FIRST SAN FRANCISCO PARTNERS

HUMAN OVERSIGHT IN AI

Designing Accountability for High Stakes Decisions

AIGOV | LEADING AI Governance

DATAVERSITY®

Welcome to Leading AI Governance – Session 4

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

Designing Accountability for High-Stakes Decisions

At the end of this session, participants will:

- ✔ Observe the evolution of risk as we move further into an agentic world
- ✔ Update the definition of high-stakes
- ✔ Confirm the role each level of the organization plays in accountable AI

Session roadmap

01 **Introduction**
As AI evolves, so does potential risk

02 **Moving Human-in-the-Loop from Singular Builds to a Living Network**
Failure of AI to scale isn't a technology problem – it's an accountability problem

03 **Exploring High Stakes (Risk)**
Redefining high stakes when the decision process is also the risk

04 **Exploring Accountability**
If the process is also the risk, accountability can't sit at the end. It must be built into how decisions get made – at every level.

05 **Closing**
Key takeaways

SECTION 01

Introduction

As AI evolves, so does potential risk

The nature of AI risk has evolved

Two cases. A decade apart. A widening accountability gap.

2014 – 2018 · Traditional ML

AMAZON AI RECRUITING TOOL

Internal system | High human oversight | Process redesign

AI ROLE	Scored and ranked candidates based on historical hiring data
DECISION TYPE	Redesign of an end-to-end hiring process
OVERSIGHT	Recruiters reviewed outputs; final decisions remained with humans
WHAT WENT WRONG	System penalized resumes containing "women" — bias encoded from historical data

CONSEQUENCE

Reputational damage · Tool scrapped

2026 · Agentic Software Solution

WORKDAY / AI RECRUITMENT SCREENING TOOL

Autonomous execution | No human override | Single deployment decision

AI ROLE	AI applicant screening module compares job posting to applicant's resume
DECISION TYPE	Support for accelerating the screening process for job applicants
OVERSIGHT	Lack of human-in-the-loop safeguards governing agentic behavior
WHAT WENT WRONG	Lawsuits allege AI screening bias with near real-time rejections proving no human-in-the-loop controls

CONSEQUENCE

TBD – litigation will determine if the software vendor, not the employer leveraging the software, could be liable for discrimination.

In both cases, **the problem was not the technology — it was the absence of accountability structures designed to match how the AI was actually operating.** As AI becomes more autonomous, that gap does not close on its own.

High stakes rules have changed

When scope of decision impact is decoupled from organizational role, the decision-making process itself poses a risk

LEGACY BUSINESS	AI ERA
RISK EXPOSURE LENS Outcome	**RISK EXPOSURE LENS** Outcome and business problem clarity
REACH / DECISION IMPACT SCOPE Correlated with organizational level and resources	**REACH / DECISION IMPACT SCOPE** Decoupled - individuals can have enterprise impact
BOUNDARY OF ORGANIZATIONAL KNOWLEDGE Absorbed through experience, culture is a guardrail	**BOUNDARY OF ORGANIZATIONAL KNOWLEDGE** AI knows only due to training and artifact availability
ACCOUNTABILITY TRACEABILITY Decision trail is human and legible (memo, meeting)	**ACCOUNTABILITY TRACEABILITY** AI's logic is opaque; requires deliberate design
SPEED OF CONSEQUENCE Emerge over time, humans are natural checkpoints	**SPEED OF CONSEQUENCE** Propagate at machine speed

SECTION 02

Moving Human-in-the-Loop from Singular Builds to a Living Network

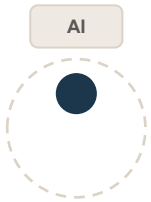
Failure of AI to scale isn’t a technology problem – it’s an accountability problem

Accountability is more than a moment

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

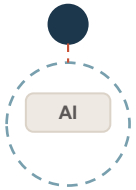
Human involvement is an accountability decision

As AI evolves in sophistication, how humans are positioned relative to it must be a deliberate, rather than default, choice



Human-in-the-loop

Human is an active participant in each cycle. Reviews, approves, or intervenes before AI proceeds.



Human-on-the-loop

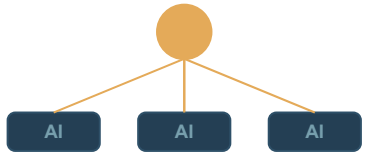
Human monitors from above. AI operates autonomously but human can intervene if needed.

Human-over-the-loop

Human sets policy and boundaries. AI operates within them. Oversight is structural, not operational.

Human-out-of-loop

AI operates fully autonomously. No human review or intervention at any stage. Highest risk.

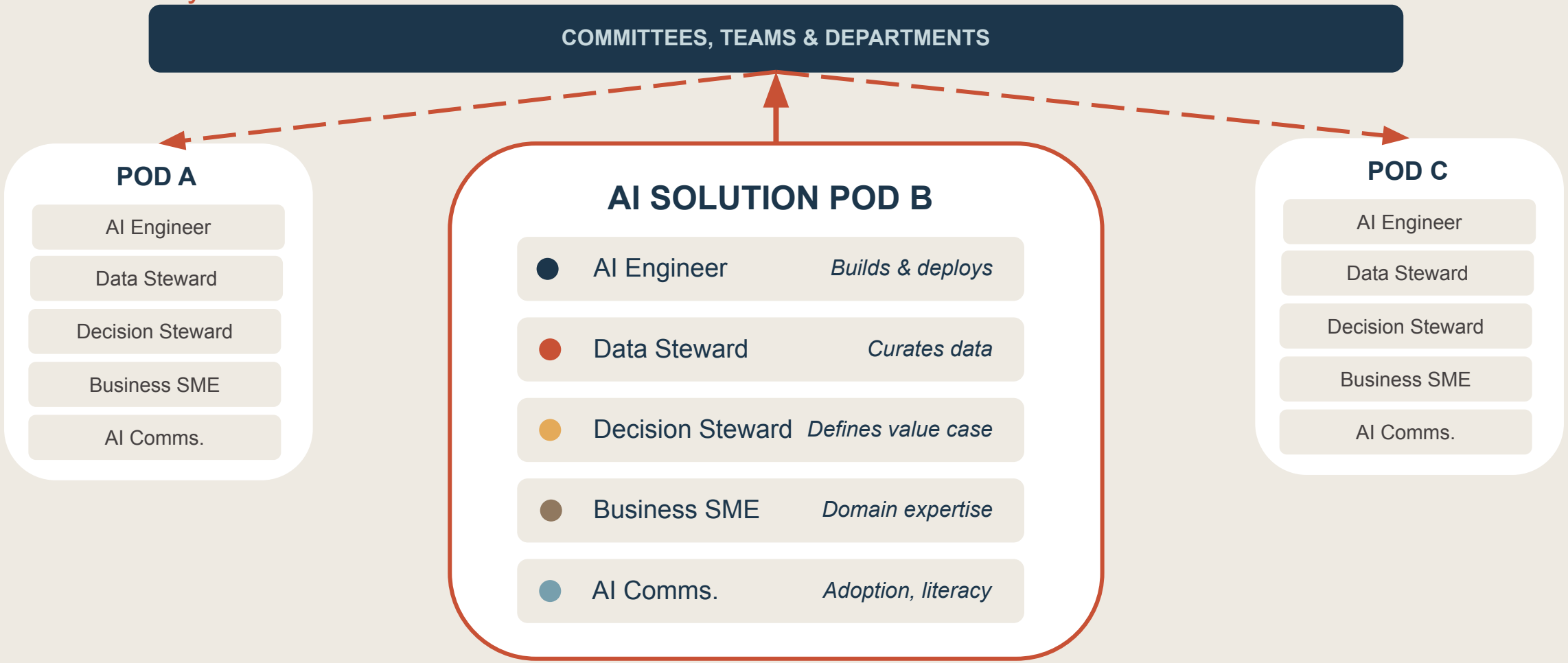


Human-in-the-lead

Human directs and orchestrates multiple AI agents. Strategic intent stays human. Execution is delegated.

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



Emerging AI responsibilities

Roles are becoming more focused, but requiring a broader knowledge of surrounding business

STRATEGIC ALIGNMENT

CHIEF AI OFFICER

Accountable for how an organization strategizes about, invests in, and derives value from AI.

CHIEF AI REVENUE OFFICER

Accountable for ensuring AI efforts are measurable and achieve stated goals.

GOVERNANCE

DECISION STEWARD

Aligns business cases with organizational value, ensuring feasibility, viability, and proper risk profiling.

SEMANTIC ARCHITECT

Translates the business into a network of AI-readable concepts to improve contextual accuracy.

BUSINESS TRANSFORMATION

AI CENTER OF EXCELLENCE LEAD

Builds shared capabilities and support tools that make AI efforts more likely to succeed.

AI LITERACY LEAD

Builds culturally relevant educational programs that close the knowledge gap between what employees understand about AI and what it can actually do.

PORTFOLIO

AI PORTFOLIO MANAGER

Oversees all AI initiatives to ensure strategic coherence.

AI DELIVERY MANAGER

Oversees execution when scope is constantly shifting.

SECTION 03

Exploring High Stakes (Risk)

Redefining high stakes when the decision process is also the risk

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.

Questions that expose high stakes decisions

"If we knew ____, we could do ____, which would make us better in this way: ____."

01 What problem are we actually solving?

Not what tool are we implementing - what business problem are we addressing?

- What is the revenue, cost, or risk result of (non)action?
- Are there intangible consequences that could also lead to this result? (e.g., reputational loss)

02 Who are we solving for? Who is indirectly impacted?

Direct stakeholders are obvious. Indirect stakeholders are where risk hides.

- Is the business process model detailed enough to identify hidden stakeholders?
- Are stakeholders represented in the training or fine-tuning data appropriately?
- Is anyone expected to make a decision based on this outcome? Will anyone be affected by that decision?

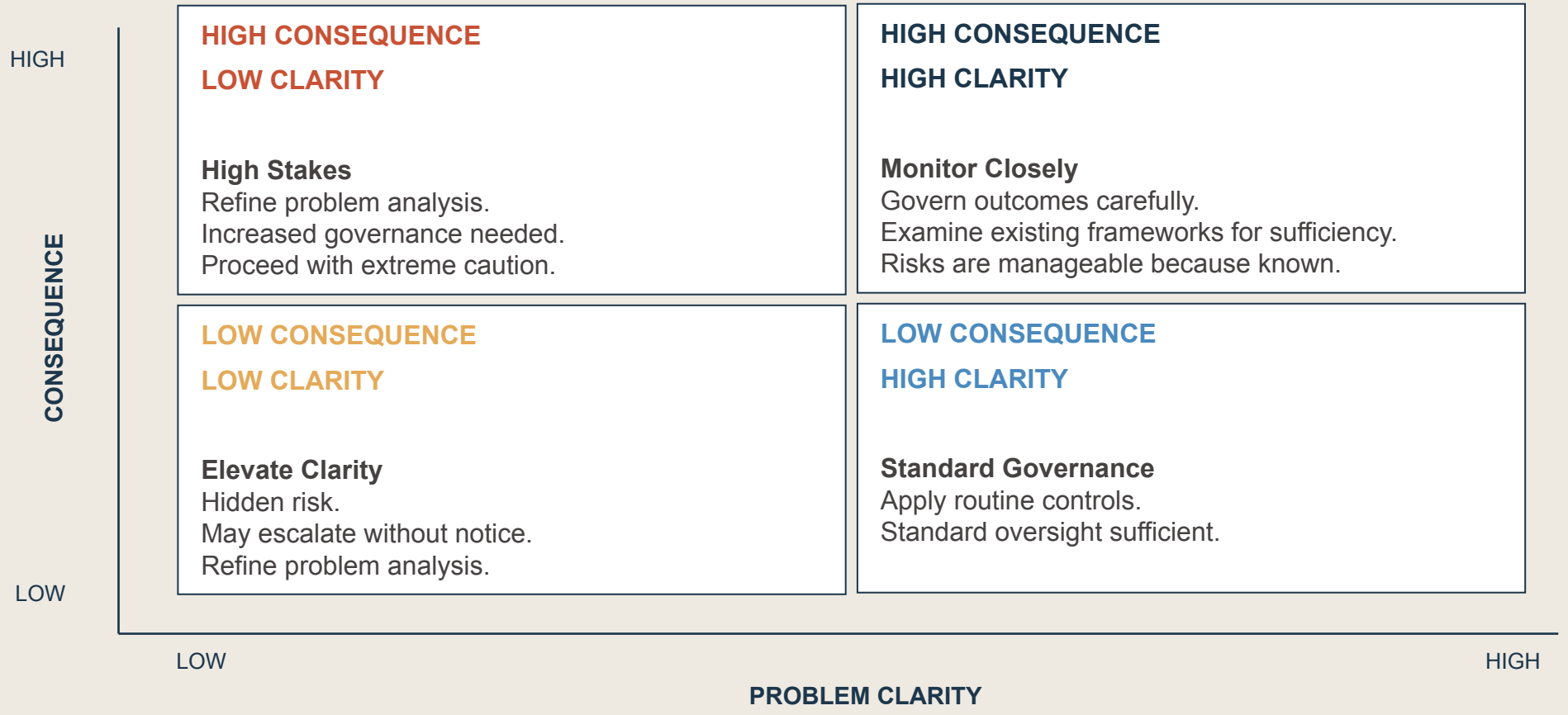
03 What should the output enable?

Define the intended outcome before evaluating the method.

- What are the quantitative and qualitative measures of success?
- Can these metrics be gamified in any way?
- How do we plan to monitor output "fitness" over time?

Proceeding with known risk levels

Consequence v. clarity matrix for decision-making



Redefining high stakes

Clarity of business problem is now a factor in determining outcome risk level.

TRADITIONAL DEFINITION

Used to describe a situation that has a lot of risk and in which someone is likely to either get or lose an advantage, a lot of money, etc.

EMERGING DEFINITION

A high-stakes decision is any decision where the consequences — intended or unintended — could materially affect people, operations, reputation, or compliance, and where the business problem clarity has not been rigorously established.

AI accountability starts with Decision Stewardship

If risk lives in the decision-making process, then governance must too

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.

HOW IT WORKS

Serves as the critical control point before execution

Establishes AI explainability, transparency, and accountability

Focuses on decisioning processes, including initial go / no-go and ongoing monitoring for “fitness”

Treats problem clarity as a measurable governance variable

Ensures key human-in-the-loop checkpoints occur

Ensures accountability structures are upheld during decisioning

Mitigates the gap AI creates when it operates on what it’s given, not what was intended

Provides proactive risk management by identifying and governing upstream decisioning

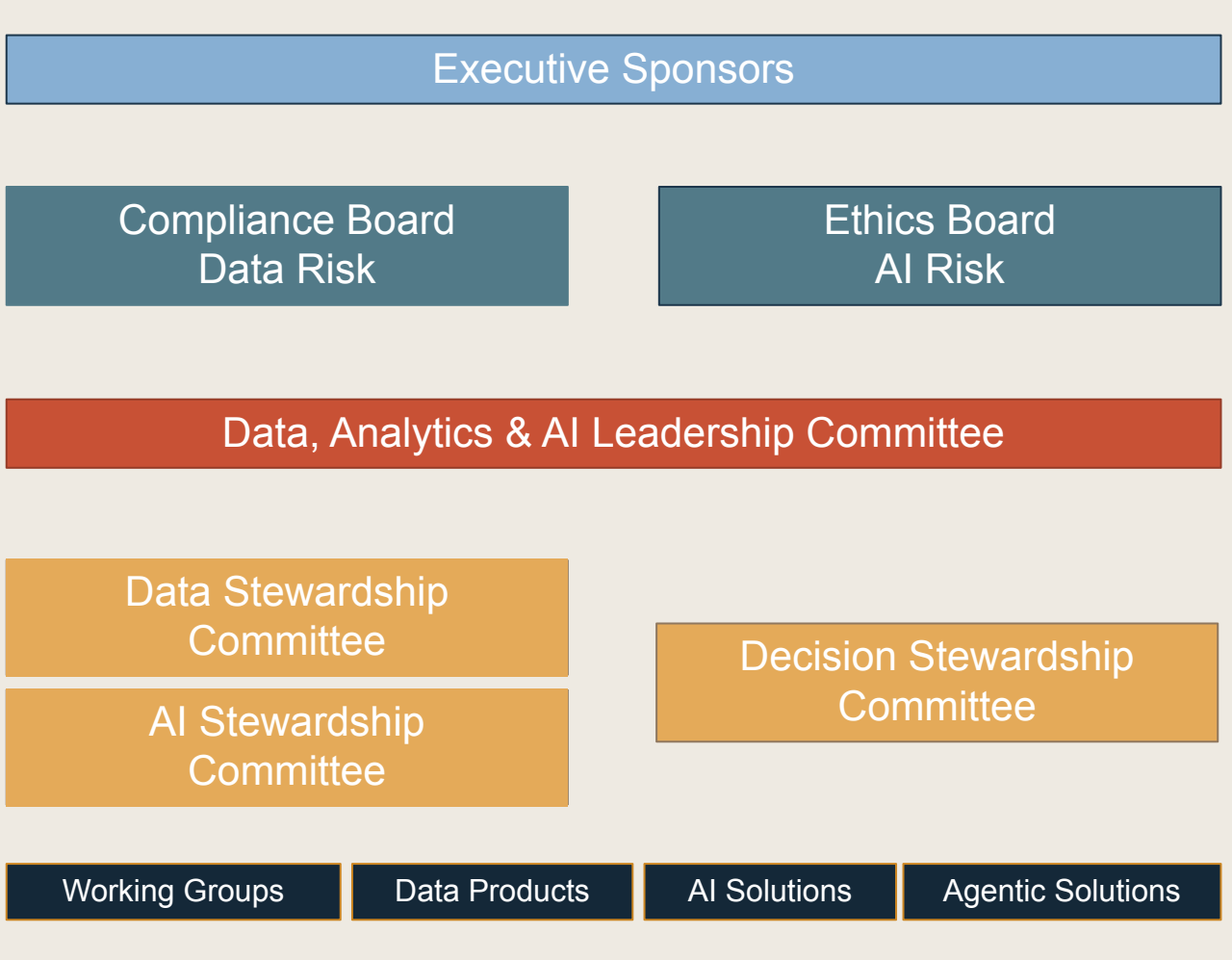
SECTION 04

Exploring Accountability

If the process is also the risk, accountability can’t sit at the end. It must be built into how decisions get made – at every level.

The Decision Framework

Details who is accountable for what types of decisions ***and*** provides a scaffolding for the accountability framework itself



ENTERPRISE LEVEL

Strategic direction, resource allocation

DOMAIN LEVEL

Risk oversight, domain-specific policies, cross-functional alignment

FUNCTIONAL LEVEL

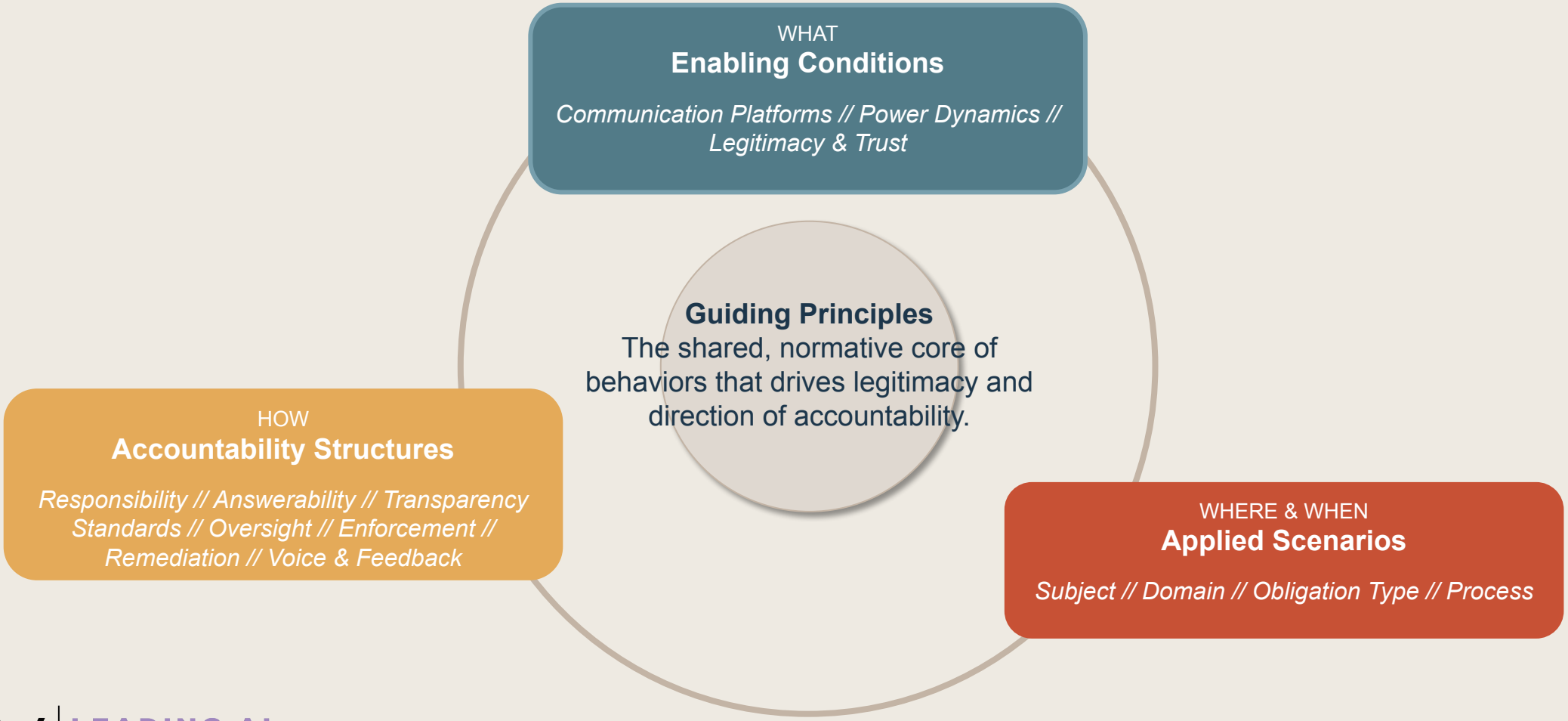
Departmental standards, value realization, compliance monitoring

OPERATIONAL LEVEL

Day-to-day execution, value production, issue escalation

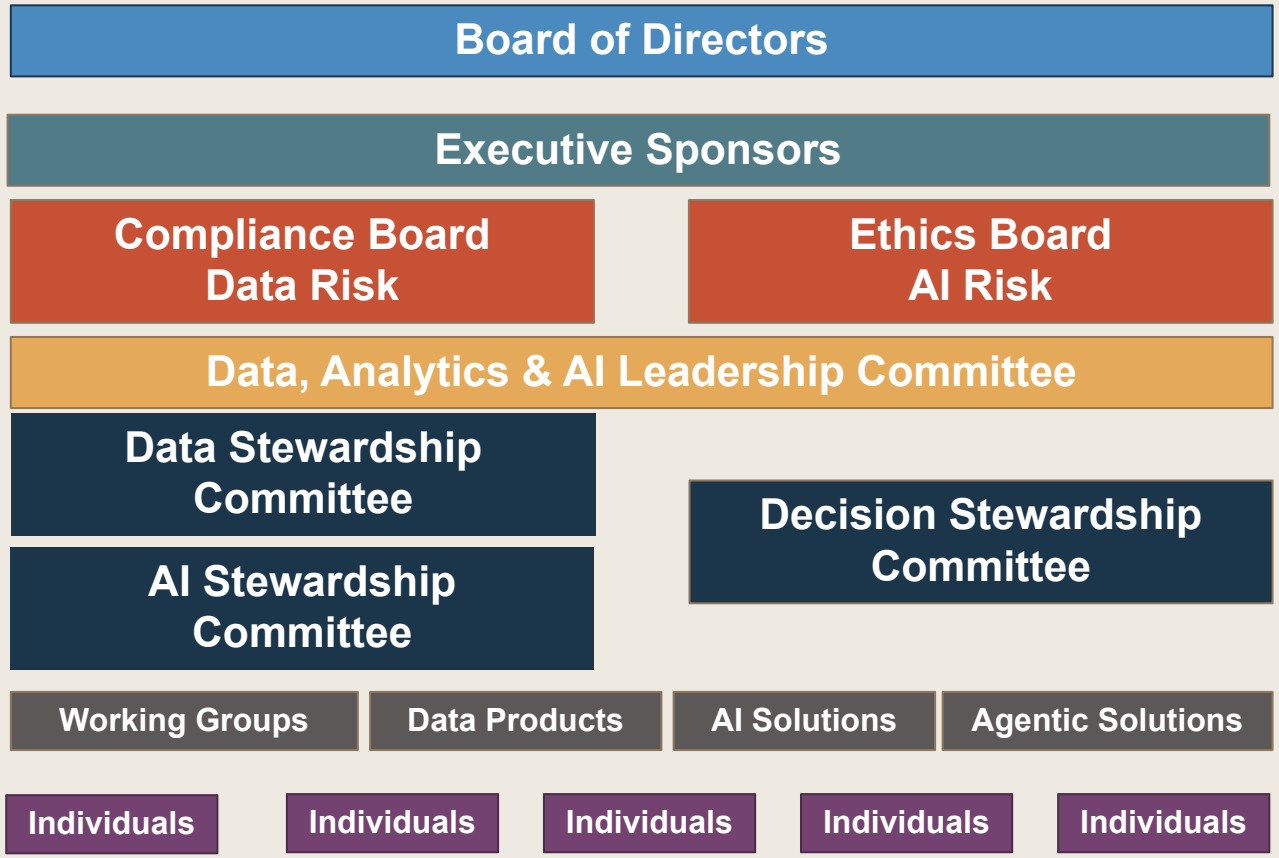
The “Accountability Orbit”

Without all components, an organization may establish control, but not true accountability. Accountability requires not just an assigned individual, but also, the conditions under which it is viewed as legitimate.



AI adds new levels of accountability

AI is now shaping strategy at the highest level, while being utilized at the individual contributor level



GOVERNANCE GAP 1

Reactive to proactive

Board-level AI oversight needs to shift from asking, “did we achieve what we said we would?” to “how is the market evolving and what does it mean for us?”.

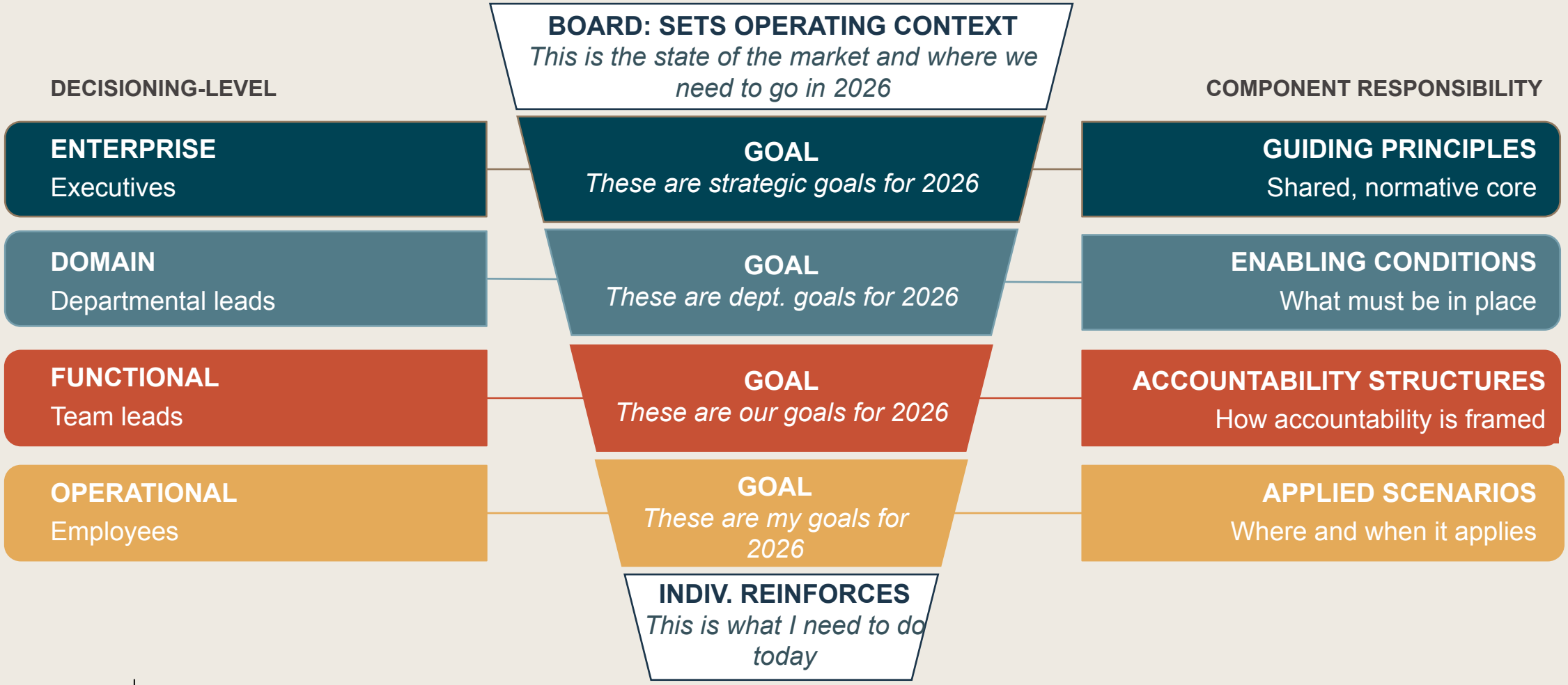
GOVERNANCE GAP 2

Operating with outsized reach

Individuals now have access to tools that execute at scale and speed that used to require significant resources. Further, the impact of their output can ripple far beyond traditional boundaries.

Accountability is upheld or undermined at all levels

Each organizational level is responsible for defining an accountability component ***and*** for reinforcing the entire framework in daily decisions.



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

SECTION 05

Closing

Key takeaways

Key Takeaways

1. High stakes is no longer just a property of the outcome – it's also a property of the business value case
2. Decision stewardship encompasses the accountability structures required to properly assess the business case risk
3. The business is a living process, but AI requires tangible artifacts; therefore, accountability must be codified to be applied
4. Humans still play a critical role in applying intuition as another control layer over AI
5. Individuals at all levels choose daily to reinforce or undermine the accountability structures

FIRST SAN FRANCISCO PARTNERS

Kelle O’Neal kelle@firstsanfranciscopartners.com

Stephanie Paradis stephanie@firstsanfranciscopartners.com