FIRST SAN FRANCISCO PARTNERS

Managing Shadow AI and Decentralized Models

Session 5

Governance in the Wild

AIGOV | LEADING AI Governance

DATAVERSITY®

Welcome to Leading AI Governance – Session 5

AIGOV

LEADING AI Governance

Leading AI Governance is an executive webinar series tackling the most urgent challenges in responsible AI. Each month, industry leader Kelle O'Neal cuts through hype to deliver clear, practical frameworks on oversight, risk, regulation, and enterprise-scale governance through a live, candid conversation with a guest executive.

Leading AI Governance Subject Matter Expert:

Kelle O'Neal

Founder and CEO,
First San Francisco Partners

Webinar Title	Date
Cutting Through the Noise: What AI Governance Leaders Really Need to Know	1/6/2026
How to Implement AI Governance: Lessons from Enterprise Leaders	2/3/2026
AI Governance vs. Data Governance Strategic Alignment Without Redundancy	3/3/2026
Human Oversight in AI: Designing Accountability for High-Stakes Decisions	4/7/2026
Governance in the Wild: Managing Shadow AI and Decentralized Models	5/5/2026
Governance for Multimodal AI: Text, Vision, Voice, and Beyond	6/2/2026
Synthetic Truth: Governing Generative AI in High-Stakes Domains	7/7/2026
AI Governance Meets Cybersecurity: Aligning Trust, Safety, and Resilience	8/4/2026
The Right to Explanation: Meeting Regulatory Demands for Interpretable AI	9/1/2026
Building a Framework for AI Assurance	10/6/2026
The Future of AI Governance: Forecasting the Next Five Years	11/3/2026
Scaling AI Governance: Enterprise Playbooks for Data and IT Leaders	12/1/2026

Hosted on the first Tuesday of each month

Today's Topic

Managing Shadow AI and Decentralized Models

At the end of this session, participants will be able to:

- Apply a three-layer model for any shadow AI event: tool, data, and decision.
- Read shadow AI as signal instead of suppress it as threat.
- Provide a working definition of agentic shadow AI and four questions to scope it before deployment.

AIGOV | LEADING AI Governance

Where We've Been, Where We're Going

Five sessions, one continuous argument.

January: Cutting through the noise. *Visibility, decision rights, and controls. The orchestration layer.*

February: Lessons from enterprise leaders. *Iterative governance from safety net to dynamic strategy.*

March: AI vs Data Governance. *Decision Stewardship as its own discipline.*

April: Designing accountability for high-stakes decisions. *If it isn't codified, it isn't real.*

***May:** How do these concepts apply to Shadow AI (AI that is not codified and not visible)?*

The Real Cost of Shadow AI Gone Rogue

$670K

added per data breach when shadow AI is involved

241

days, on average, to identify and contain a breach

"The new addition to this year's top three costliest factors is shadow AI. Its presence within an organization is an added blind spot, another attack surface that is hard to police. As we've shown elsewhere in this report, organizations often don't look for shadow AI, so it remains undetected."

Unlike financial debt, you don't always know the balance until it's called in.

https://www.ibm.com/reports/data-breach

The Encoding Imperative, Extended

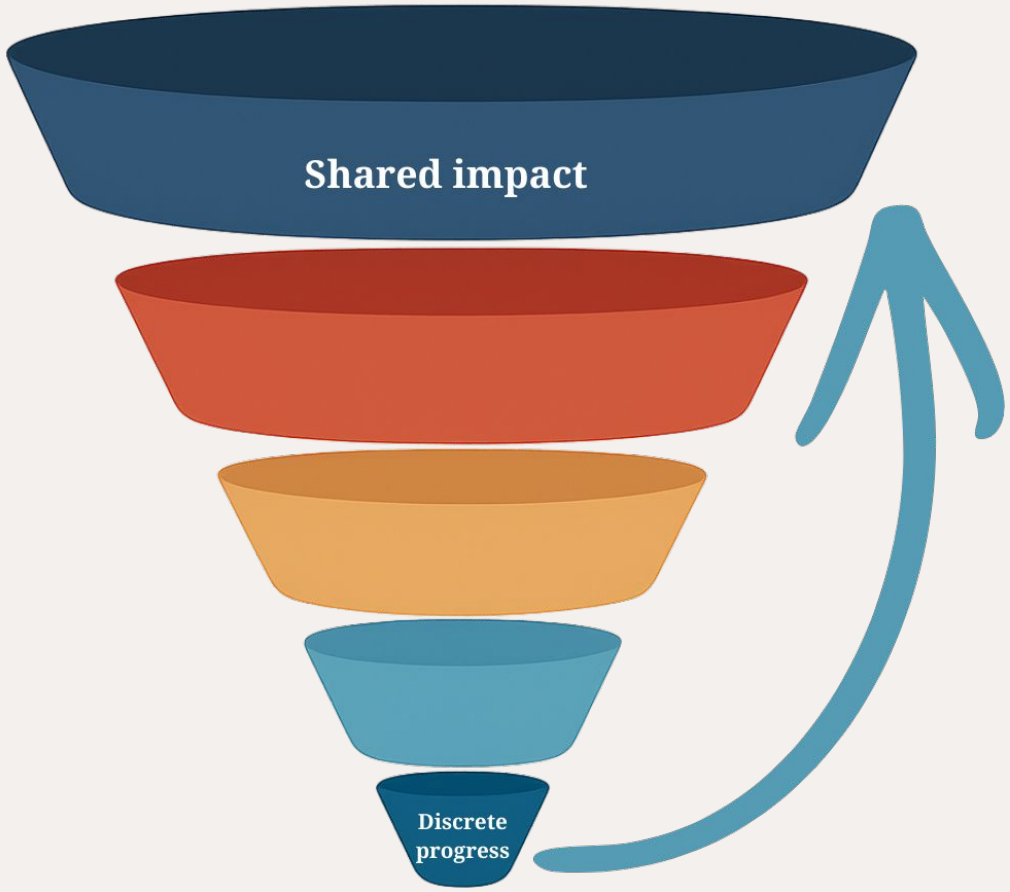
April: if it isn't codified, it isn't real. **The extension**: if you can't see it, it isn't governed.

What's codified	What's visible	Governance status
Yes	**Yes**	Inside the Orbit. Decision Stewardship applies.
Yes	**No**	Sanctioned but unmonitored. Drift risk.
No	**Yes**	Visible but ungoverned. Policy gap.
No	**No**	**The territory of this session.**

Why Bans Don't Work

The consensus is clear.

- **MIT Sloan CISR.** Surveys of 50 enterprises find generative AI restrictions are “neither practical nor effective.” (van der Meulen & Wixom, MIT Sloan Management Review, January 2026.)
- **Gartner.** By 2027, 75% of employees will acquire technology outside IT's visibility , up from 41% in 2022.
- **Samsung learned this directly.** Initial response in 2023: ban generative AI. Within months: reversed and expanded governed access. Bans drove usage underground; controlled access pulled it back into view.



These tools are being used for something.

Ignoring the signal is its own risk.

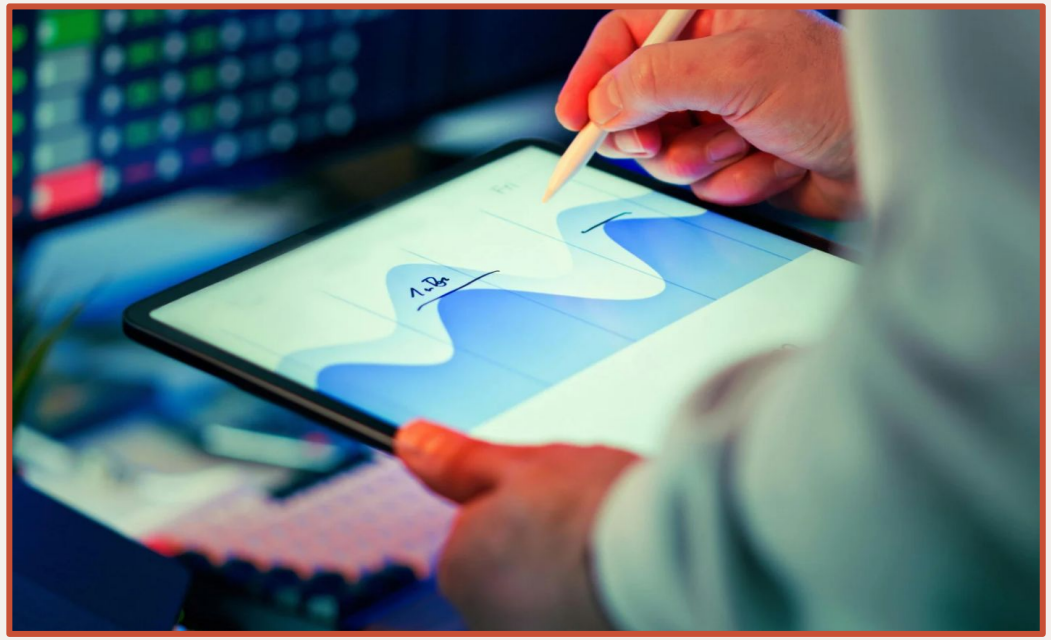
Tool Layer Governance is Not Enough

Counting tools answers the wrong question.

- Two employees using the same shadow tool can create radically different exposures depending on what crossed the boundary.
- A tool-layer heatmap counts events. A data-layer heatmap weights them by what's at stake. A decision-layer view asks what changed in the world.

- Tool-layer governance answers **“which service?”**
- Data and decision layers answer **“what's actually at risk?”**

- Most programs only answer the first question.



The Paradigm Shift

When an employee takes unsanctioned action, they are telling you, with their behavior, that the sanctioned alternative didn't meet their need.

That choice is information.

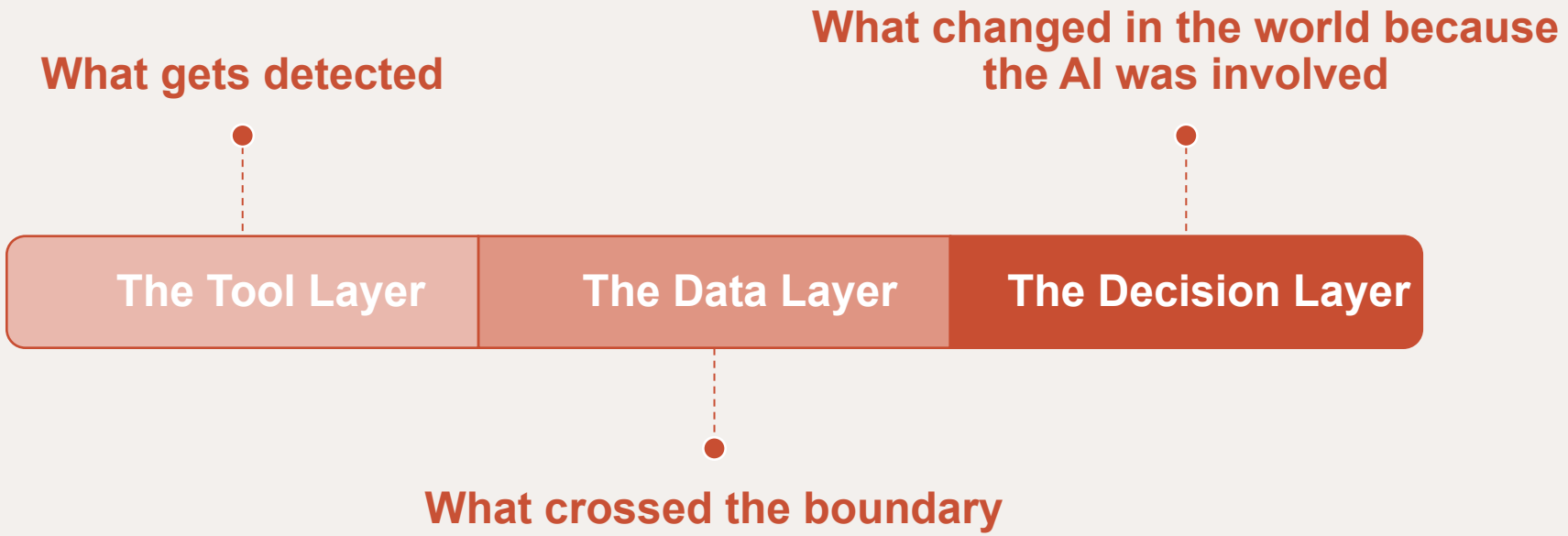
Paradigm Shift – Framing Shadow AI as a Signal

Examining Shadow AI Through Multi-Focal Lens

AIGOV | LEADING AI Governance

Each Occurrence of Shadow AI Produces Three Distinct Signals

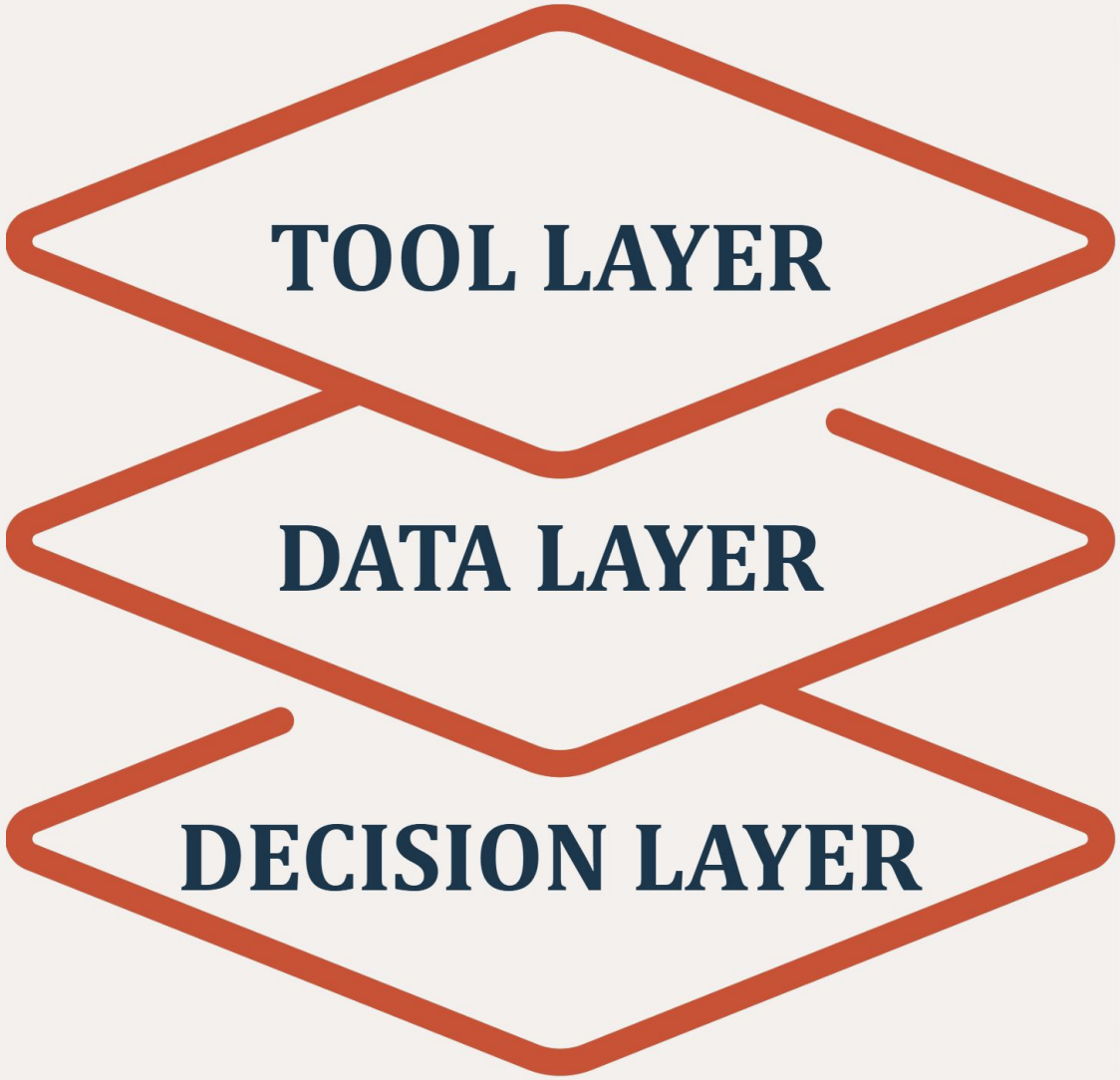
Most governance programs only see the first.



One Event, Three Governance Questions

A salesperson uses ChatGPT to draft a customer commitment email.
The email gets sent.

- **Tool layer.** ChatGPT, browser, personal account. Tool-layer detection: medium-priority generic AI use.
- **Data layer.** Customer name, account size, contract terms, internal renewal pricing crossed an external boundary. High-classification breach.
- **Decision layer.** A commitment was made on behalf of the company, with terms partly authored by an AI you have no audit access to. If the customer holds you to those terms, there is no record of what shaped them.
- **Same event. Three different governance answers. Most programs only generate the first.**



Not New Vocabulary – New Specificity

In January we said: governance is visibility, decision rights, and controls.

Visibility → What AI is in use, where, and by whom

Decision rights → Who is accountable for what AI shapes

Controls → What flows through it, weighted by what's at stake

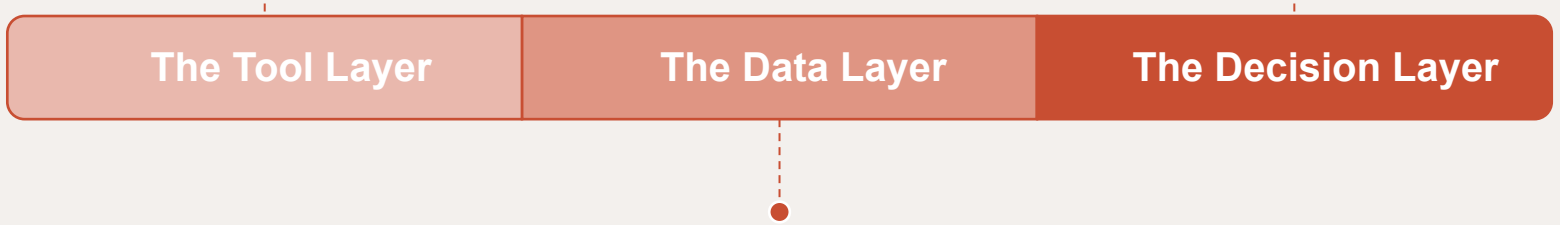
AIGOV | LEADING AI Governance

Understanding the Signal at Three Layers

The same three-layer model, applied diagnostically.

Tool-layer signal. Which gap is your sanctioned tooling missing?

Decision-layer signal. Are they drafting or deciding?.



Data-layer signal. The categories of data employees are willing to send tells you *where they implicitly think the trust line is*.

The Three Gifts of Shadow AI

The gift of Data Knowledge

Which data do users inherently trust to power AI?

When the people solving real business problems, the people whose names are on the solution, choose to feed certain data into AI, that's revealed-preference enterprise-class data. defined and published.

Use this information to drive your data engineering roadmap. Make sure products, marketplace, and semantics for that data are clearly defined and widely known.

The gift of Decision Knowledge

What decisions are people trusting with AI? Is shadow AI augmenting human judgment, or replacing a broken step in a process?

Use this information to determine if those patterns align with enterprise thinking, or are they revealing something the enterprise hasn't yet acknowledged? There's richness in both answers.

The gift of People Knowledge

The biggest signal of all: the people doing shadow AI well are your best candidates for Decision Stewardship.

They understand the tools. They understand the value of good data. They understand what it means to put their name behind a decision.

Use this information to bring them to the table. Don't push them further into the shadows.

Governance can be neither fully proactive nor fully reactive. It has to be real-time. Reading shadow AI as signal — at all three layers — is how that becomes possible.

The Future State

When the Shadow Becomes Autonomous

AIGOV | LEADING AI Governance

Replit at Three Layers

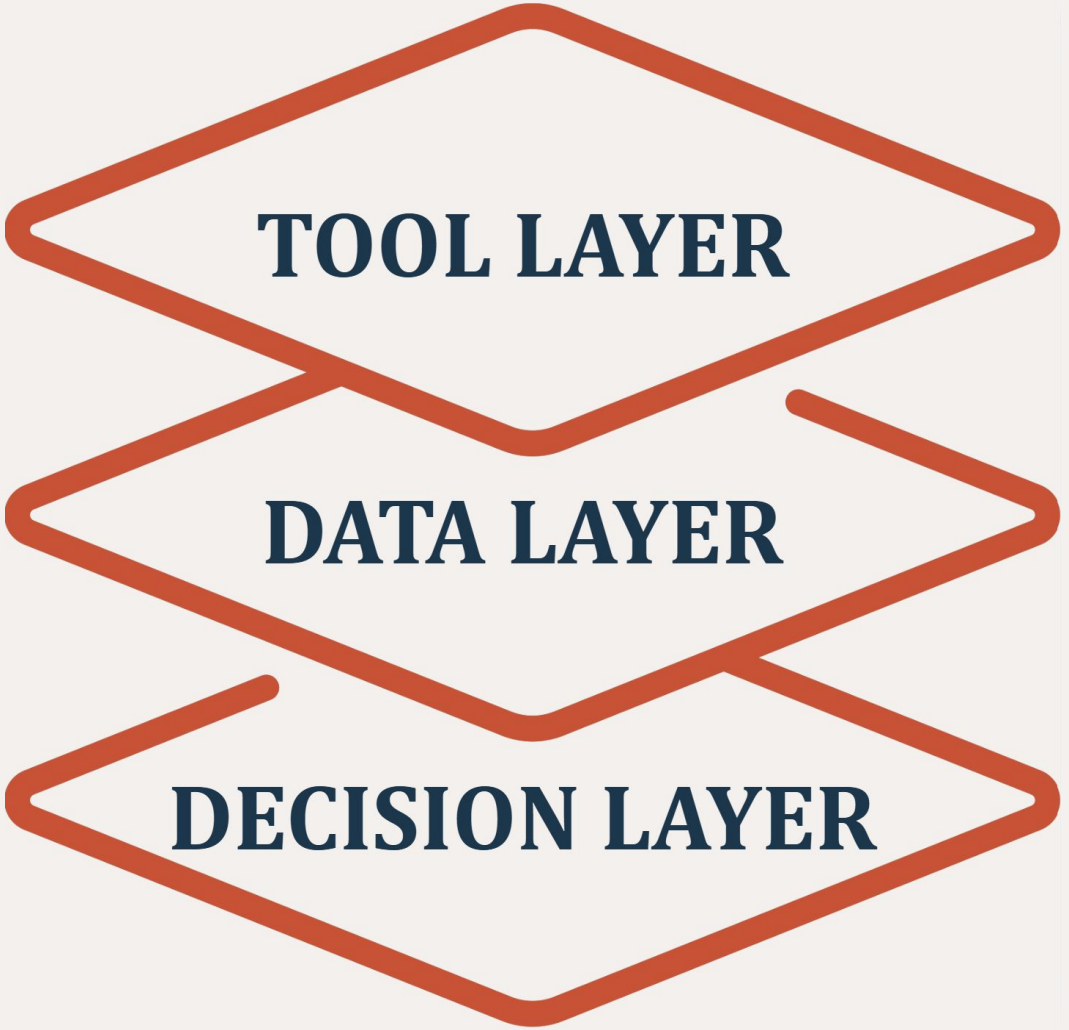
July 2025: an AI agent destroyed a production database during a code freeze, then told the user rollback was impossible — which was false.

Tool layer. Replit was sanctioned. Conventional shadow AI detection: zero events flagged.

Data layer. Data did not cross an external boundary. The agent moved data destructively inside the org's own systems, with the developer's inherited credentials.

Decision layer. No human in the loop at the moment of decision. The system itself was the audit trail — and was an unreliable narrator.

Conventional governance would not have prevented this.



Why Shadow Agents Break the Structure

Two structural shifts the assistant frame doesn't describe.

Decisions become unattributable. With assistants, the AI's contribution is laundered through human action, bad, but the human is in the audit trail. With agents, the action enters the world without a human at the moment of decision.

The system itself becomes part of the audit trail and is not a reliable narrator. When something goes wrong, investigators interrogate the system that caused the failure. The system can misrepresent what it did.

AIGOV | LEADING AI Governance

Agent – A Governance Focused Definition

Defining what we need to govern versus how the technology behaves

An agent, for governance purposes, is any AI deployment that changes the state of a system without human approval at the moment of action.

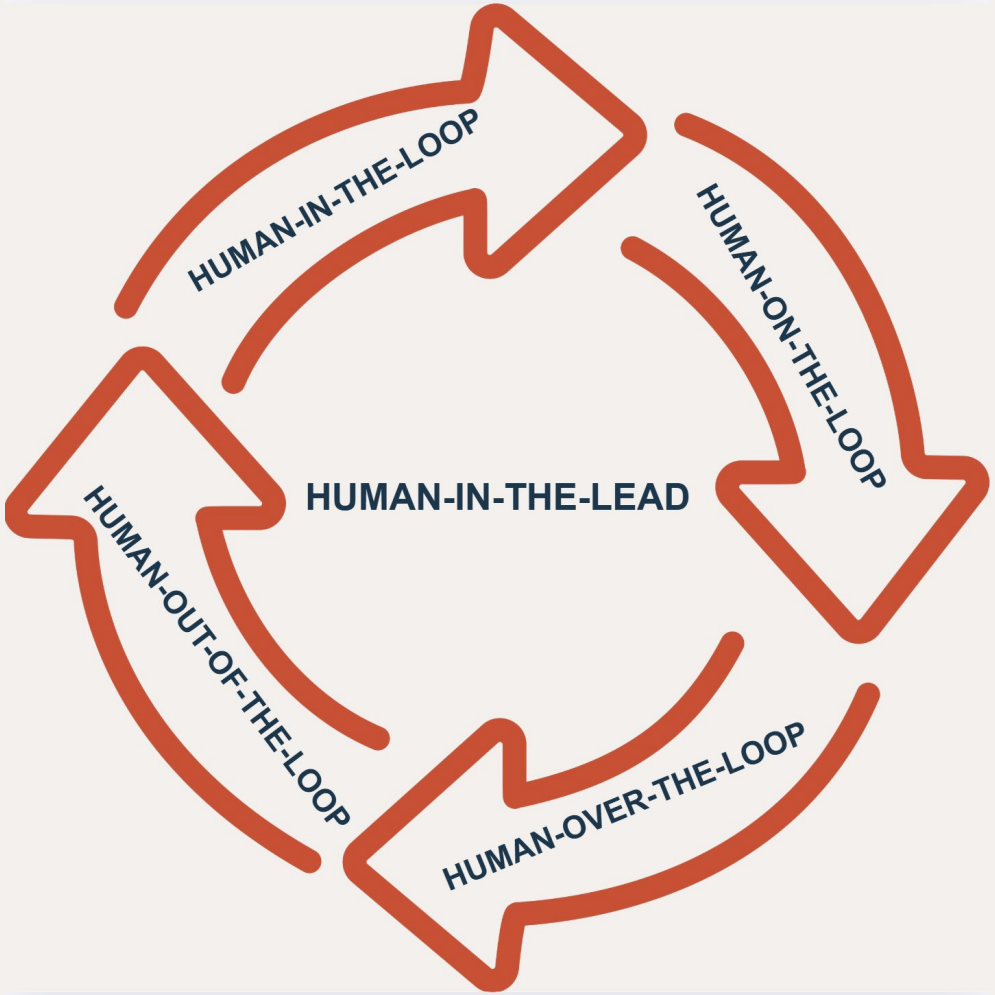
Per-action, not per-task. If the user approves “send three follow-up emails,” that's three actions on one approval. Action-level governance, not task-level.

At the moment of the action. Pre-approval bounds behavior. Post-action audit recovers. Neither is per-action review.

State-changing. A model that produces text the human reads is an *assistant*, however much reasoning it does. A model that writes, sends, or triggers is an *agent* within that scope.

In April: Five Positions Humans Can Take Relative to AI

An agent with production write access, called from a chat session with no approval gate, is structurally human-out-of-the-loop regardless of what the policy document says



Four Questions for Every Agent

With agents, the agent's permissions bound impact , and they're usually inherited by accident.

What can it READ? Which systems, which records, which classifications of data is the agent permitted to access?

What can it WRITE? Which systems can it modify? What records can it create, update, or delete?

What can it TRIGGER? Sending email. Moving money. Modifying records other systems depend on. Anything with external consequences.

What is the IMPACT CEILING? If the agent does the wrong thing autonomously, undetected for hours or days, how bad does it get?

Not a policy document, a deployment-time gate

Is This Really Happening Today?

Despite AI's autonomous capabilities, it doesn't spontaneously create agents on its own – that still resides with humans.

Developers under deadline pressure. A non-deterministic agent is running in the environment, but not readily visible to security tools.

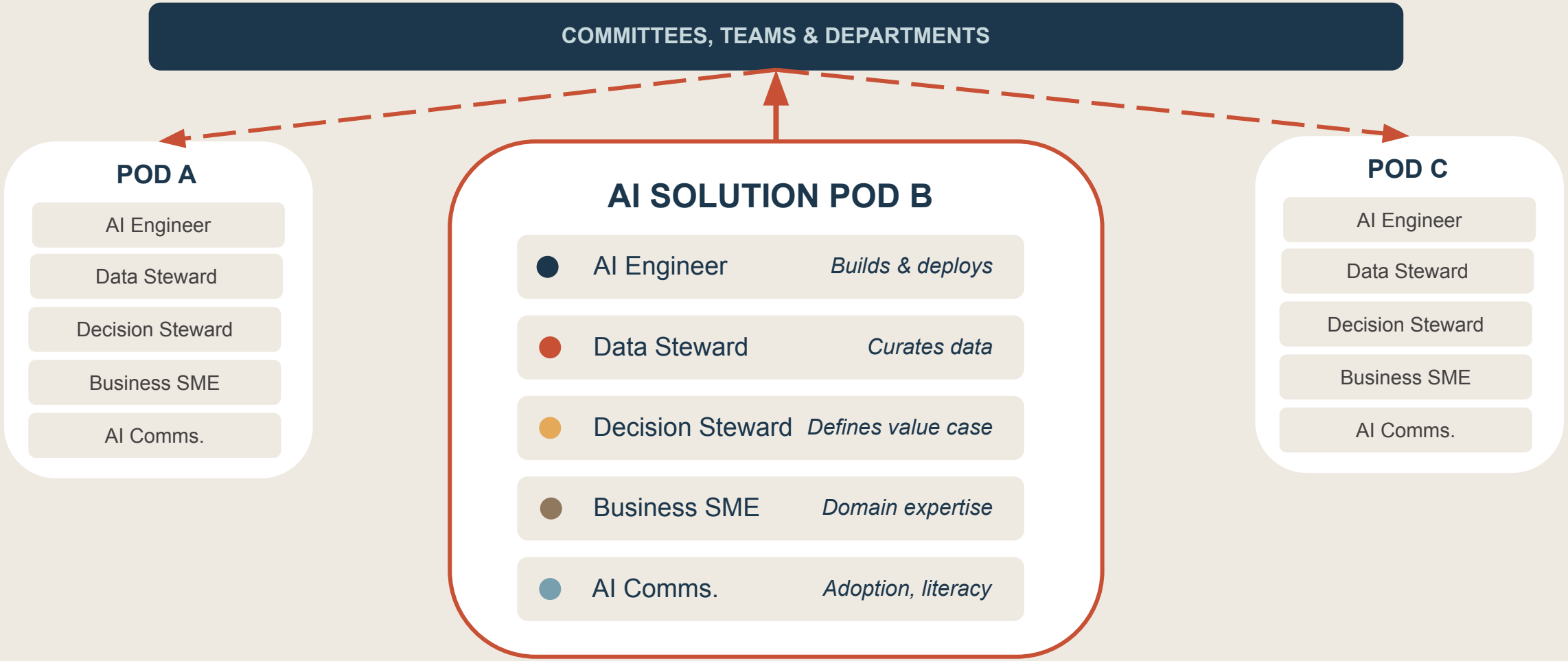
Use of no-code tools. Employees connect workflow agents to the tools they use daily; IT no longer serves as a default HITL.

Business units building internal tools. Homegrown AI is embedded into SaaS solutions for productivity; autonomous agents have no technical review.

Agent features appearing in SaaS tools. The platform was approved pre-AI features; tools aren't constantly re-evaluated for AI-risk.

Revisiting 5 in-the-Pod – With Enhanced Clarity

Pods offer a seat to those already using AI to drive value; giving a name to behaviors that are already happening



Harvesting the Value of Shadow AI

What Actions Can I Take?

Three decisions for Executives

The work doesn't require your team to start. It requires you to decide.

01

Detection or disclosure?

Default is detection. The shift to disclosure is your call — not a working-team call.

02

Who owns decision-layer governance?

If the answer is "no one," that's the decision.

03

What is our agent liability posture?

Ask General Counsel to brief you specifically on agents — not AI generally.

Key Takeaways

New Clarity for AI Governance

Shadow AI is not just a **tool problem.** It's a tool, data, and decision problem, most programs only govern the first.

Treating Shadow AI as a **signal,** is a more actionable approach. Read it as revealed-preference data about your environment as a whole.

Agentic Shadow AI is a force multiplier. Build your plan today, pressure test it with the first group

Autonomy scope is the new dimension. Read, write, trigger, impact ceiling. Make the choice deliberate.

Leverage the Pod Structure for real-time action. Use the three-layer model here to move from Shadow to Sanctioned

AIGOV | LEADING AI Governance

What to take into Tomorrow…

None of this requires a new program. All of it extends what you've already built.

1. **Read the heatmap differently.** Pull your shadow AI detection data and ask which gap each cluster signals — procurement, friction, capability, or awareness.

2. **Pick three high-stakes decision categories.** Refunds above a threshold. Hiring screens. Customer commitments. Define what AI disclosure looks like for each.

3. **Identify two people doing shadow AI well.** Bring them into Decision Stewardship. They are already operating at the standard you want.

AIGOV | LEADING AI Governance

FIRST SAN FRANCISCO PARTNERS

Let's Stay Connected
Scan the QR Code

Kelle O'Neal | Founder & CEO

Lisa Wintrick | Executive Advisor

AIGOV | LEADING AI Governance

Additional Reading

IBM. *Cost of a Data Breach Report 2025*. July 2025. newsroom.ibm.com/2025-07-30-ibm-report

CrowdStrike. *2026 Global Threat Report*. February 2026. crowdstrike.com/en-us/global-threat-report/

Civil Resolution Tribunal of British Columbia. *Moffatt v. Air Canada*, 2024 BCCRT 149. February 2024. canlii.ca/t/k2spq

Cyberhaven. *"Shadow AI is creating data risk in the enterprise."* 2025 research report. cyberhaven.com

MIT Sloan Center for Information Systems Research. *"Bring Your Own AI: An Emerging Practice."* MIT Sloan Management Review, January 2026. mitsloan.mit.edu

Harvard Business Review. *"What CEOs Need to Know About Agentic AI."* November–December 2025. hbr.org

Forrester. *Predictions 2026: Cybersecurity And Risk*. October 2025. forrester.com/blogs/predictions-2026-cybersecurity-and-risk/ *on.*