



Synthetic Truth

Governing Generative AI in High Stakes Domains



Welcome to Leading AI Governance – Session 4



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



CUTTING THROUGH THE NOISE: WHAT AI LEADERS REALLY NEED TO KNOW

January

AI governance = visibility, decision rights, and controls that keep AI aligned with business, risk, and regulatory expectations.

Defining AI Governance in 2026

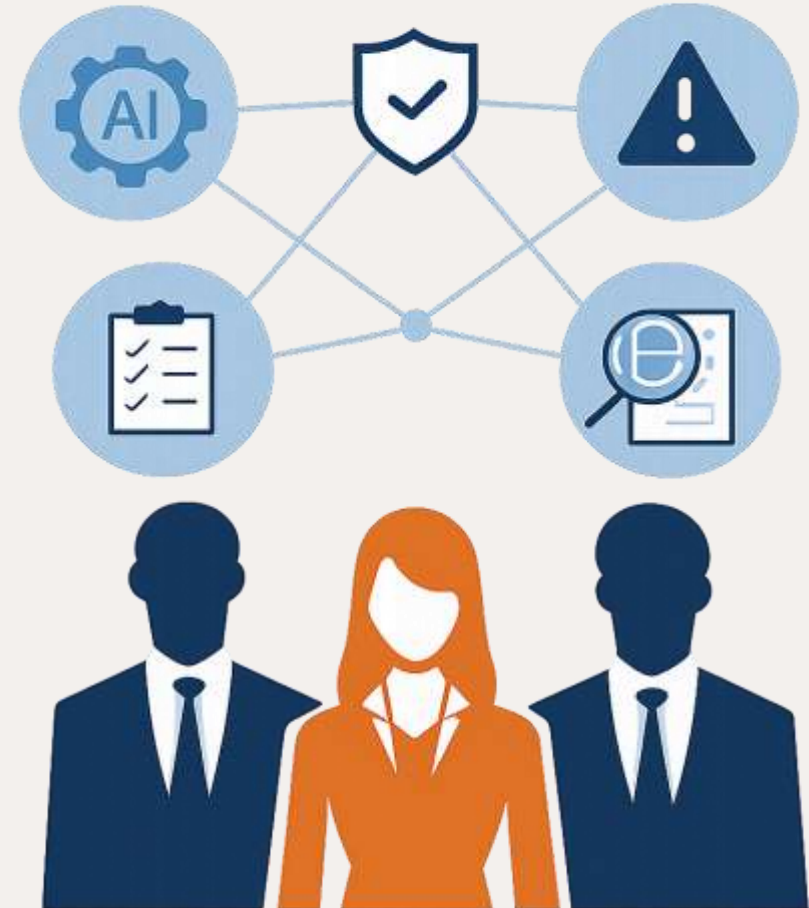
What it is, what it covers, and who's accountable

AI Governance is:

- AI Governance is the organizing framework for establishing the strategy, people, and processes for responsible creation and management of AI solutions in support of organizational goals.
- A combination of visibility, decision rights, and controls that ensures AI aligns with business, risk, and regulatory expectations.

AI Governance covers:

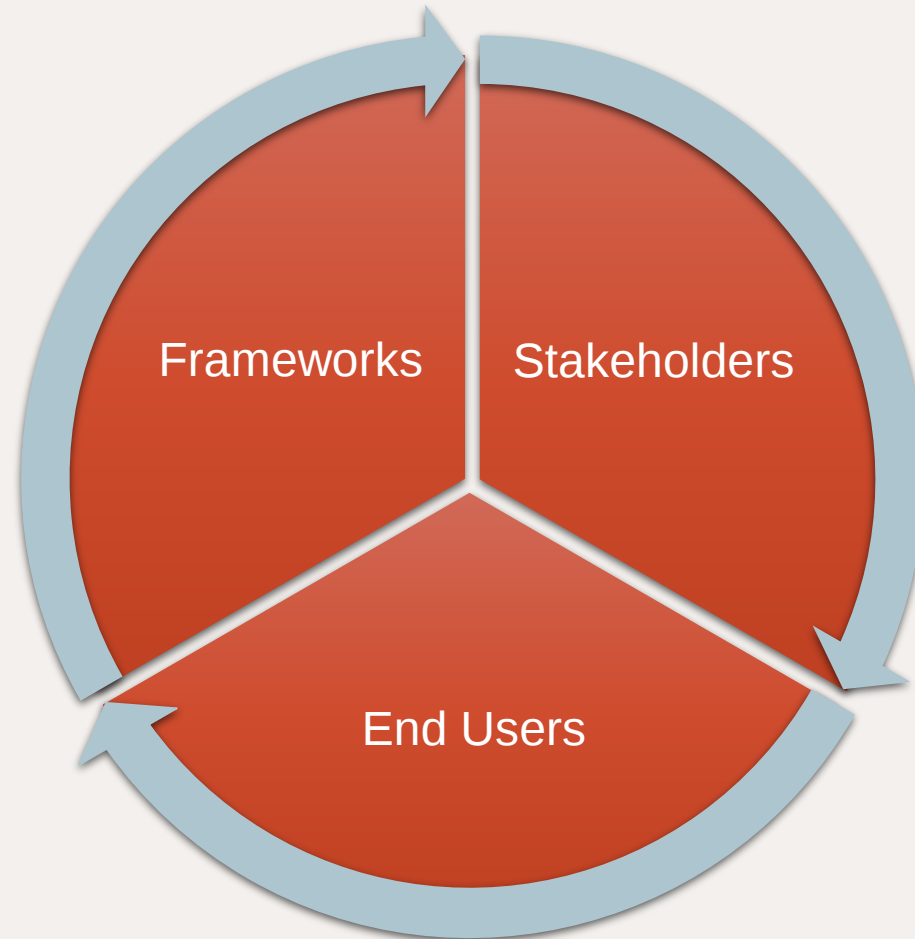
- Decision making rights
- Use-case visibility and ownership
- Risk tiering and approval paths
- Ongoing monitoring and escalation
- Evidence and accountability (not just intent)



AI governance turns AI ambition into accountable, repeatable, actionable decision-making.

Success by Design

Creating an Orchestration Layer that cuts through the clutter, gives everyone a voice and aligns effort to value is the single most effective way to manage chaos.



What are your “20 Questions”?

The journey forward starts with a few thought-provoking questions that require purposeful answers and spark meaningful action.

How will this model behave as it scales?

What decisions are we delegating to AI?

What does AI failure look like?

How do you know?

How are we communicating our intentions clearly and honestly?

How would we explain this decision to ...?

How do we know those decisions align to our values?

Can management assure us this system is safe?

How are we reskilling employees for an AI-augmented future?

Are you prepared?

HOW TO IMPLEMENT AI GOVERNANCE: LESSONS FROM ENTERPRISE LEADERS

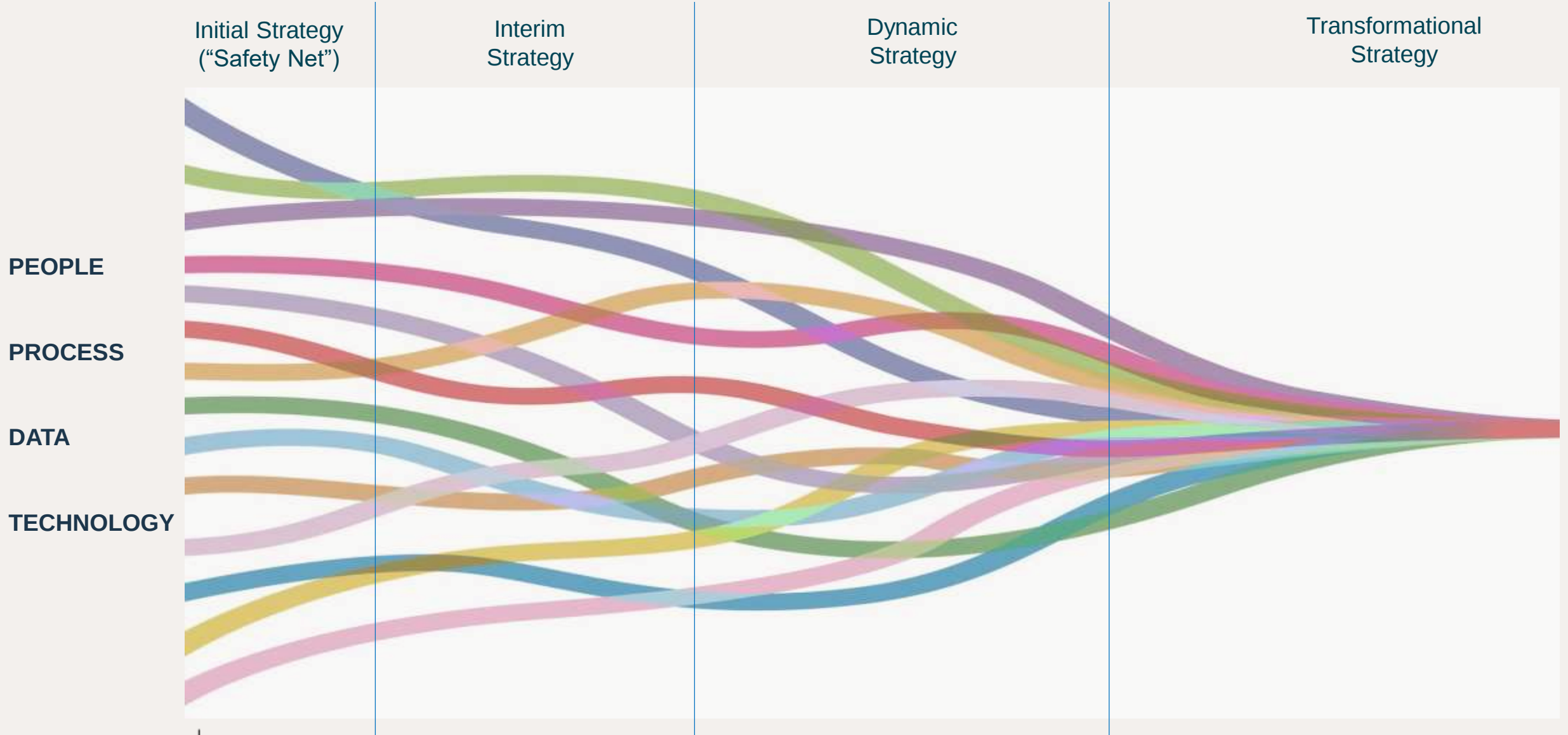
February

A New Paradigm for “People” Governance

“People” needs to adapt to include agentic peers.

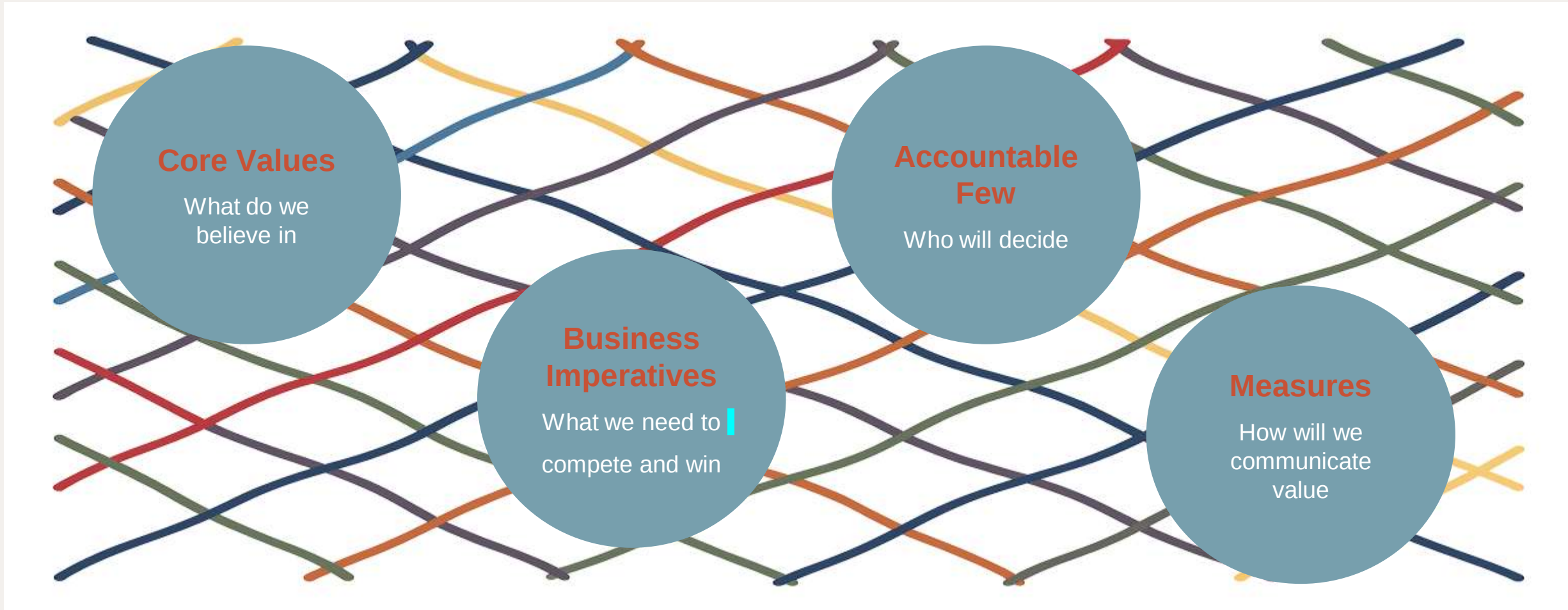
In addition to considering governance for the “what” you need to consider governing for the “how”.

AI Governance is Iterative, Scalable and Adaptable

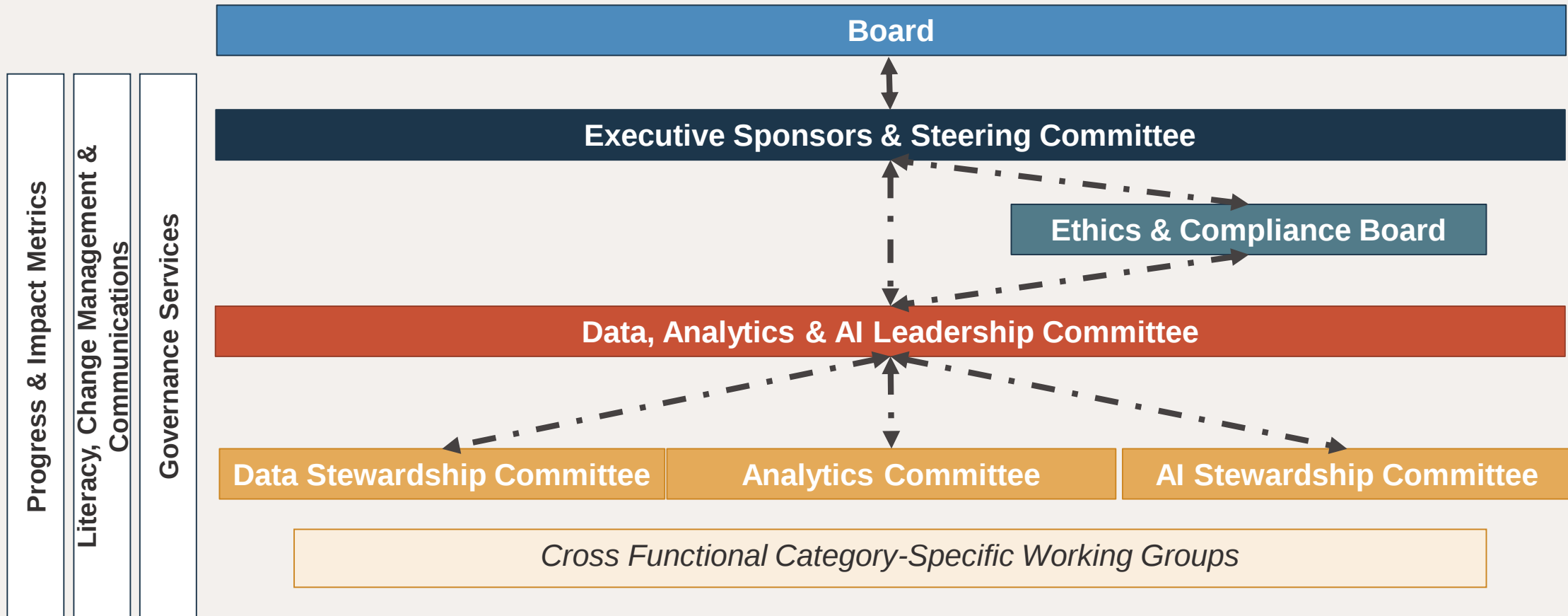


Start with a Safety Net

Creating an initial set of non-negotiables keeps you secure and allows for evolution



The Orchestration Layer = Decision-Making and Accountability



A background image showing three people in a meeting. A man with a beard and glasses is on the right, looking towards the left. A woman with glasses is in the center, looking towards the left. A man with a beard is on the left, looking towards the center. They are all sitting at a table, and there is a glass of water in the foreground on the left.

AI GOVERNANCE VS DATA GOVERNANCE: STRATEGIC ALIGNMENT WITHOUT REDUNDANCY

March

Initial Belief:

Data Governance would evolve into AI Governance

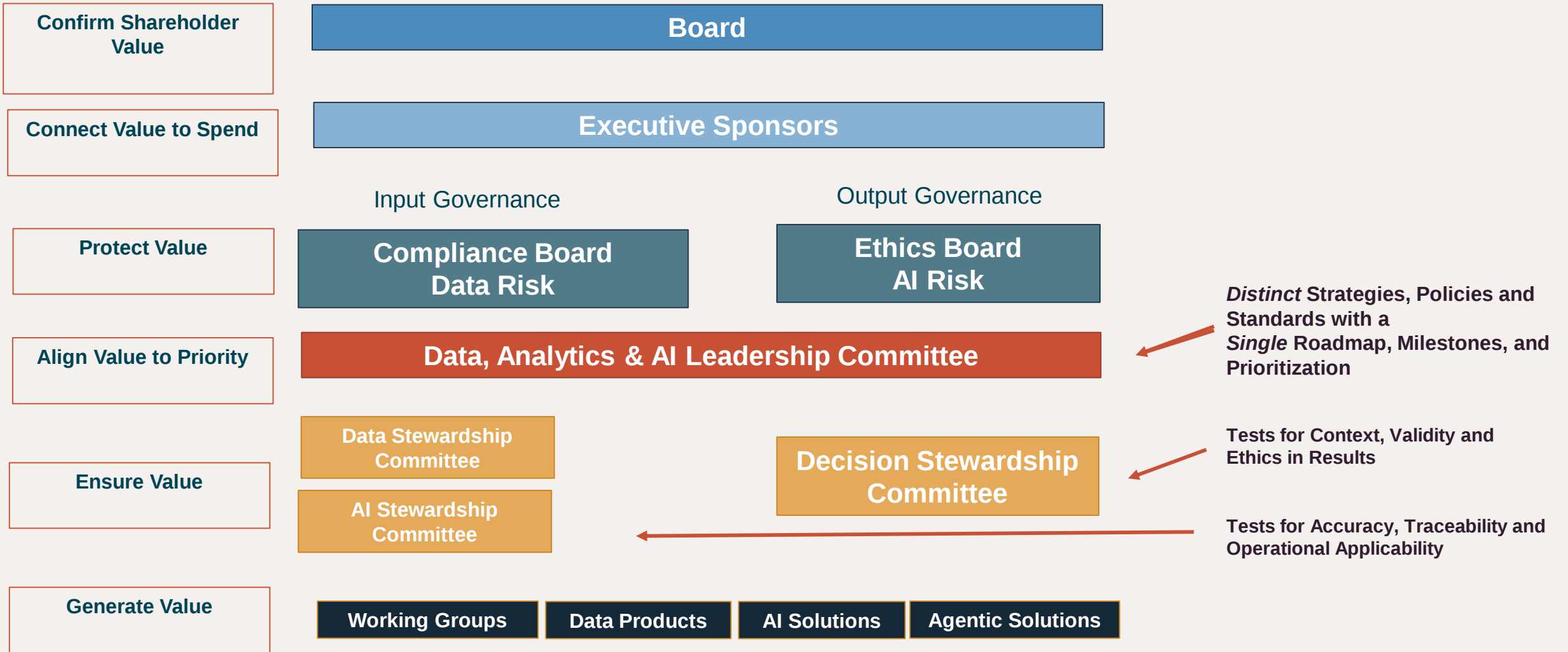
Current Understanding:

Data Governance evolves in service to AI. AI Governance runs in parallel

Data Governance vs AI Governance – Comparison Matrix

Dimension	Data Governance	AI Governance
Purpose	Organizing framework to establish strategy, objectives, and policy for managing enterprise data	Organizing framework to establish strategy, people, and processes for responsible AI solutions
Scope	Managing enterprise data assets	Responsible creation and management of AI solutions
Key Mechanisms	Policies, standards, accountabilities, and technologies [<i>policy-driven</i>]	Visibility, decision rights, and human-in-the-loop controls [<i>human-centered</i>]
Decision Rights	Decision-making rights around data	Decision-making rights around AI solutions
Risk & Oversight	Risk identification and mitigation across the data lifecycle [<i>lifecycle-scope</i>]	Risk/Ethics identification and mitigation across the AI lifecycle [<i>ethics + risk</i>]
Accountability	Standards, directives, and stewardships	Evidence and accountability beyond intent [<i>evidence-based</i>]
Value Delivered	Trusted Data	Explainable Solutions

Decision-Making in an AI-Centric Organization



Decision Stewardship Defined

The practice of ensuring the organization knows exactly what problem it is solving, for whom, and at what risk – before any AI solution is built or deployed.

Decision Stewardship is:

- The critical examination of decisions to ensure alignment to organizational values, societal ethics, and government compliance.
- Ensuring logical and explainable results that are grounded in the specificity of the enterprise yet expansive to create new opportunities and advancements that can be leveraged for competitive advantage and growth.

Decision Stewardship Covers:

- Human in the Loop
- Reasoning
- Ethics Evaluation
- AI as a Co-Worker Assessments

HUMAN OVERSIGHT IN AI: DESIGNING ACCOUNTABILITY FOR
HIGH STAKES DECISIONS

April

**Risk is no longer
just a property of the outcome.
It's also a property of the decision process.**

Unintended consequences can have an outsized impact if the right questions aren't asked before building the solution - even if execution is perfect. Therefore, risk also lives in the decision-making process, not just its result.

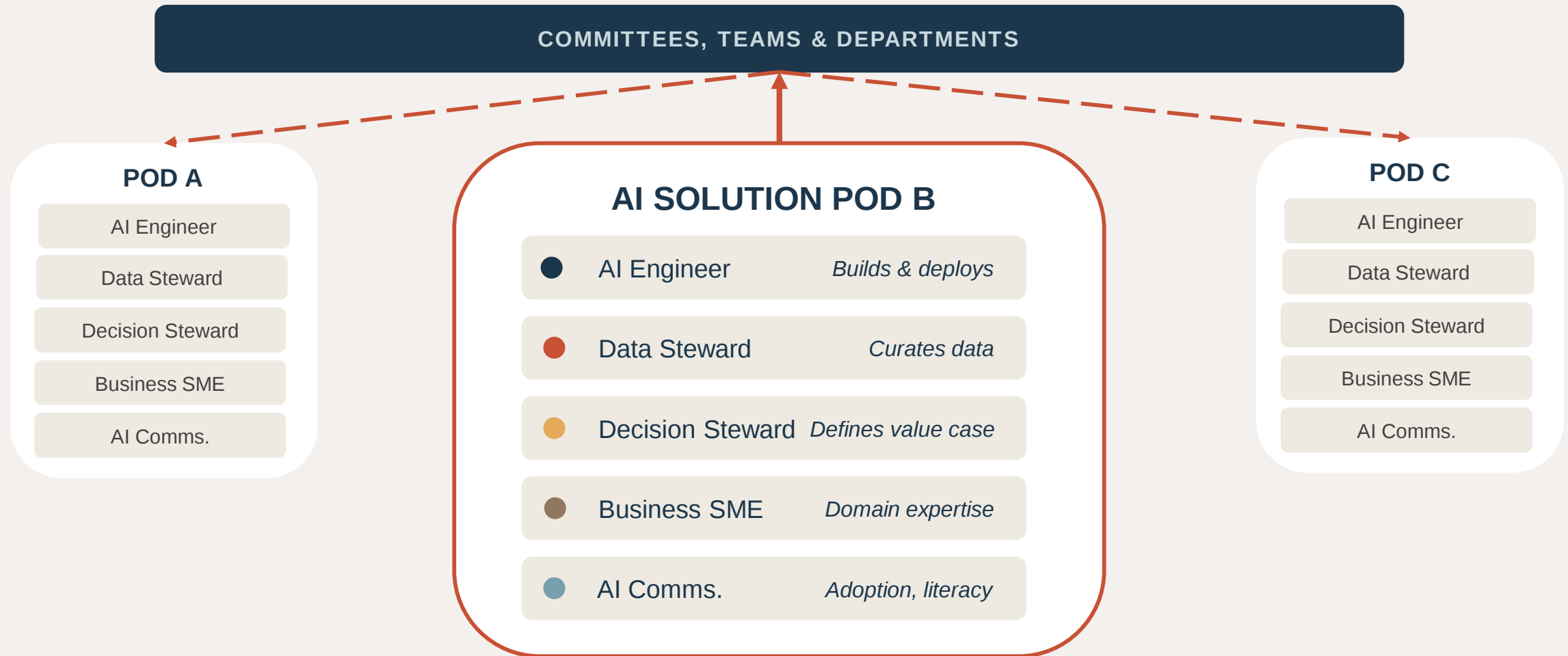
Accountability is more than a moment

It's a deliberate, consequential structure across the entire lifecycle



5 in-the-Pod – the connective tissue for governance at scale

Each pod builds targeted AI solutions at speed, while enabling lateral thinking, upward alignment, and accountability



For AI, if it isn't codified – *it isn't real*

Why this matters for AI

HUMAN WORKER

Humans have instincts, AI does not.

- Humans **notice, question, and seek out** missing accountability components
- Draw on **intuition, relationships, and experience** to navigate ambiguity
- Nuance is noticed by humans **not by AI** – unless it is translated

TWO FORMS OF BUSINESS

Humans inhabit two worlds, but AI only one.

- **Business as artifact** – org charts, policies, process maps
- **Business as living process** – collective decisioning, behaviors, negotiation
- Humans read the org chart **and** the room
- AI operates only on what has been encoded to be **machine readable**

THE ENCODING IMPERATIVE

Defend AI's decision. Don't guess why.

- AI **can and will infer** from whatever context is available
- **Explainability and transparency** are ethical and risk requirements
- Encoding the accountability framework gives you a **defensible basis** for every decision AI makes
- Codification holds ourselves – and **AI as our coworker** – accountable

GOVERNANCE IN THE WILD: MANAGING SHADOW AI AND DECENTRALIZED MODELS

May

The Paradigm Shift:

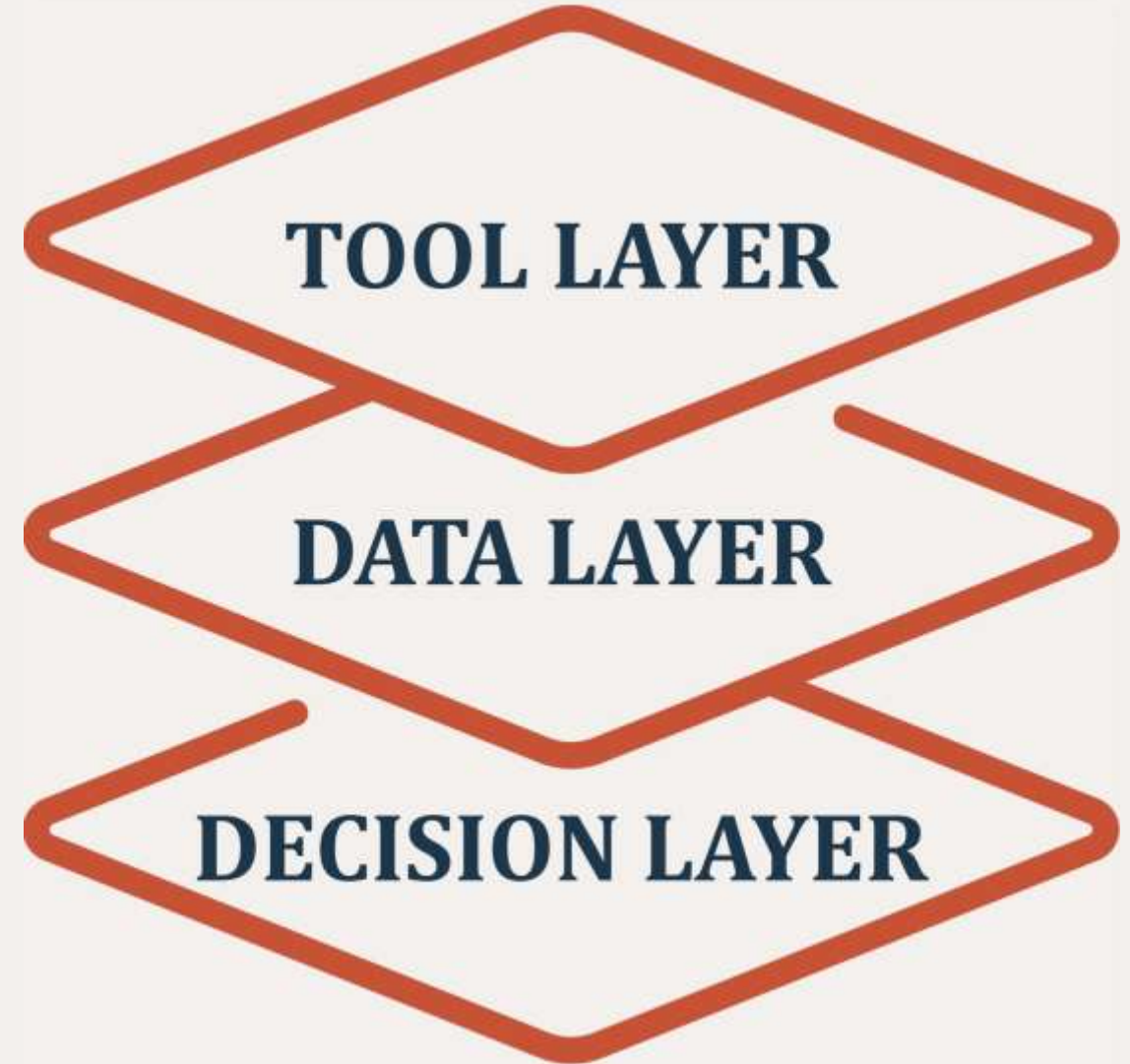
When an employee takes unsanctioned action, they are telling you, with their behavior, that the sanctioned alternative didn't meet their need.

That choice is information.

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.

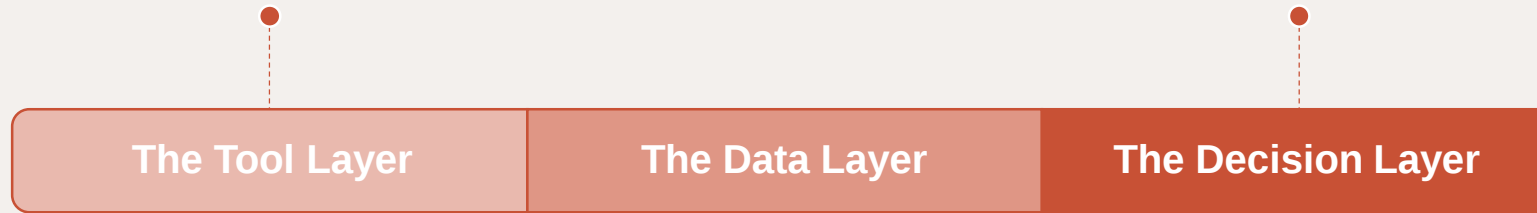


Understanding the Signal at Three Layers

The same three-layer model, applied diagnostically.

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

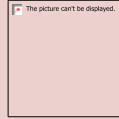
Decision-layer signal. Are they drafting or deciding?.



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

Key Takeaways

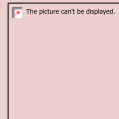
New Clarity for AI Governance



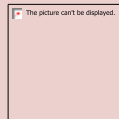
Shadow AI is not just a tool problem. It's a tool, data, and decision problem, most programs only govern the first.



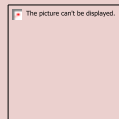
Treating Shadow AI as a signal, is a more actionable approach. Read it as revealed-preference data about your environment as a whole.



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



GOVERNANCE FOR MULTIMODAL AI: TEXT, VISION, VOICE AND BEYOND

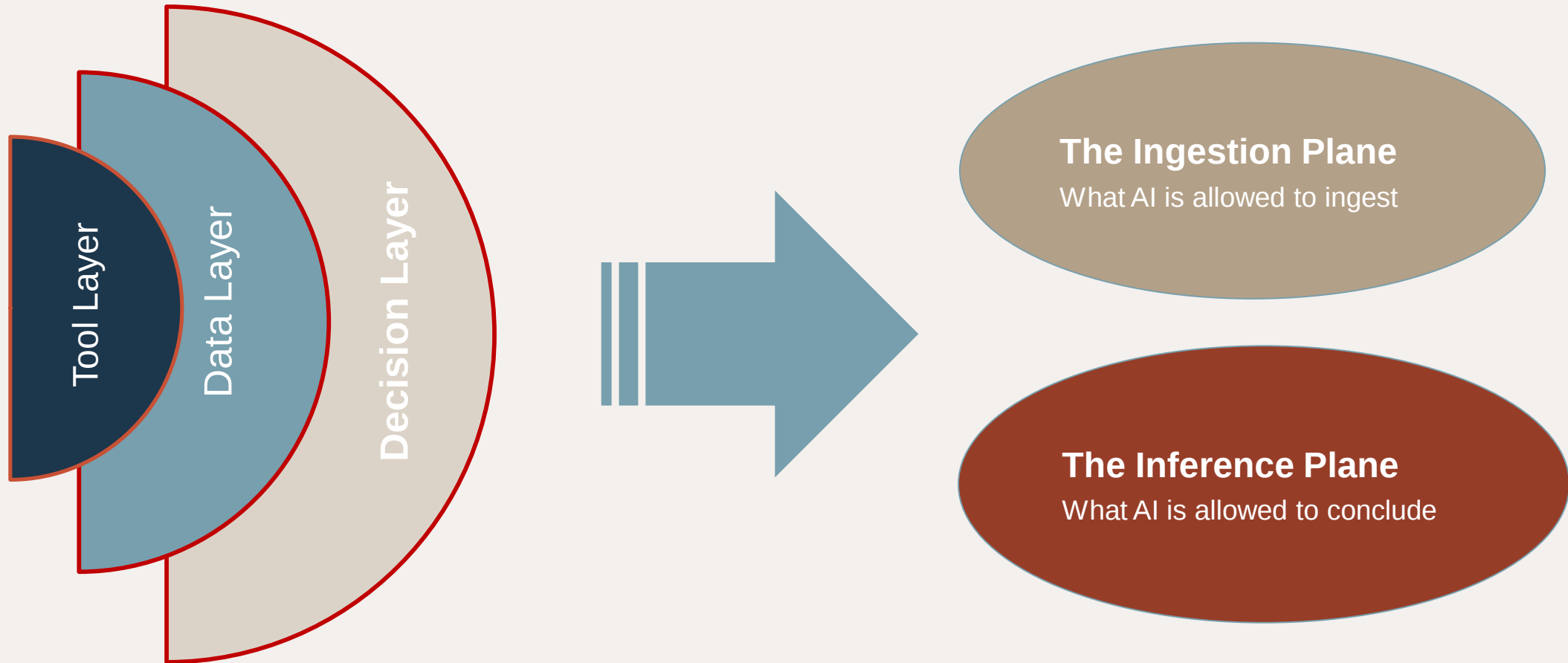
June

Multimodal AI Requires a New Level of Enterprise Collaboration

Data Governance, Privacy, Ethics and Compliance have new work. AI

Governance is at the center and connects requirements to decisions, in real time.

Applying the Decision Layer to Multimodal Data



One Event > Two Planes > One Decision

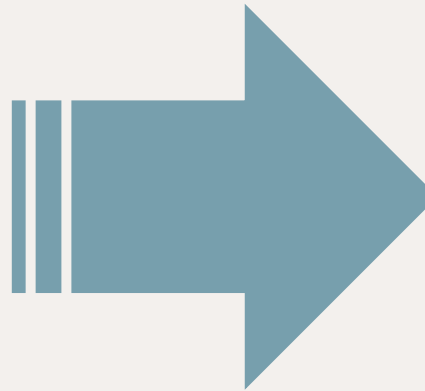
A meeting transcript under the microscope

The Ingestion Plane

The bot takes in: Tone.
Speakers. Words. Pace.
Pauses. Voice.

The Inference Plane

The bot concludes: Who
dominated. Who disengaged.
Sentiment. Action items.
Deal probability. Emotional state.

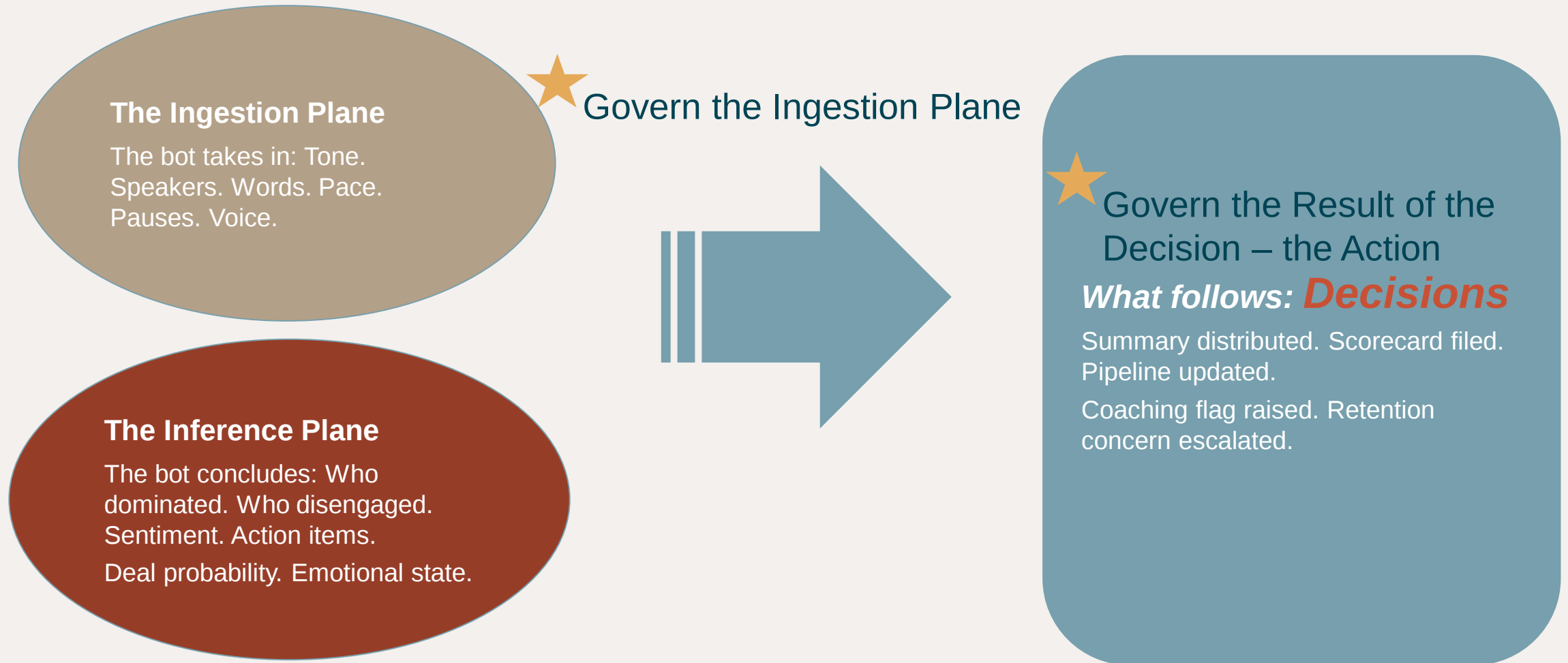


*What follows: **Decisions***

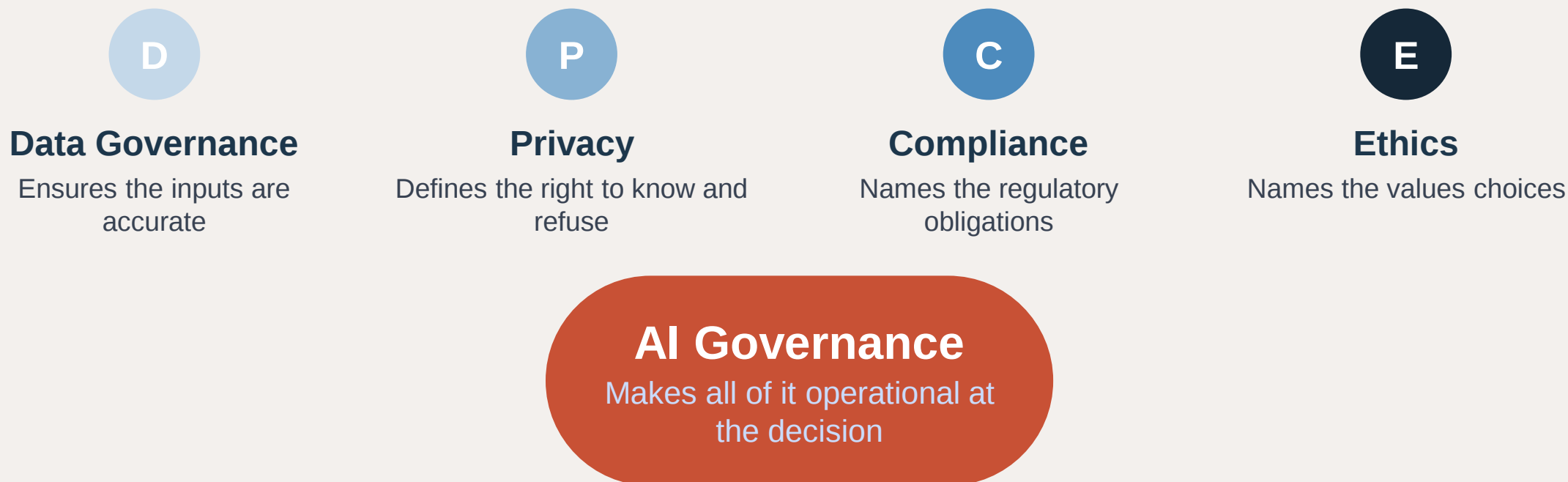
Summary distributed. Scorecard filed.
Pipeline updated.
Coaching flag raised. Retention
concern escalated.

One Event > Two Planes > One Decision

A meeting transcript under the microscope



AI governance is the discipline that makes the requirements operational at the moment a decision gets made.



The pod is where it happens. Multi-discipline roles. Every decision.
The decision is what differentiates AI Governance from Data Governance.

A woman with long dark hair is shown in profile, looking at a laptop screen. The screen displays various data visualizations, including a pie chart and bar graphs. The background is slightly blurred, showing other people in a meeting setting. The overall image has a dark, muted color palette with a purple tint.

GOVERNING GENARATIVE AI IN HIGH STAKES DOMAINS

Synthetic Truth

Welcome to the Board Room...

This is the question many of our clients are facing today

Do you *trust* what AI is telling us?

The truth is... “it’s complicated”

Wrong vs. fabricated

- We've always known AI can be wrong, that's why a human owns the decision.
- Now we need to address the fact that it can invent:
 - A citation for a case that doesn't exist
 - A drug interaction that was never documented
 - A research study, a quote, or a figure with no real source

"Synthetic truth": output that looks true but isn't.



The plot twist we haven't discussed yet...

AI didn't get it wrong. It made it up.

Bad input is a real value that happens to be wrong, like a stale figure or a typo. It came from somewhere, so you can trace it and fix it.

A fabrication came from nowhere. The model produced a fact that was in no file, no feed, no record. It arrives in the same confident, fluent voice as everything true, so nothing looks off.

In this case, clean data doesn't save you. Every control we built checks the decision and trusts the input and inference facts to be real.



Double clicking on the consequence axis

Determining the consequence of synthetic truth

Is it reversible?

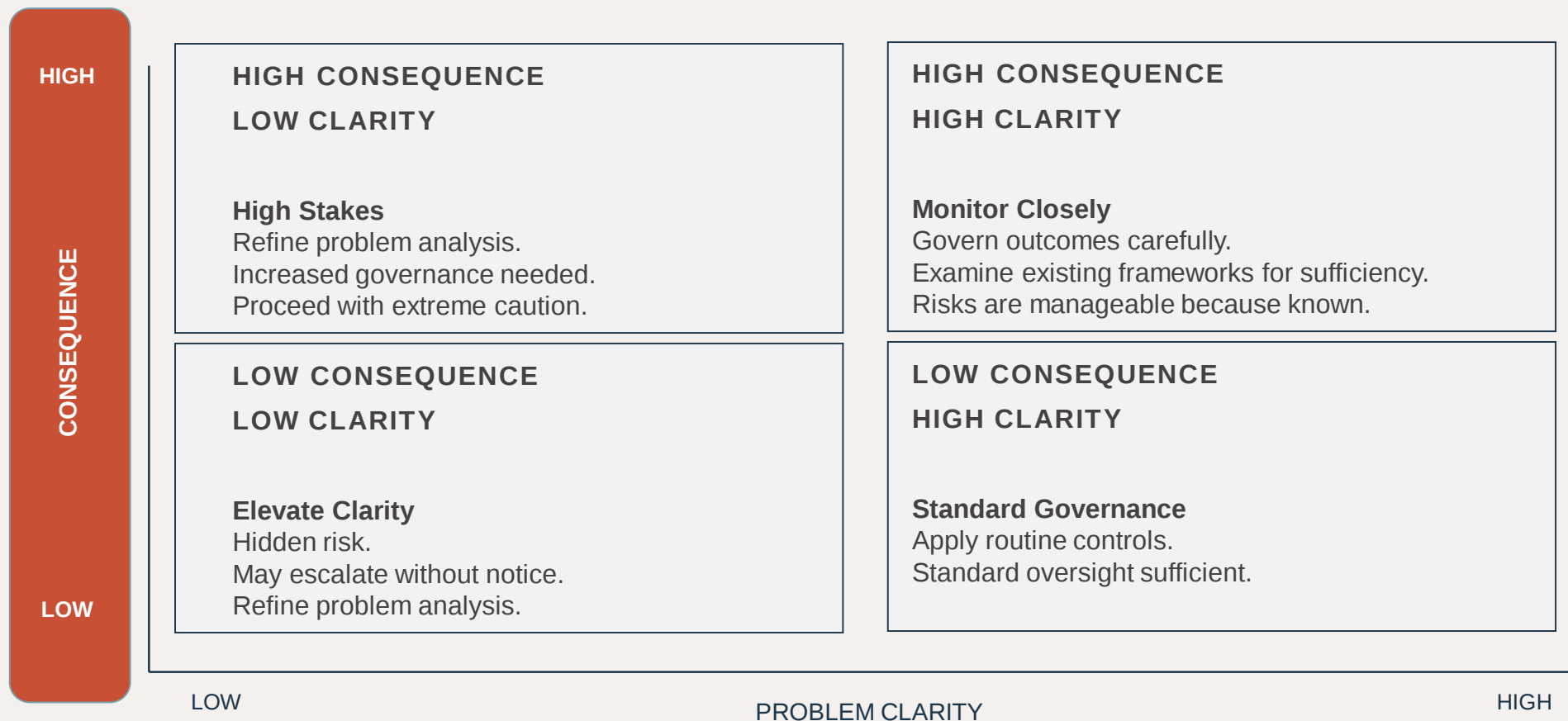
(can you change the outcome?)

Is it defensible?

(can you attribute the decision to its inputs?)

Is it evident?

(how do negative outcomes surface?)



The Legal Precedent is Being Set

Courts aren't punishing error; they're punishing unowned assertion.

Mata v. Avianca (2023): a lawyer filed a brief citing cases ChatGPT invented and was fined. The court faulted the failure to verify and standing behind the fakes, not the use of AI.

Johnson v. Dunn (2025): a major firm filed fabricated citations. The court signaled that fines alone are no longer deterring this.

Fletcher v. Experian (2026): sanctions at the appellate level over roughly 16 fabricated quotations. The person who signed owned the claim. Where it came from didn't matter.



Stop promising trust. Promise a trail.

The way to move toward confidently answering the board room question lies here.

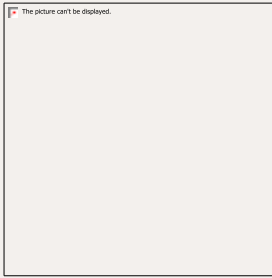
Move trust from the decision to the process:

- Grounded
- Traced
- Owned

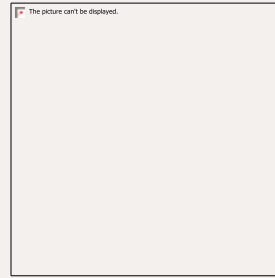


In practice: a credit-decline process flow

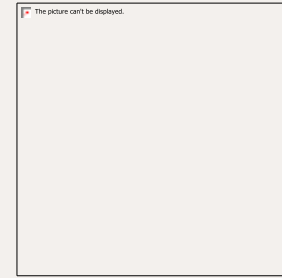
A bank uses AI to draft adverse-action letters, the notices that tell a customer why they were declined



Grounded: AI can only cite decline reasons from an approved list tied to the decision data, not free text it composes.



Traced: Every letter records the model version, the data fields used, and the template.



Owned: A named compliance lead approves the flow and owns the exception queue. *A letter that can't map to an approved reason doesn't go out. It routes to a person.*

Govern the process flow, and the decision

Creating Decision Governance that Scales

The instinct is to govern outputs: review answers, catch the bad ones. Outputs are infinite and growing, so that never scales.

Governing process flows is finite (recall the decision and inference plane from May)

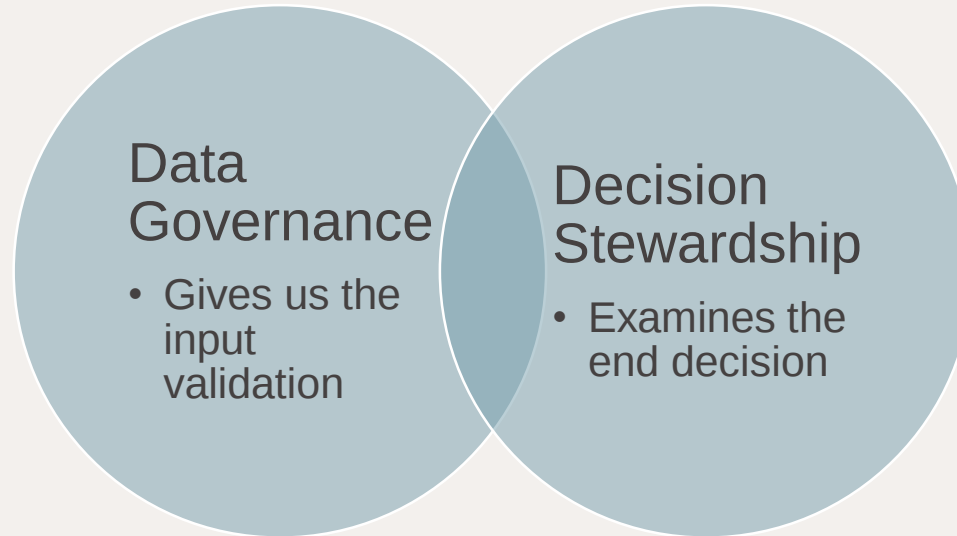
The consequential decision paths are few, and they stay few.

Design those few so grounded, traced, and owned are built in, and every output inherits them.

Decision Stewardship and the Pod

Extending the Scope

We've Established:



What's left is the decision making that happens inside the model, which can not be governed, so it must be attributed.

The Pod Evolved: The pod now must own the evidence of adherence to the flow. They work the exceptions. Their signal is the attribution gap: no source, no owner.

This is the scalable solution.

Accountability

A defensible record beats a trust score

- **Accountability means every consequential claim has an owner and an origin you can trace.**
- **When something's wrong, you catch it, you know where it came from, and you can fix it. Nothing fails silently.**
- **No system can guarantee truth. This makes errors visible and answerable instead of hidden. The judgment stays human.**



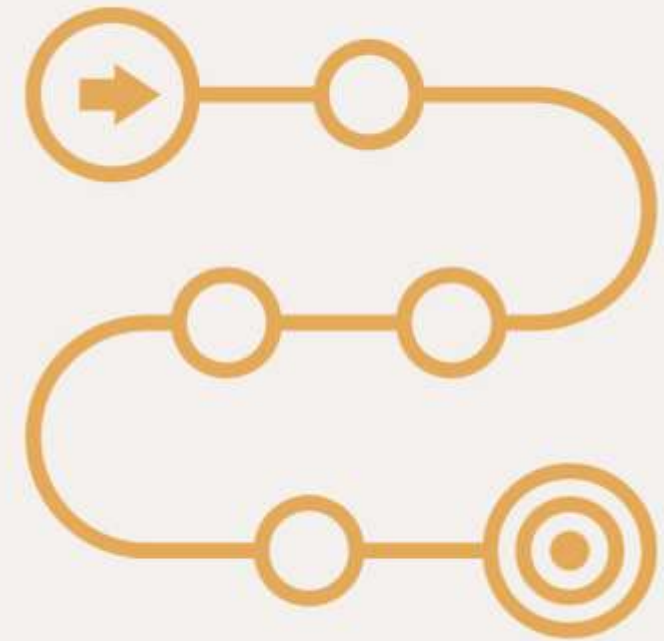
Proof beats a plan

One decision, with a documented path will anchor the journey

Governing every output is impossible. Governing one consequential process flow is attainable.

Pick the flow where a silent fabrication hurts most: hard to reverse, hard to check, slow to surface.

Make it grounded, traced, and owned with a named owner and a composable pod around it.



Now you have evidence for the board: here's one decision we can stand behind, and here's the record.

Welcome to the Board Room...

This is the question many of our clients are facing today

Do you *trust* what AI is telling us?

"I can't certify every answer is true. I can show you that every consequential decision is grounded, traced, and owned, and walk you through one."



Let's Stay Connected
Scan the QR Code



Kelle O'Neal | Founder & CEO
Lisa Wintrick | Executive Advisor

