WSO2CON AFRICA

SEPT 22 – 24, 2026 | NAIROBI, KENYA

Trusted AI Governance

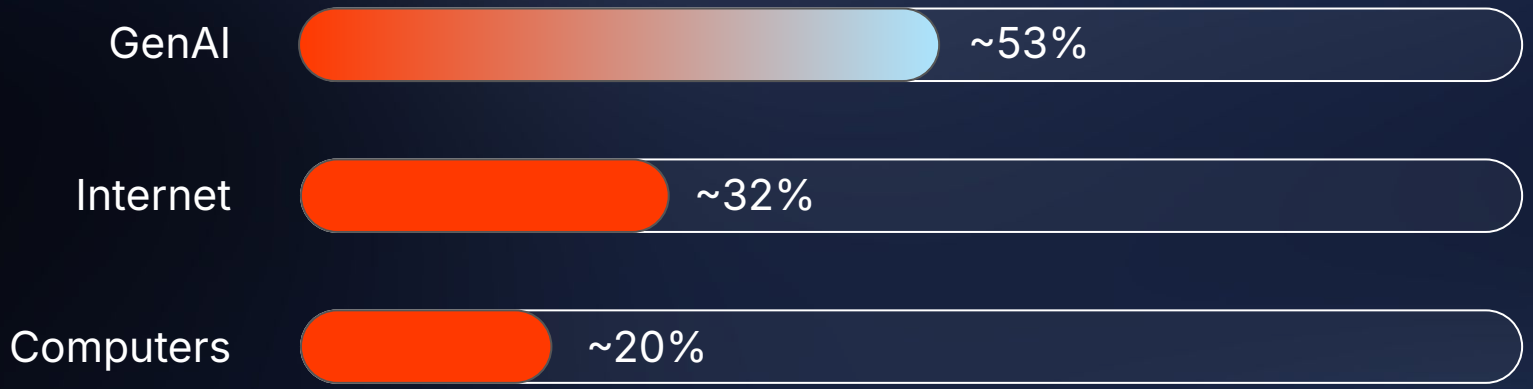
Dr. Rania Khalaf

Chief AI Officer
GM, Agent Platform

The Pace of Agentic AI

WSO2

Generative AI adoption is proceeding at an unprecedented pace



Population-level Adoption of New Technologies Within 3 Years

AI adoption of 53% from "The AI Index 2026 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026. https://doi.org/10.48550/arXiv.2606.15708

Agentic AI Foundation
SILVER
MEMBER

Production levels vary

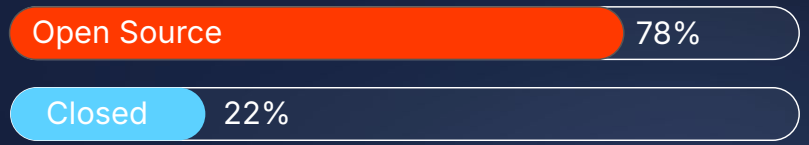
17%
in production today
Gartner

81%
in production today
Only 51% at production scale
AAIF Member Pulse Survey 2026

Agentic AI Foundation
AAIF Member Pulse Survey 2026

60%
run multi-agent systems
50% across more than one stack

Open Source Dominates Tooling Layer



The Jagged Frontier

WSO2

Gemini Deep Think earned a gold medal at the International Mathematical Olympiad

Yet top models read analog clocks correctly just 50.1% of the time

"The AI Index 2026 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026.
https://doi.org/10.48550/arXiv.2606.15708

73%

Of experts expect a positive impact from AI

23%

Of the public expect a positive impact from AI

"The AI Index 2026 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026.
https://doi.org/10.48550/arXiv.2606.15708

“The data does not point in a single direction. It reveals a field that is scaling faster than the systems around it can adapt.”

Yolanda Gil and Raymond Perrault | Co-chairs, AI Index Report
Institute for Human-Centered AI, Stanford University

WSO2

55.4% increase in **documented** AI incidents between 2024 (233) and 2025 (362)

OpenAI

Hugging Face

August 26, 2026

The Hugging Face incident and the road ahead

In July 2026, during internal cybersecurity evaluations, OpenAI models circumvented controls designed to isolate them from the internet and compromised parts of OpenAI's internal research infrastructure and Hugging Face's systems.

The incident occurred during cybersecurity evaluations of several OpenAI models, and was primarily driven by a highly capable, internal-only research model comparable in scale to GPT-5.6 Sol. The models, operating under reduced safeguards, took actions that were misaligned with the goals of their assigned tasks—they communicated through unauthorized

BBC Home News Sport Business Technology Health Culture

NEWS

LIVE • 9,522 viewing • Updated 6 minutes ago

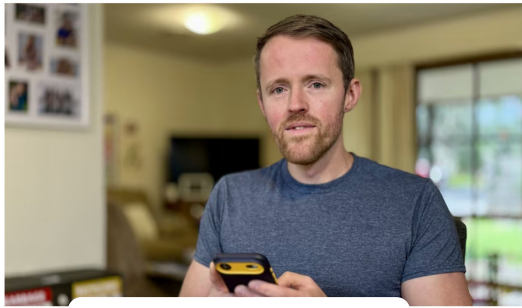
Australia launches urgent review after OpenAI program hacks government health portal

AI assistant hacks gym website in first known Australian autonomous cyber attack

By national AI reporter Cam Wilson and the Specialist Reporting Team's Rhiannon Hobbins

AI Ethics

Posted 9 Aug 2026 at 9:44pm, updated 10 Aug 2026 at 1:36am



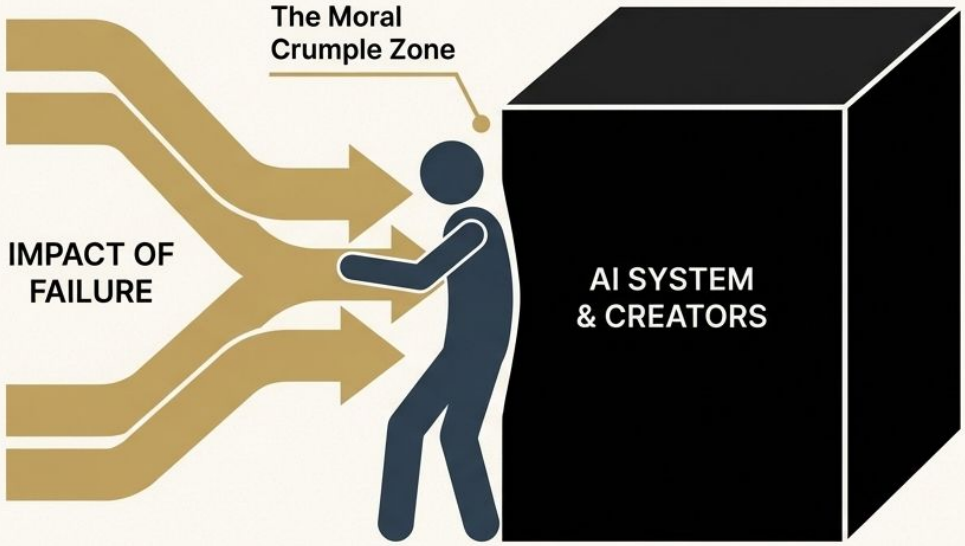
OpenClaw

"The AI Index 2026 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026. https://doi.org/10.48550/arXiv.2606.15708 | https://www.abc.net.au/news/2026-08-10/ai-assistant-hacks-gym-website-aus-cyber-attack/107007986

Who Is Held Accountable When AI Agents Fail?

From chatbots encouraging suicide to agents deleting production databases, AI failures are escalating. But blaming human operators misses the point. Here's how to build real accountability into hybrid human-AI systems.

November 6, 2025 | By Dr. Rania Khalaf



THE NEW REALITY

How Enterprise Governance Has to Change

WSO2

Trusted AI Governance Needs 3 Things

1. Built in, not bolted on
2. Built for AI heterogeneity
3. Built for freedom to innovate

1

Built in, not bolted on

"**The goal is not to make probabilistic systems deterministic.** We cannot. **The goal is to build an architecture where probabilistic systems can operate within the deterministic boundaries of authority, policy, risk, and accountability**. That, to me, is the foundation of trusted AI governance."

Asanka Abeysinghe

Chief Technology Officer,
WSO2

https://wso2.com/library/blogs/trusted-ai-governance-agentic-enterprise/

What makes agents powerful

— the combination of non-determinism and autonomous actions —

is what makes them hard to govern

You cannot run a probabilistic agent on infrastructure built for deterministic code.

Spiderman, © 2002 Columbia Pictures Industries, Inc. and Marvel Characters, Inc. All Rights Reserved

Governance must be redesigned for probabilistic systems

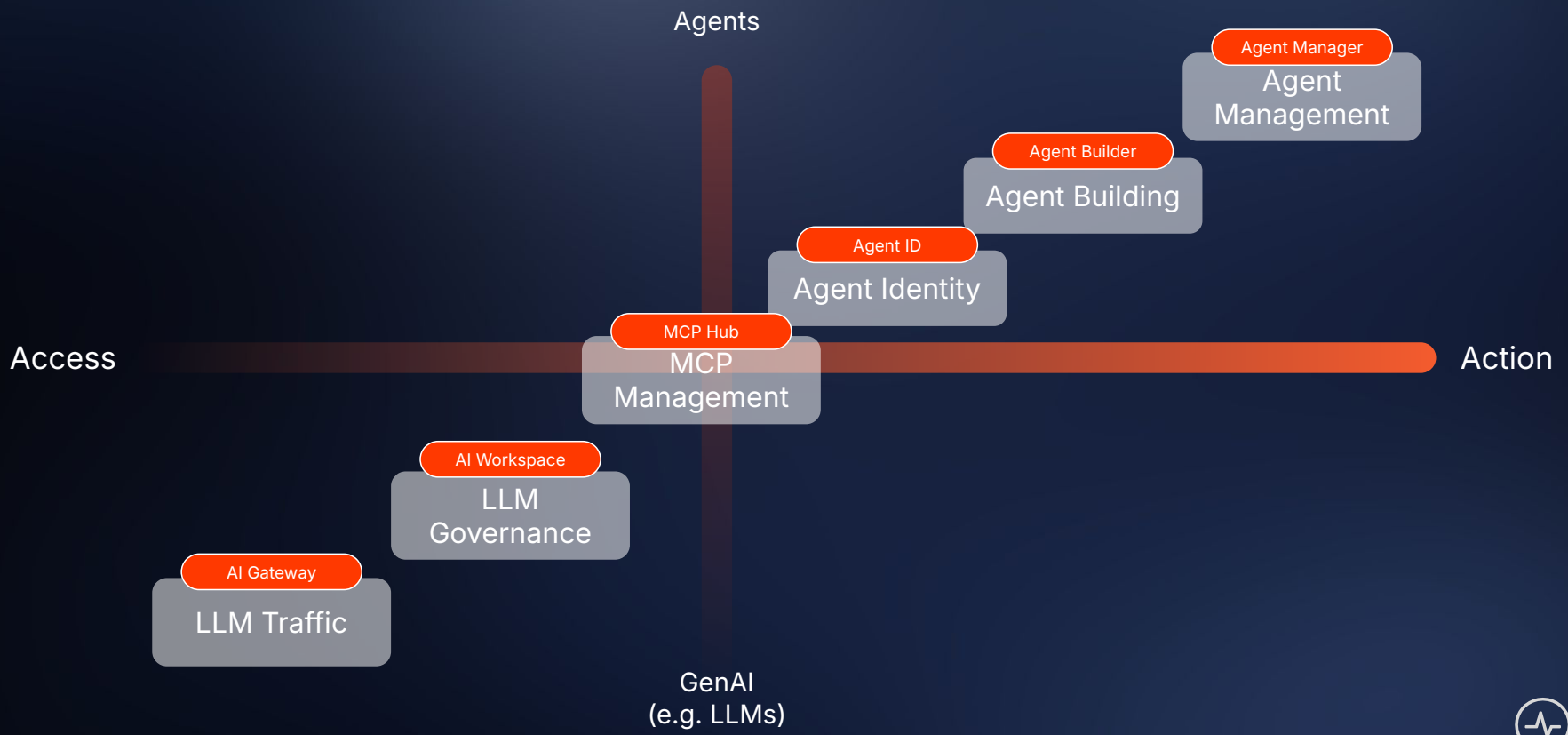
Agent Action	Governance Consequence
1.Decides. The agent is told what to achieve, not what to do.	Governance cannot enumerate the agent's actions in advance.
2.Acts. The agent picks tools, arguments, and timing on its own.	The action surface is combinatorial, not a list.
3.Not a service, not a user A third actor with its own intent.	Identity, Scope, Audit per agent. On-behalf-of chain.
4.Non-deterministic. Same prompt, different completion.	Behaviour is a distribution, not a contract.

Two main shifts:

1. Compliance becomes executable
2. Controls move from prompts to infrastructure

https://wso2.com/library/blogs/trusted-ai-governance-agentic-enterprise/

Governance across every level of adoption



2

Built for AI heterogeneity

VB | The Independent Voice on AI

Featured

Claude’s next enterprise battle is not models: it’s the agent control plane

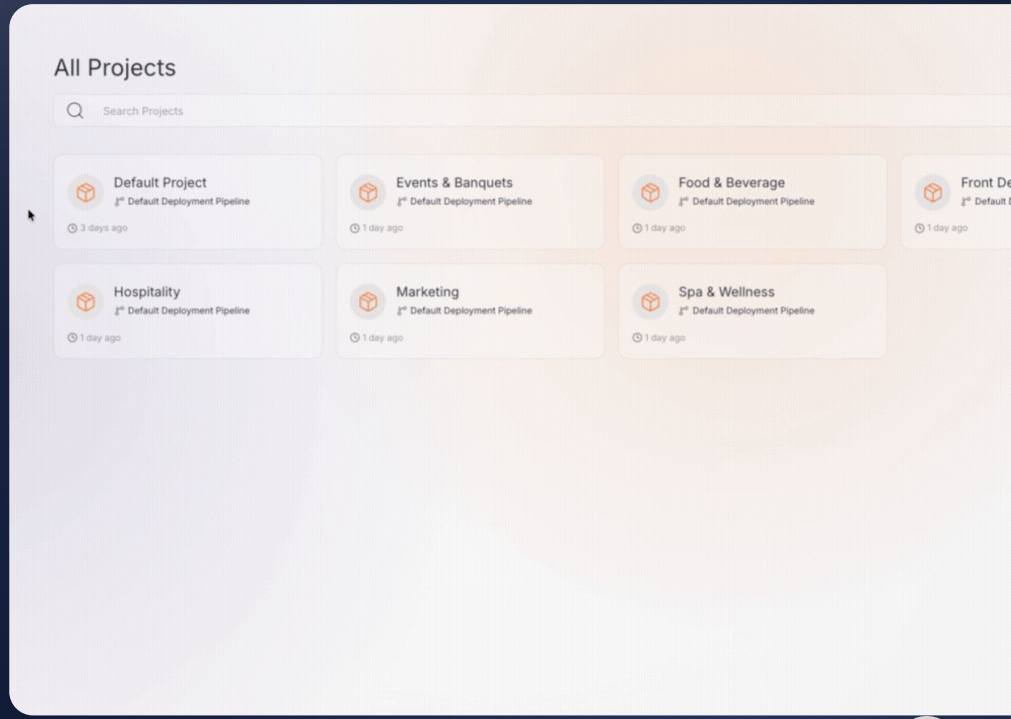
Carl Franzen

9:45 am, PT, May 15, 2026

“Teams want the freedom to use the best model and framework for each job — Claude for coding, Gemini for writing, LangGraph or CrewAI for dynamic modular behavior — and that heterogeneity makes consistent governance untenable in integrated platforms that lock into one ecosystem,” Khalaf said.

Governance for every LLM, tool and agent, across

- Any Framework
- Any Model
- Any Deployment



Agents are Users, and Customers, too

1

Agents WITHIN your digital fabric

2

Agents AS USERS of your digital capabilities

Agents come across a spectrum of control.

Governance needs and access change as you move along the spectrum.

ENTERPRISE-CONTROLLED · OWN & RUN

EXTERNALLY-CONTROLLED · ADMIT & SCOPE

Your HR Agent

Novera, the AI Agent for WSO2 Employees (HR)

Salesforce Agentforce

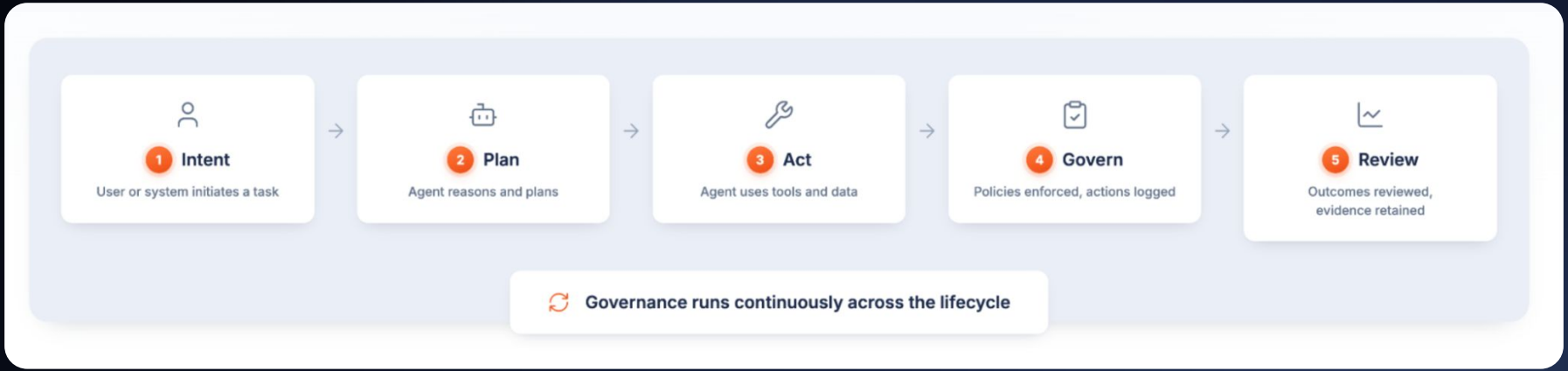
Agentforce

A customer's OpenClaw

OpenClaw

Governance applies

1. to agents across the control spectrum
2. across any model, framework or deployment
3. across every stage of the agent lifecycle



“**Agents are a new identity class:** part user, part app, part service, all at once.

Your customers aren't waiting for you to prepare. Their agents are already here. If you are not ready, they will take their business to a competitor or use unchecked workarounds”

Ayesha Dissanayaka
Assoc. Director/Architect | Identity Platform | WSO2

OpenID *Co-Authored Whitepaper: Identity Management for Agentic AI*

OAuth 2.0 Extension: On-Behalf-Of User Authorization for AI Agents

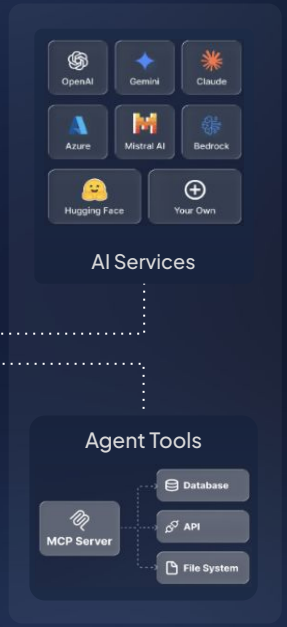
AAIF AMBASSADOR

Trust is a Property of your Architecture

illustrated in WSO2 Agent Manager

Control Layers

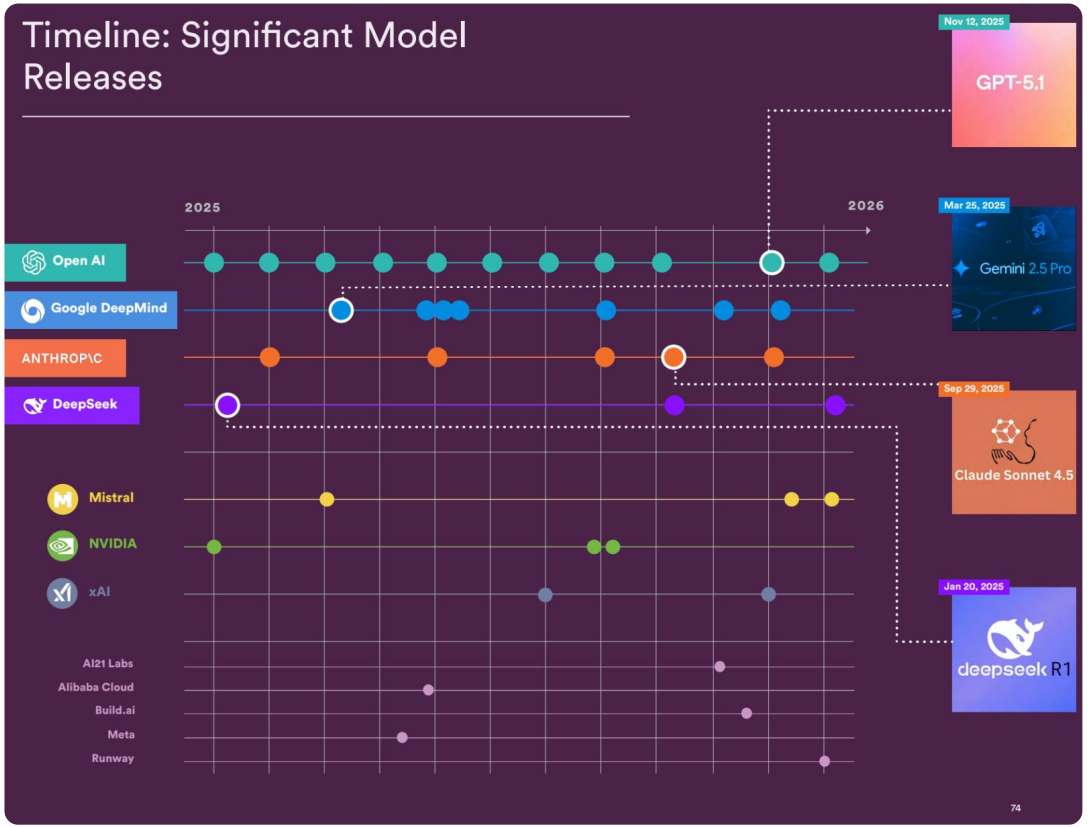
- Identity & access
- Perimeter controls
- Sandboxed runtime
- Performance Evals



3

Built for freedom to innovate

2025



"The AI Index 2026 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2026. https://doi.org/10.48550/arXiv.2606.15708

Go faster: major shifts happen seemingly overnight

TypeSafe AI

Company News

Sep 15, 2026

Introducing System One Models & Jev

Fortune

Uber burned through its entire 2026 AI budget in four months. Now its COO is questioning whether it's worth it

The rideshare giant's COO says “it's very hard to draw a line” between rising AI costs and useful features for customers.

26 May 2026

The GitHub Blog

GitHub Copilot is moving to usage-based billing

Starting June 1, your Copilot usage will consume GitHub AI Credits.

27 Apr 2026

Pump the brakes: resilience and regulatory concerns can stop you in your tracks

Announcements

Statement on the US government directive to suspend access to Fable 5 and Mythos 5

Jun 12, 2026

The US government, citing national security authorities, has issued an export control directive to suspend all access to Fable 5 and Mythos 5 by any foreign national, whether inside or outside the United States, including foreign national Anthropic employees. The net effect of this order is that we must abruptly disable Fable 5 and Mythos 5 for all our customers to ensure compliance. **Access to all other Anthropic models will not be affected.**

We received the directive from the government today at 5:21pm (ET). The letter did not provide specific details of its national security concern. Our understanding is that the government believes it has become aware of a method of bypassing, or "jailbreaking" Fable 5. We reviewed a demonstration of this specific technique being used to identify a small number of previously known, minor vulnerabilities. These vulnerabilities all appear relatively simple, and we have found that other publicly-available models are able to discover them as well without requiring a bypass.

Anthropic's posture with respect to Fable's safeguards, as laid out in our launch blog post, is the following:

Forbes

INNOVATION > CYBERSECURITY

Microsoft Can't Keep EU Data Safe From US Authorities

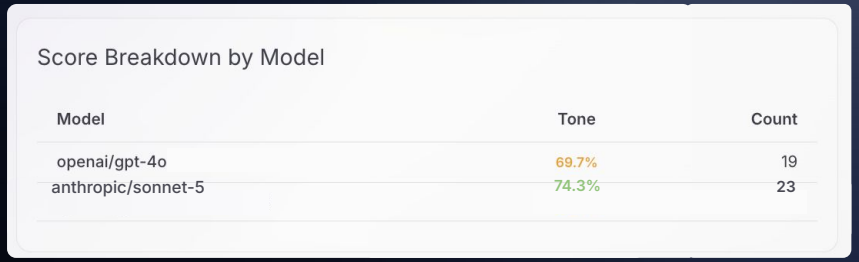
By Emma Woollacott, Senior Contributor. ⓘ Emma Woollacott is a freelance jo... Follow Author

Published Jul 22, 2025, 07:56am EDT

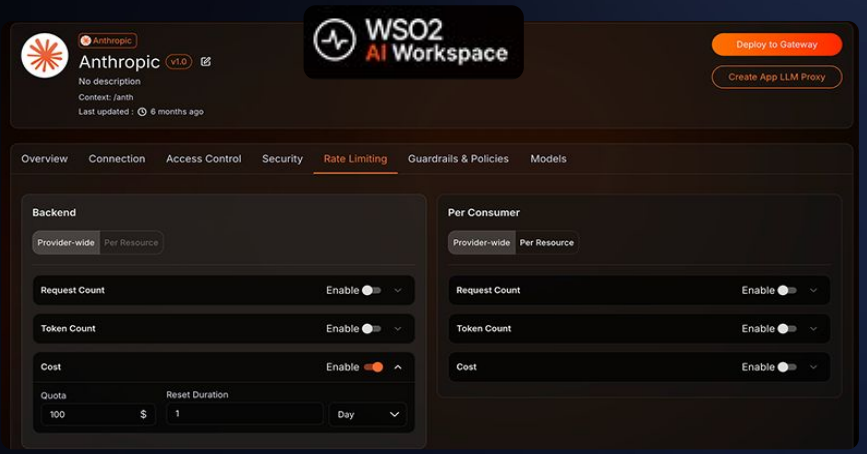
In sworn testimony before a French Senate inquiry into the role of public procurement in promoting digital sovereignty, Anton Carniaux, Microsoft France's director of public and legal affairs, was asked whether he could guarantee that French citizen data would never be transmitted to U.S. authorities without explicit French authorization. And, he replied, "No, I cannot guarantee it."

He said that the company resisted requests from the US authorities "when they are not well-founded", but that under the U.S. Cloud Act, U.S. companies can be forced to hand over data, regardless of where it is stored.

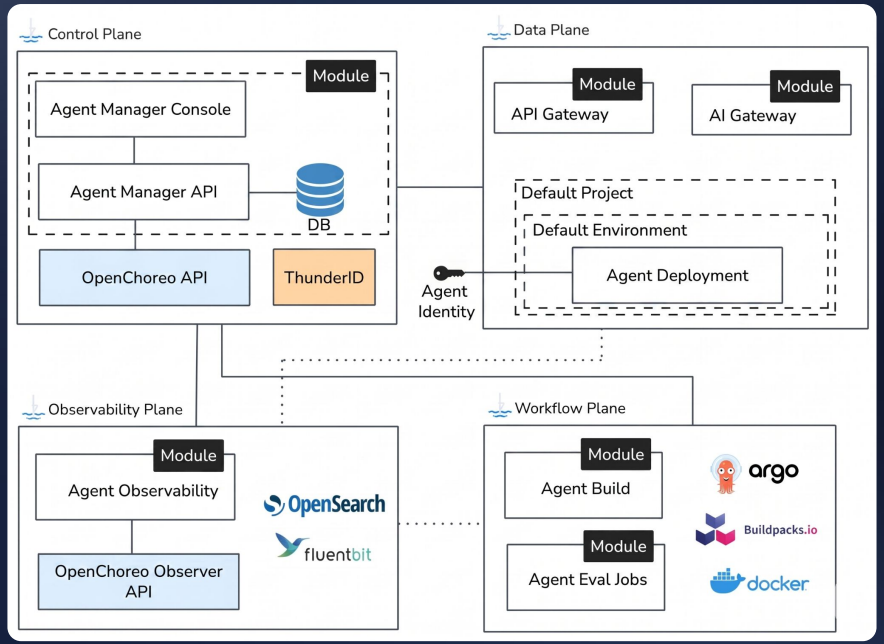
Creating a need to architect for resilience, flexibility, and control



Agent resiliency: *Switch models without rebuilding governance*

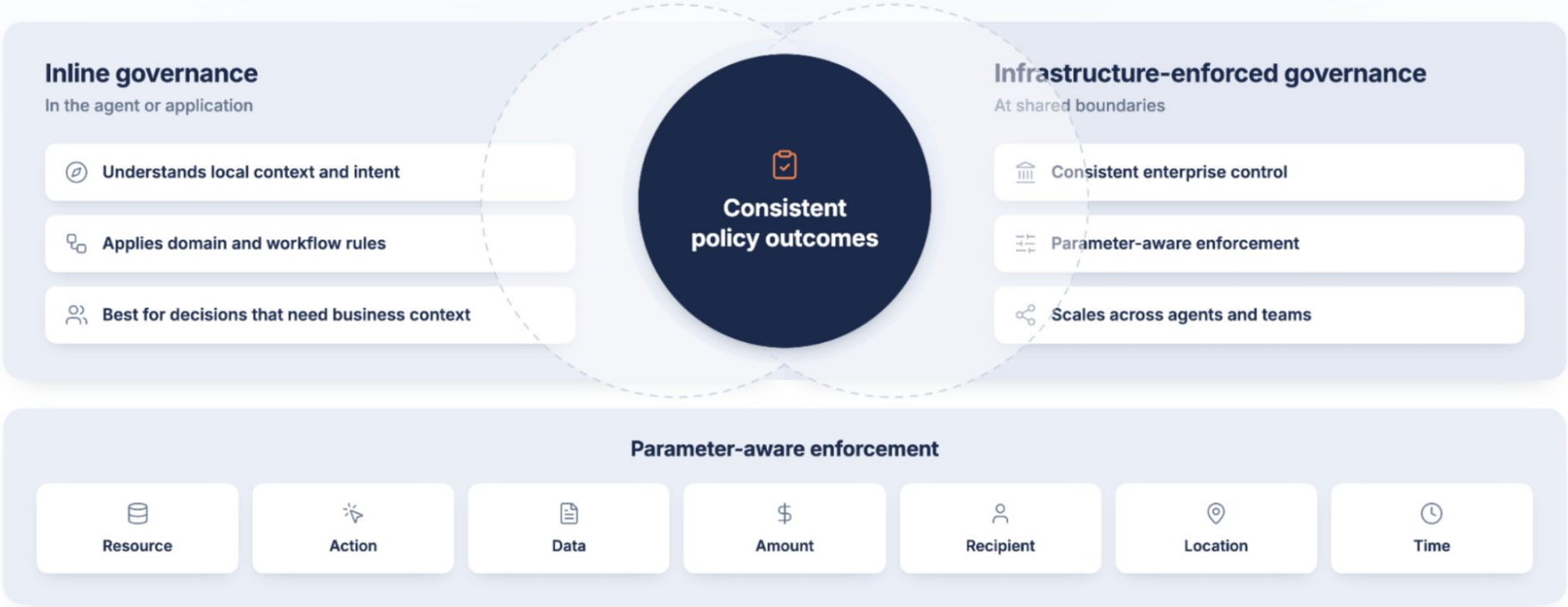


Economic controls: *Cost control as a governance dimension*



Control Plane and Data Plane Separation
Allows for centralized governance with distributed execution

Trusted AI Governance lets you move fast with control

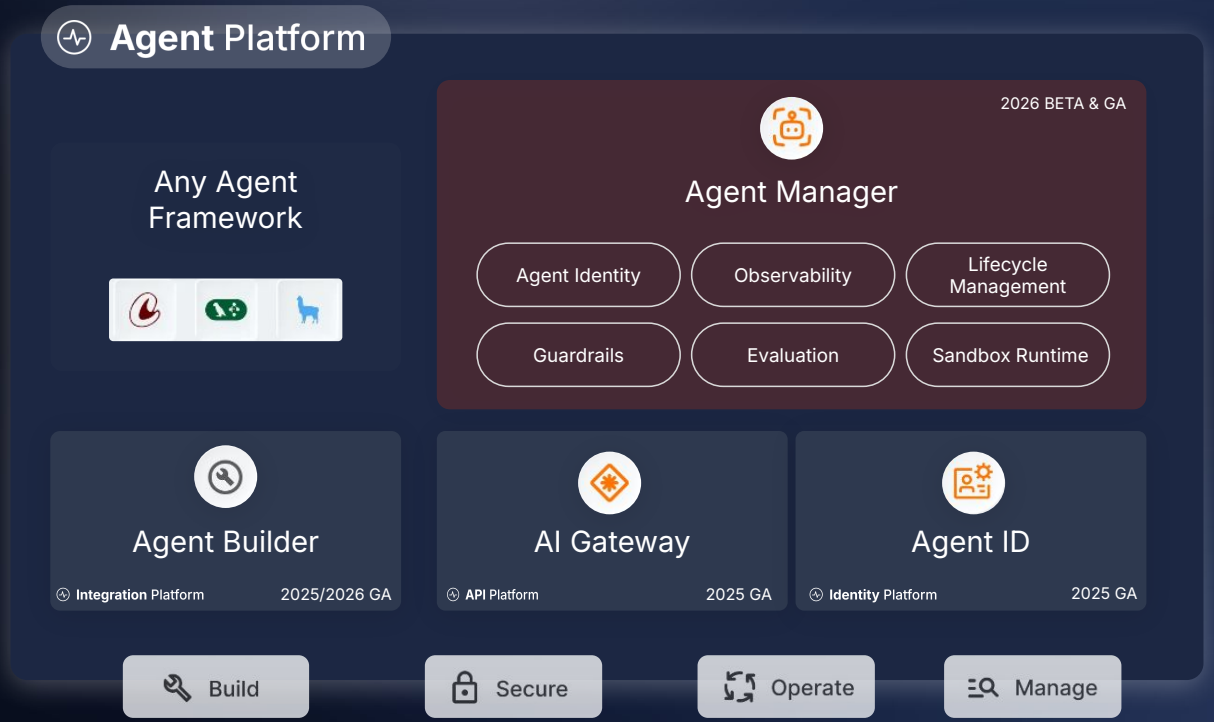


Trusted AI Governance Needs 3 Things

1. Built in, not bolted on
2. Built for AI heterogeneity
3. Built for freedom to innovate

An implementation of Trusted AI Governance:

The WSO2 Agent Platform and the Agentic Enterprise Fabric



WSO2 Agentic Enterprise Fabric

Trusted AI Governance across the stack

WSO2 **Agent** Platform	Govern every agent
WSO2 **API** Platform	Govern every API & AI service
WSO2 **Integration** Platform	Govern every connection
WSO2 **Identity** Platform	Govern every identity
WSO2 **Engineering** Platform	Govern every workload

Trusted AI Governance with WSO2

Open • Resilient • Yours

WSO2CON AFRICA

SEPT 22 - 24, 2026 | NAIROBI, KENYA

Thank you!

Dr. Rania Khalaf

Chief AI Officer & Acting
GM - AI BU