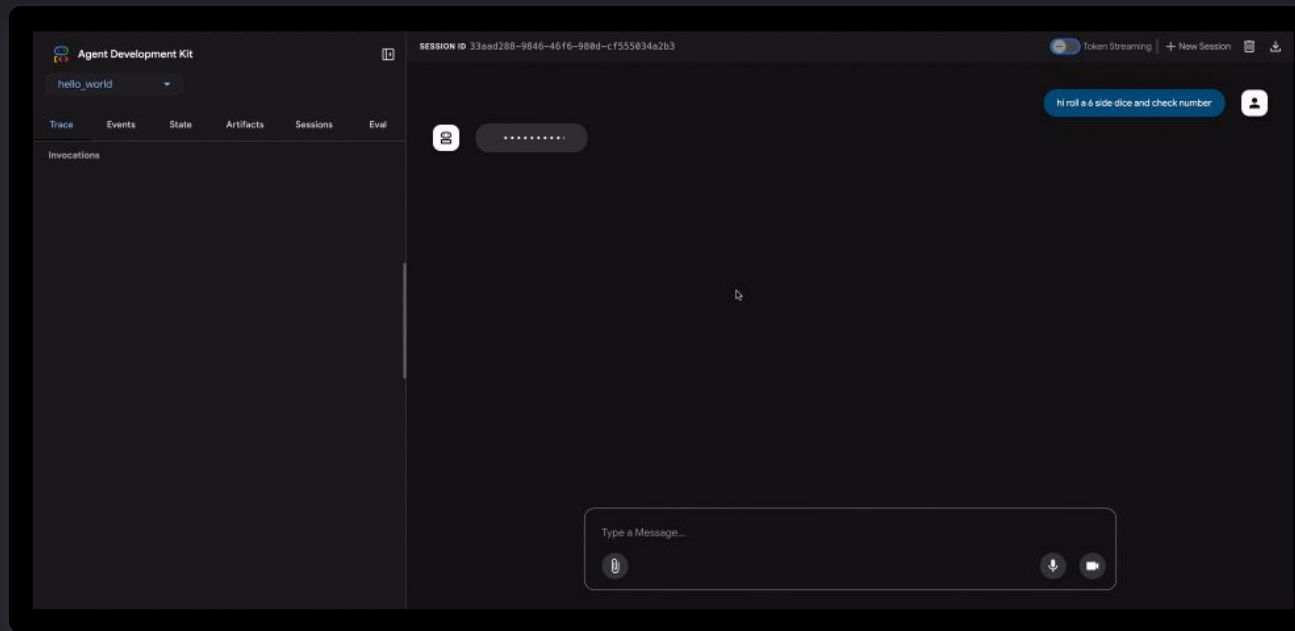
# ADK

## Agent Development Kit

Google's ADK is the agent framework Google uses for our own products; it's fully open source and extensible.

Equip agents with resources and actions via MCP and Skills.

[adk.dev](https://adk.dev)

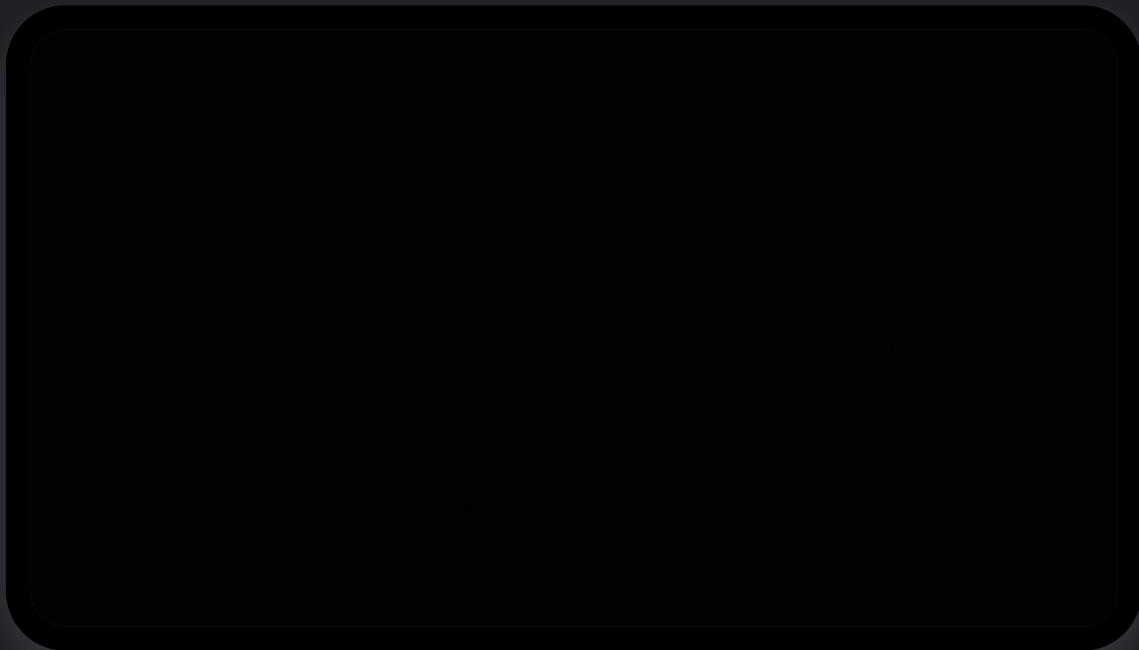


# Antigravity

Google Antigravity is our agentic development platform, evolving the IDE into the agent-first era.

Bring your MCP servers to your IDE.

[antigravity.google](https://antigravity.google)

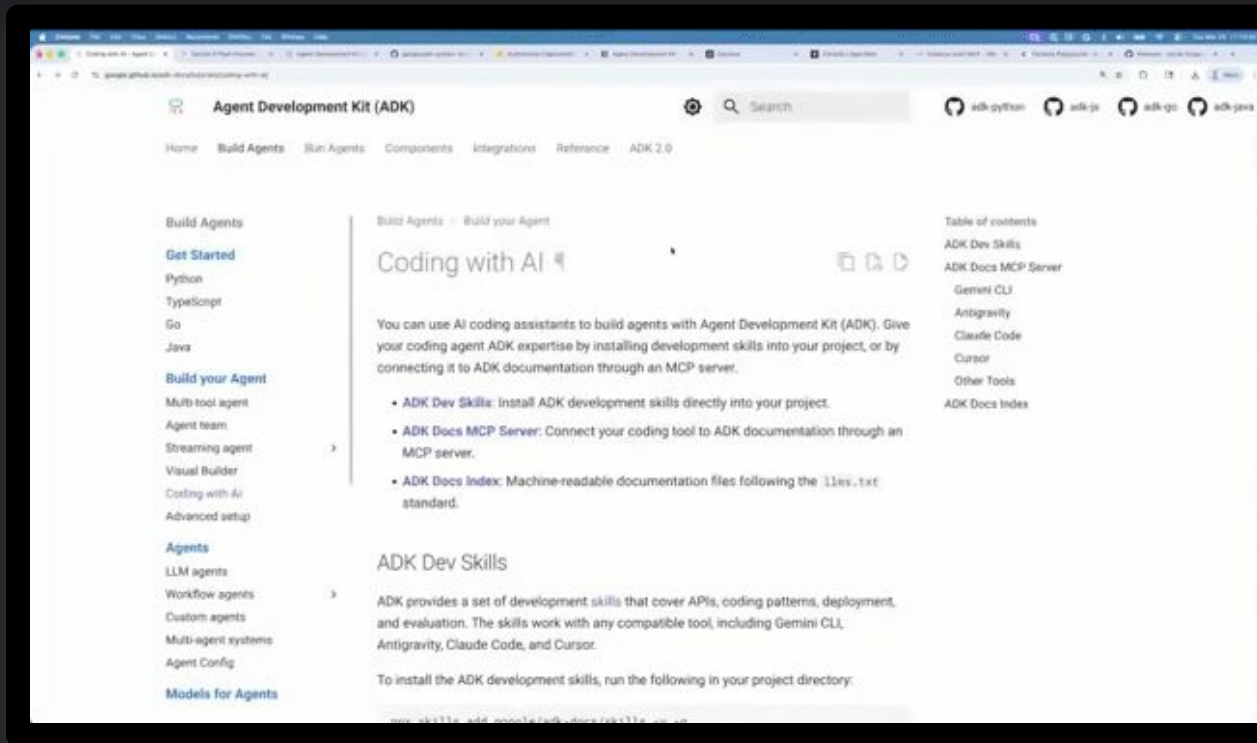


# ADK Dev Skills & Docs MCP

Devs build agent skills wired into Antigravity or Gemini CLI.

Bring your MCP servers to your IDE.

[adk.dev](https://adk.dev) → Coding with AI

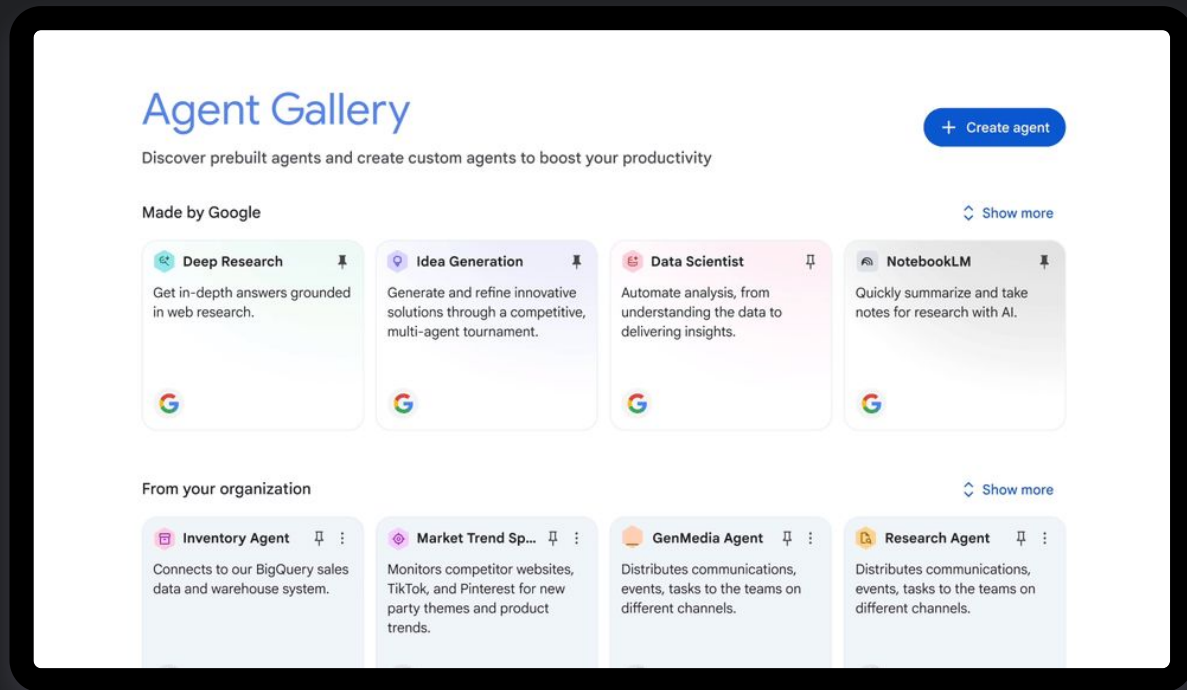


# Gemini Enterprise

Gemini Enterprise is an Employee Productivity application built on top of this platform.

Bring and connect data and agents, and build agents, with MCP support.

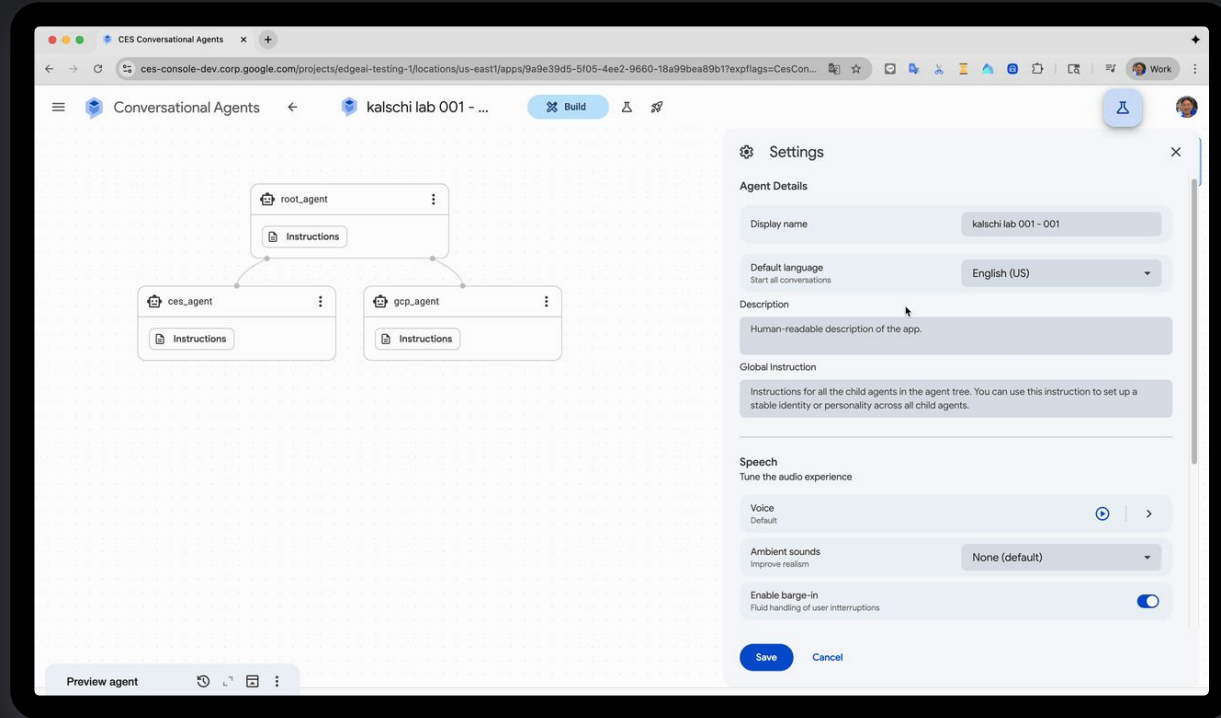
[cloud.google.com/gemini-enterprise](https://cloud.google.com/gemini-enterprise)



# Gemini Enterprise for Customer Experience

Gemini Enterprise for Customer Experience is an application for the full CX agent lifecycle. Discover OOTB agents, customize them, build and optimize them - all built on top of our platform with MCP support.

[gecx.cloud.google.com](https://gecx.cloud.google.com)



# MCP Toolbox for Databases

## Open Source & Extensible

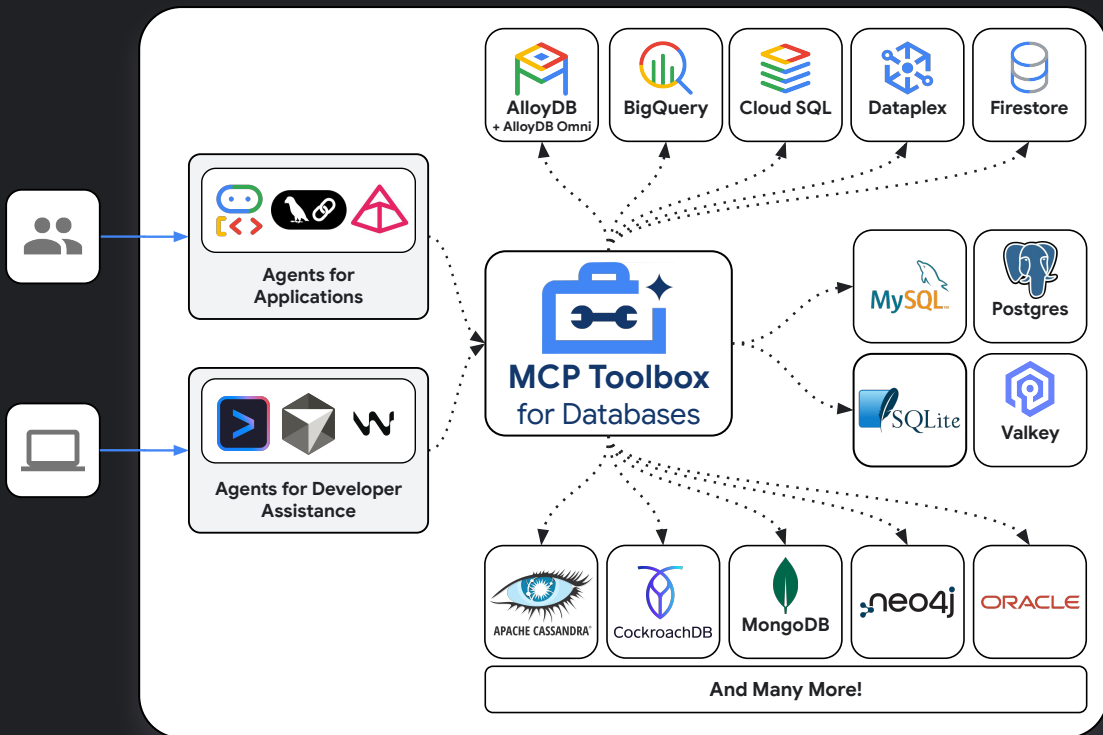
- ★ 13k+ Stars on GitHub, +127 contributors
- 🌐 40+ different databases

## Highly Customizable

- Pre-built tools for common operations
- Configure specific, narrow access for production-grade applications.

## Scalable and Secure

- Secure connection management & credential handling
- Built-in connection pooling and OpenTelemetry support

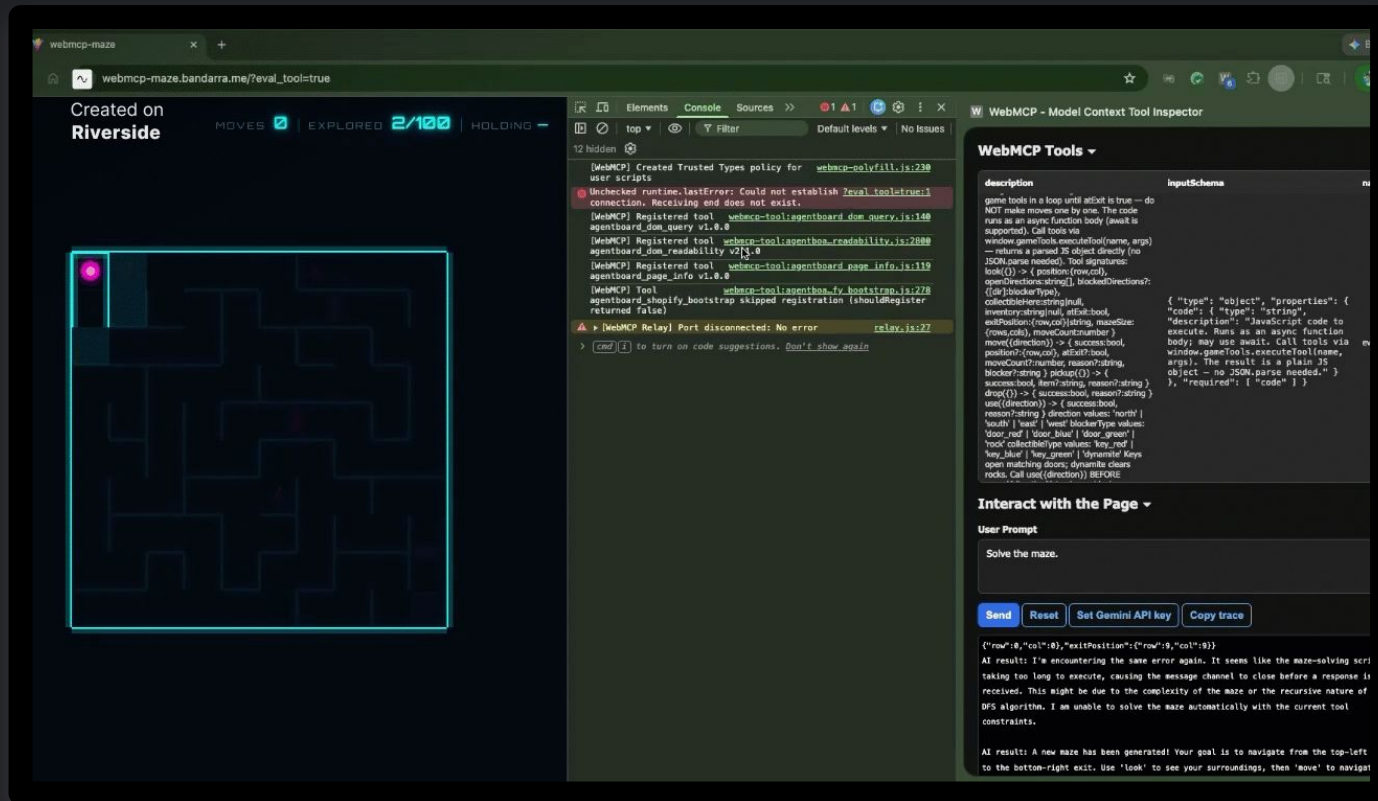




# WebMCP

WebMCP bridges the gap between AI models and the web by enabling agents to interact with browsers and APIs using the Model Context Protocol (MCP).

Developers can build "web-aware" AI applications that autonomously navigate, extract data, and perform complex tasks across any website.



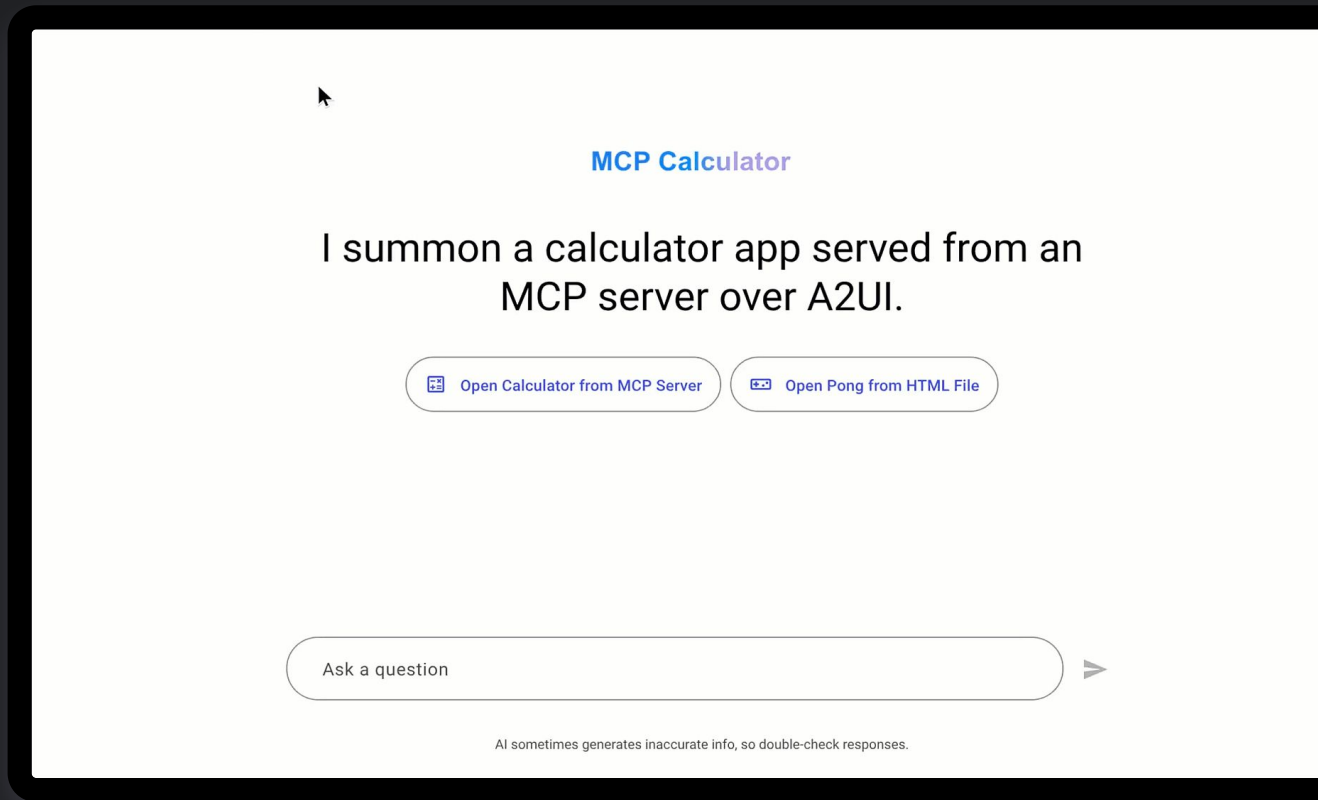
# A2UI + MCP (apps and transport)

This is a declarative, generative UI protocol & toolkit with renderers for web and mobile and more.

Safe like data, expressive like code.

Now working over MCP Apps and MCP transport (*experimental*)

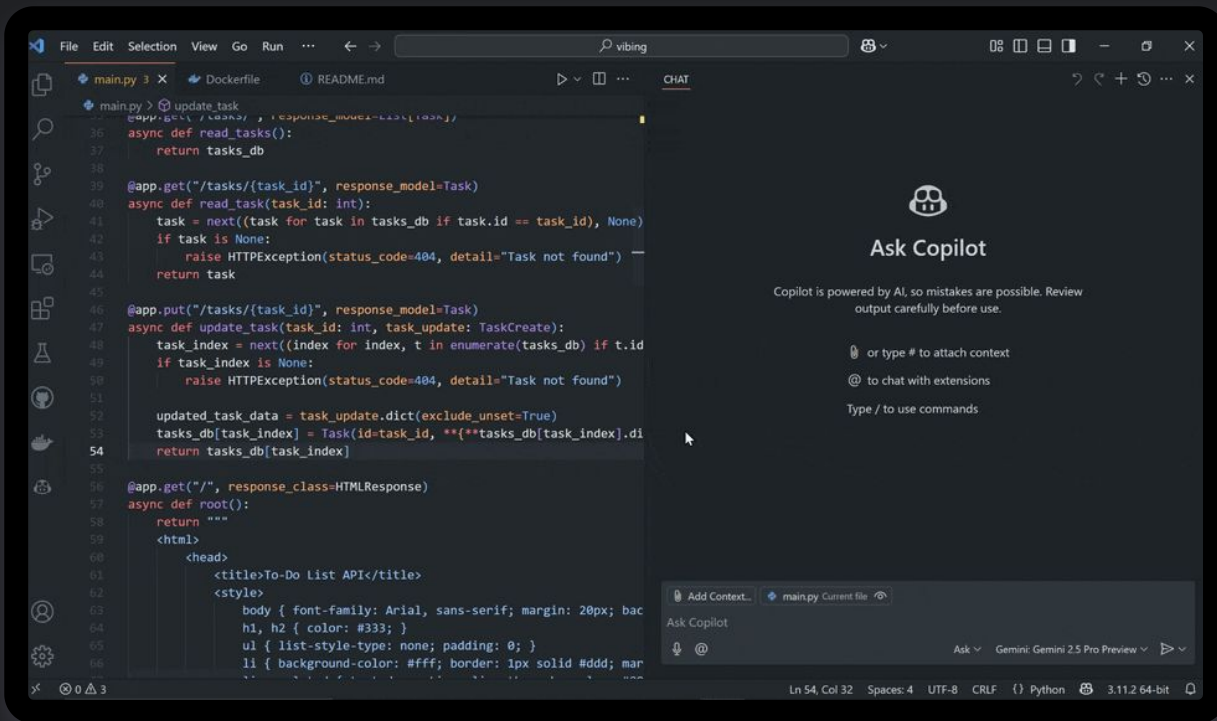
[a2ui.org](https://a2ui.org)



# Cloud Run MCP

Many Google APIs are exposed as hosted MCP servers.

For example here is Copilot using our Cloud Run MCP server to deploy an application.



# Deploy your own MCP servers

Build an MCP server. Deploy to:



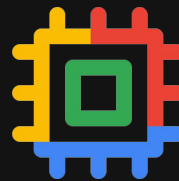
**Cloud Run**

[cloud.run](https://cloud.run)



**Google Kubernetes  
Engine (GKE)**

[cloud.google.com/gke](https://cloud.google.com/gke)



**Compute Engine**

[cloud.google.com/compute](https://cloud.google.com/compute)




02

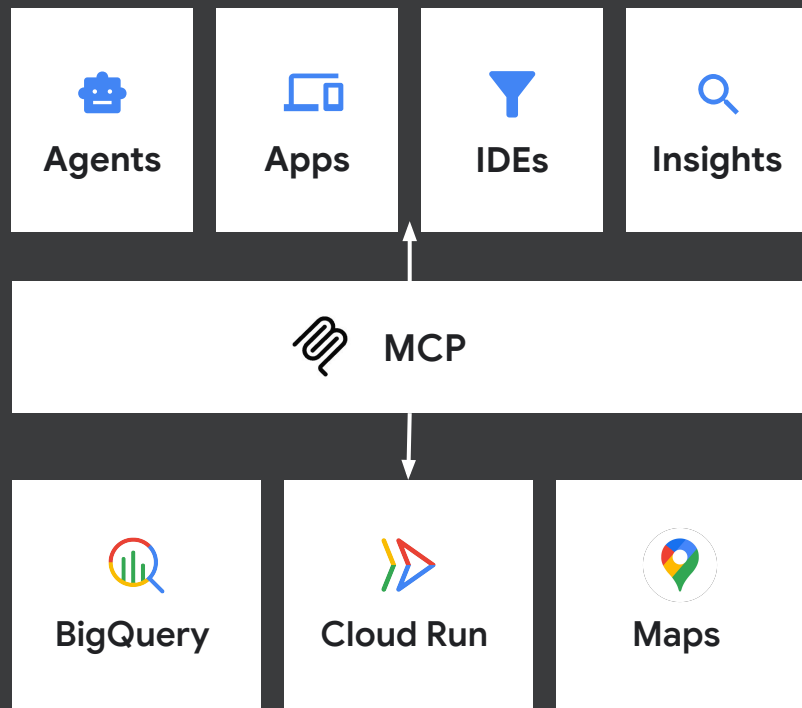
# Scale







# Google MCP Servers

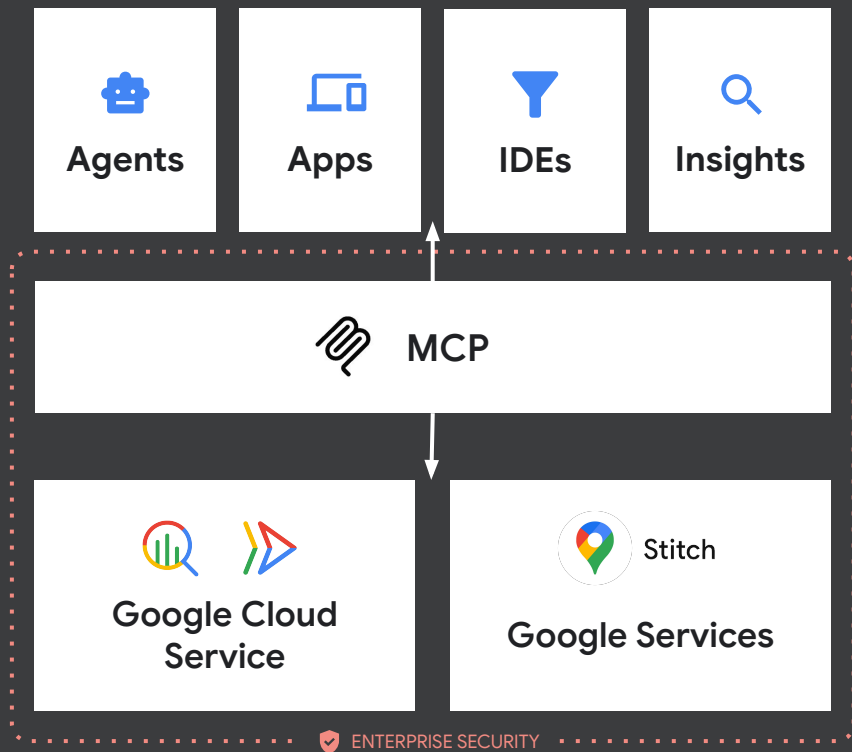
The unified, fully managed MCP platform for Google Cloud

-  Unified interface and simplified discoverability with Registry
-  Fully managed: No infrastructure to provision. Google handles hosting, scaling, and security.
-  Easy access your favorite client: Gemini CLI, Claude, ADK Agents, and more...



# Built-in Enterprise Security for Google MCP Servers

-  OAuth Authentication
-  Fine-Grained Access & Organization Policies
-  AI Security: Detect & Block Prompt Injections
-  Audit Logs



# Calling Google MCP servers from ADK

You can integrate Google MCP Servers via the StreamableHTTP class in the Agent Development Kit.

```
agents > cityscape > agent.py > ...
1  from google.adk.agents import LlmAgent, ParallelAgent, SequentialAgent
2  from google.adk.tools.mcp_tool import McpToolset, StreamableHTTPConnectionParams, StdioConnectionParams
3  from mcp import StdioServerParameters
4  from google.adk.tools import google_search
5  from google.adk.tools.tool_context import ToolContext
6
7  import datetime
8  from google.genai import types
9  import os
10
11  DEFAULT_MODEL='gemini-2.5-flash'
12  NANO_BANANA_MODEL='gemini-3-pro-image-preview'
13
14  get_weather = McpToolset(
15      connection_params=StreamableHTTPConnectionParams(
16          url="https://mapstools.googleapis.com/mcp",
17          headers={"X-Goog-API-Key": os.environ["MAPS_API_KEY"]}
18      ),
19  )
20
```

Google Maps MCP server



General

Account

Privacy

Billing

Usage

Capabilities

**Connections**

Claude Code



GitHub

https://github.com

Connected



Slack Calendar

https://calendar.slack.com

Connected



On Fisk

https://onfisk.com

Connected



LinkedIn connect

https://www.linkedin.com/company/anthropic

Configure



LinkedIn CP connect

https://www.linkedin.com/company/anthropic

Connect



LinkedIn connect

https://www.linkedin.com/company/anthropic

Connect

[Add custom connector](#)



# Apigee

## No-code composition of APIs as MCP Tools

- ✓ Convert existing REST APIs into MCP servers with No code
- ✓ Compose a new MCP server from multiple REST APIs with no code
- ✓ Generate Agent-optimized specs using SpecBoost
- ✓ Continuous feature enhancement to support evolving MCP specification including Authz Info, Tool Set etc

[cloud.google.com/apigee](https://cloud.google.com/apigee)

# Optimally Serve & Govern

## Runtime policies optimized for MCP & beyond

- Enable robust Security Policies like OAUTH Apigee on MCP Server
- Leverage Semantic Caching Policies to quickly serve very chatty agentic flows
- Protect you backend tools from agentic abuse via Quota or Rate Limiting Policies
- Additionally you can use Apigee's 30+ built-in security and governance policies for MCP tool guardrails
- Native ability to differentiate between different JSON-RPC protocols: e.g. `tools/list` and `tools/call` operations for MCP.
- All Apigee policies still apply for MCP/A2A/AP2



**AVAILABLE NOW!**

Example: List Tools

```
{
  "jsonrpc": "2.0", "id": 1,
  "method": "tools/list",
  "params": { "cursor": "..."}
}
```

Example: Call Weather Tool

```
{
  "jsonrpc": "2.0", "id": 2,
  "method": "tools/call",
  "params": { "name": "get_weather", ... }
}
```

*MCP servers created from APIs in customer's catalog*

03

# Govern



# Registry

## Discover, share, and scale agent capabilities

- ✓ Tool registry: centralized discovery/reuse of capabilities
- ✓ Expanding ecosystem of ready-to-use Google/3P tools
- ✓ Agent Marketplace for 3P A2A agents and MCP tools

The screenshot shows the Google Cloud Vertex AI Tools interface. The left sidebar contains navigation options like Colab Enterprise, Workbench, Vertex AI Studio, Overview, Multimodal, Vision, Translation, Speech, Prompt Gallery, Prompt management, GenAI Evaluation, Tuning, Agent Builder, Agent Garden, Agent Engine, Tools (highlighted), RAG Engine, Vertex AI Search, Vector Search, Tutorials, Migrate to Vertex AI, and Marketplace. The main content area is titled 'Tools' and includes a 'Manage organization policy' link. Below this is a table of tools with the following columns: Name, Description, Type, Server Source, Org policy, and MCP status. The table lists 11 tools, including Acme API, Acme access entry API, Course-registry API, SAP-key-verify API, SAP-KVN-service-API, Inventory-function, Registry, Mock-server, Messaging service, and Translate. The 'Org policy' column shows 'Allowed' for all tools, and the 'MCP status' column shows 'Enabled' for Acme API, Course-registry API, and Translate, while others are 'Disabled'. A filter 'Org policy : Allowed' is applied. The bottom right of the table shows 'Rows per page: 20' and '1-10 of 241'.

Name	Description	Type	Server Source	Org policy	MCP status
▶ Acme API	Tools for querying and managing large datasets in BigQuery	MCP_SERVER	Apigee	✓ Allowed	✓ Enabled
▶ Acme access entry API	Tools for understanding text using Google Cloud Natural Language AI	MCP_SERVER	Apigee	✓ Allowed	⊖ Disabled
▶ Course-registry API	A toolset for managing relational databases on Google Cloud	MCP_SERVER	Apigee	✓ Allowed	✓ Enabled
▶ SAP-key-verify API	Tools for storing and retrieving objects in Google Cloud Storage	MCP_SERVER	Apigee	✓ Allowed	⊖ Disabled
▶ SAP-KVN-service-API	Tools for analyzing images with Google Cloud Vision AI	MCP_SERVER	Apigee	✓ Allowed	⊖ Disabled
▶ Inventory-function	This Cloud Run service hosts a high-performance API	MCP_SERVER	Cloud Run	✓ Allowed	⊖ Disabled
▶ Registry	A set of tools for managing registry	MCP_SERVER	GCE	✓ Allowed	⊖ Disabled
▶ Mock-server	Tools for interacting with mock server	MCP_SERVER	GKE	✓ Allowed	⊖ Disabled
▶ Messaging service	A set of tools for accessing messaging Platform APIs	MCP_SERVER	Android Manag..	✓ Allowed	⊖ Disabled
▶ Translate	Tools for translating text between languages	MCP_SERVER	Compliance Ma..	✓ Allowed	⊖ Disabled

# Core requirements for Agent Governance

01

## Authentication and Authorization

Who is the user & agent? What is it allowed to do?  
Which actions need human authorization?

02

## Auditing

You can't secure what you can't see.

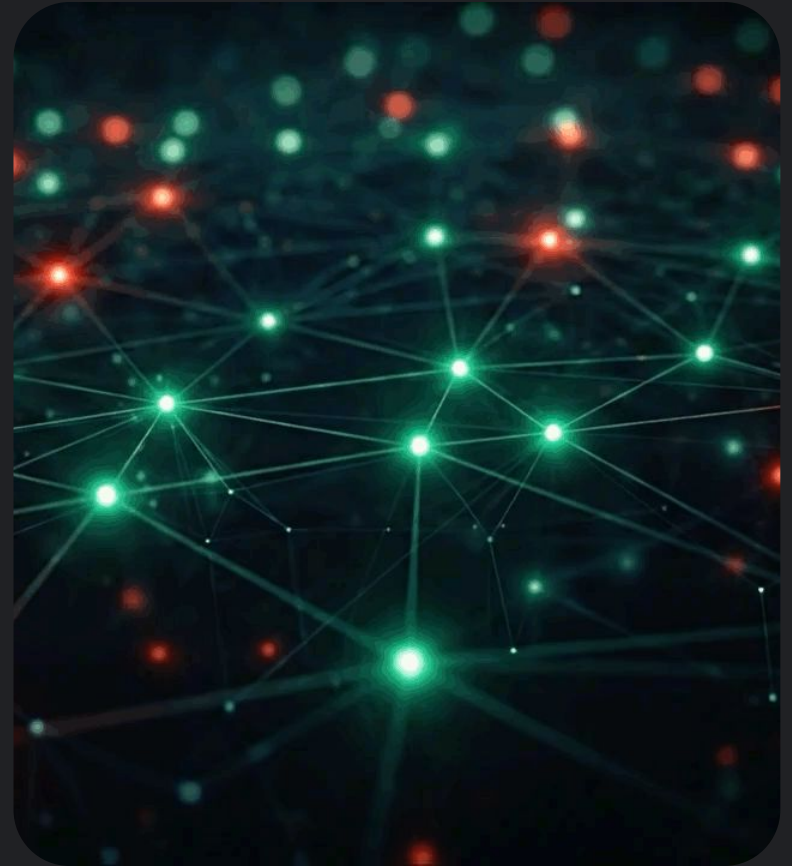
03

## Agent Data & Context security

Sanitizing what agents see and say.

# The Governance Gap in Agentic AI

Feature	Traditional Proxy
Route by URL	✓
Rate-Limit by IP	✓
Quota by Model Name	✗
Access control by Tool	✗



# Envoy Proxy: Agentic AI ready



## Open & Extensible

- ✓ Protocol parsing:  
OpenAI API, MCP, A2A
- ✓ Extensible policies & orchestration
- ✓ Open-source



## Pervasive

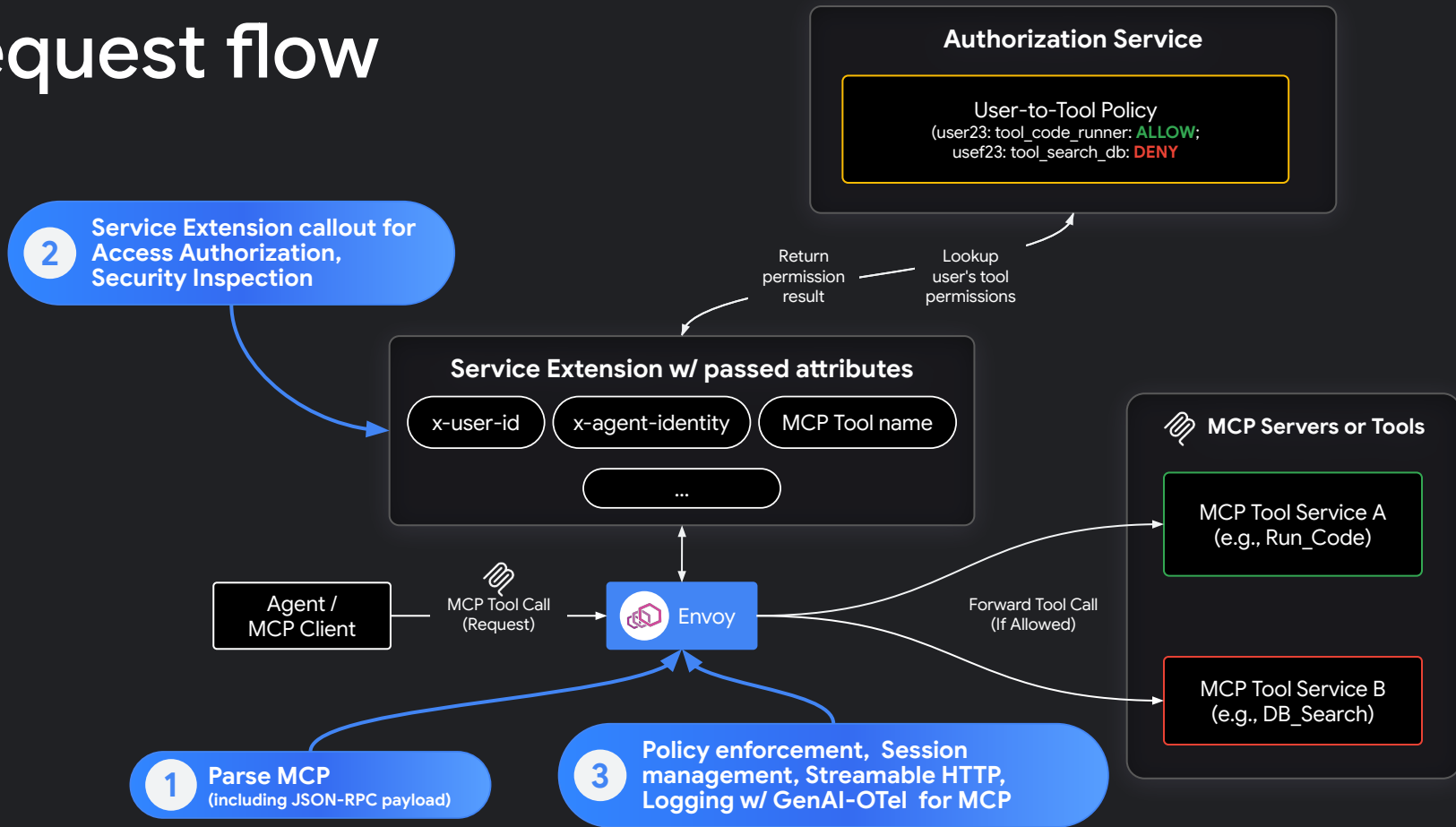
- ✓ Cloud & Kubernetes
- ✓ On-Premises
- ✓ On devices



## Scalable & Resilient

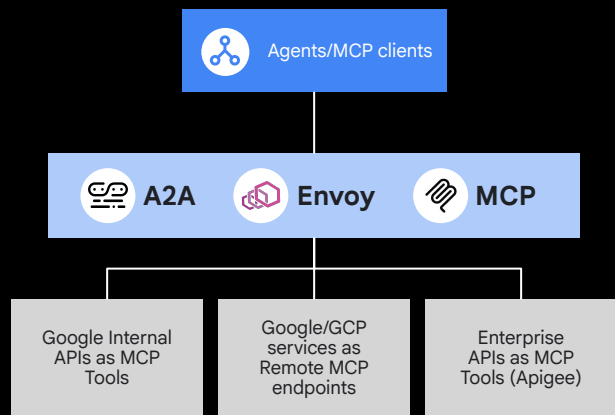
- ✓ Horizontally scalable
- ✓ Zero-downtime updates
- ✓ Low latency & high throughput

# Request flow

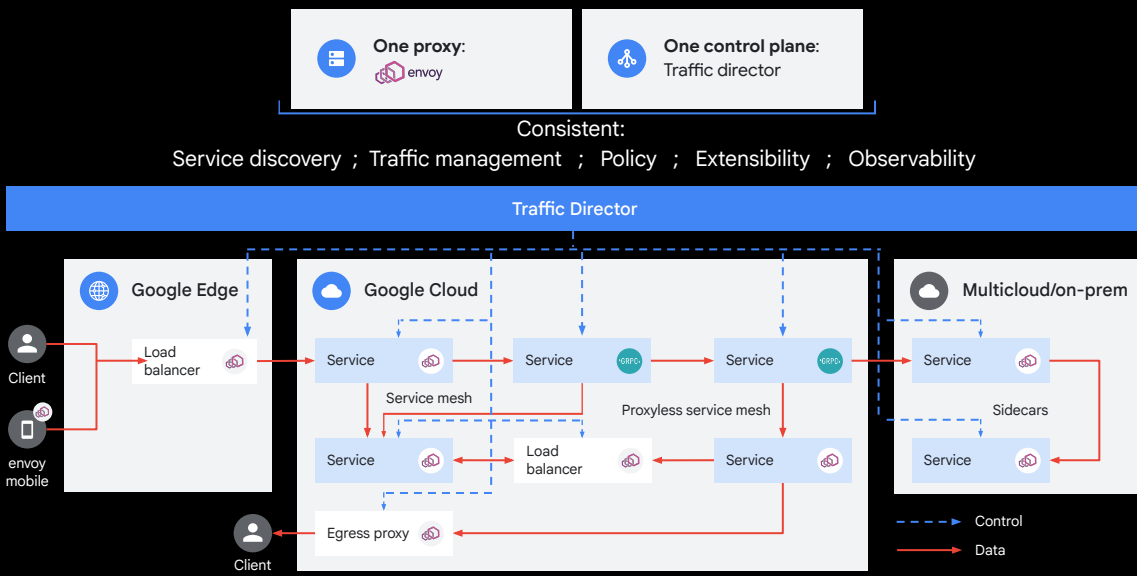


# Envoy @ Google for Agentic AI governance

## Securely publish MCP Tools as Remote endpoints



## Service connectivity w/ secure MCP (Svc-Svc, Svc-Internet)



# Thank you

---

Scan, register and  
stop by our booth  
for Google swag

