OPERATIONAL EXCELLENCE NEWSLETTER

The AI Governance Failure: Why Technology Isn't the Problem.

Moving beyond pilot hype to a disciplined adoption framework built around real-world stakes.

Corporate Taboos Series 3 · Part 1 · AI Governance · Process Improvement · Risk Management · Quality Systems

> The newsletter for leaders who prefer reliable process over flashy, uncontrolled technology.

The Promise vs. The Pitfall: Why AI Governance is the Real Job

We've all seen the presentations. They show sleek dashboards, glowing lines of code, and a promise that Artificial Intelligence will magically solve our yield issues or automate away our bottlenecks overnight. In those meetings, it sounds like a miracle—a way to leapfrog over years of incremental improvement in a single afternoon. But I want to talk about what happens when you actually bring that "miracle" onto the shop floor.

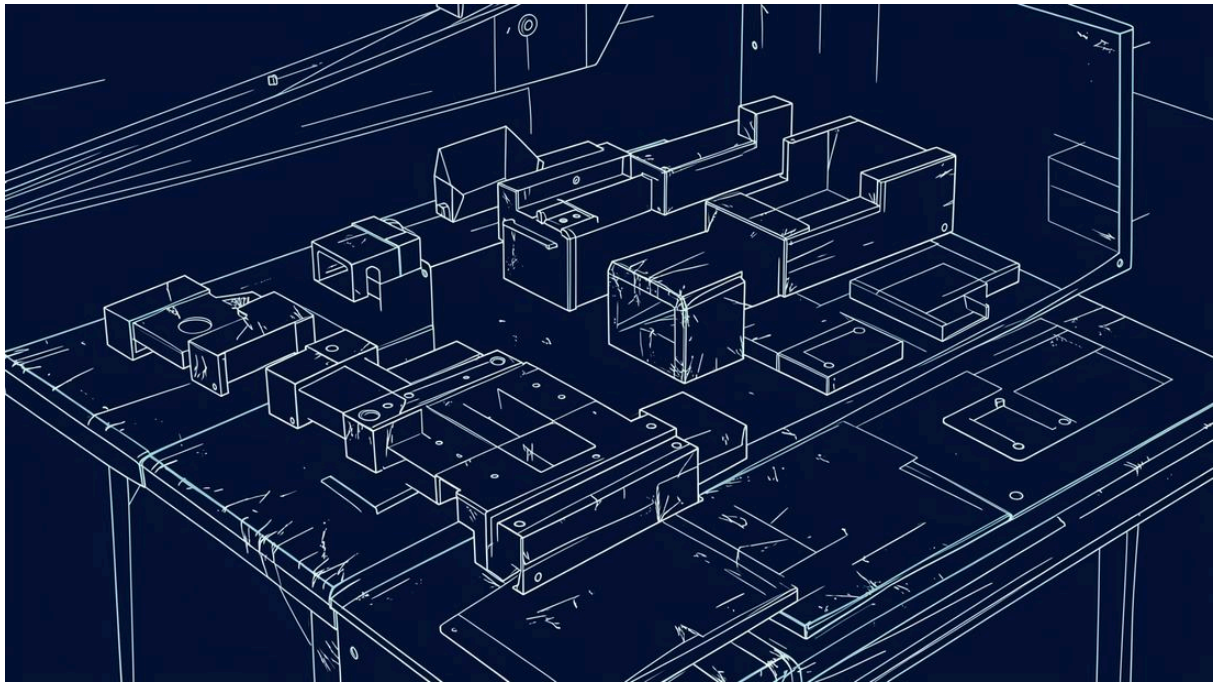
The mistake most leaders make is focusing entirely on the performance metrics of the model itself: accuracy rates, F1 scores, and processing speeds. Those are academic numbers. They tell you how well a machine can think in a vacuum, but they don't tell you if that machine can actually *work* in your facility.

AI governance isn't about choosing which algorithm is the smartest; it's about operational readiness and risk mapping. It is the discipline of figuring out what happens when the model gets it wrong. If an AI-driven vision system flags a part as "good" that should have been scrapped, who owns that error? If the

system slows down because of a network hiccup on the third shift, does the line stop or do we lose our first-pass yield?

Real governance is about building the guardrails around the technology so it doesn't become a runaway train. It’s moving from "Can this work?" to "How do we ensure this stays safe and useful when things go sideways?" We need to stop treating AI as a plug-and-play solution and start treating it as a new, complex piece of equipment that requires its own maintenance plan, safety protocols, and standard work.

Naming the Failure: The Illusion of Automated Control



I want to call this failure **The Illusion of Automated Control**. It’s the belief that because an algorithm is making a decision, the process has somehow become "self-managing." In reality, it often just hides the lack of control until it's too late.

I watched a pilot project go south at a facility that tried to automate quality inspections using a new deep-learning model. On paper, it was a triumph. The metrics looked great in the lab. But they didn't account for the "messy" reality of the floor—dust on the lenses, slight variations in ambient lighting during different shifts, and occasional material inconsistencies that weren't in the training data.

Because there wasn't a governance framework to define human intervention points, the operators were left guessing. When the model started making weird calls, they didn't have a clear "reaction plan" taped to the machine or an escalation path. They just let it run because "the computer said so," until a batch of sub-standard parts made it through to the next station.

The technology wasn't broken; the governance was missing. We mistook the speed of the software for the stability of the process. When you automate a decision without defining who owns the consequence of that decision, you haven't gained control—you’ve just outsourced your accountability to a black box.

Why This Keeps Happening: The Tech Hype Cycle vs. Process Ownership

Why do we keep falling into this trap? It usually comes down to a desire for quick wins over long-term process ownership. It is much easier (and looks better on a slide) to buy a piece of software than it is to redesign the underlying workflows that feed that software data.

We see teams looking for an external "fix" because they are tired of the hard work of cleaning up messy data or standardizing manual inputs. They want the shortcut. But AI doesn't fix a broken process; it just accelerates it—including its flaws. If your data lineage is a mess, your AI will just produce high-speed junk.

The Comfortable Rationalization	The Underlying Reality
"The software handles the complexity for us."	We have lost sight of the underlying logic and cannot intervene when it fails.
"It’s just a pilot, so there's no real risk yet."	Every pilot is a live test on your production environment with real-world consequences.
"The model is 'smart' enough to adjust itself."	A model that adjusts without human oversight will eventually drift into an unrecoverable state.
"We’ll figure out the rules after we see it working."	If you don't have the rules before you start, you aren't testing a solution; you're gambling on luck.

What It Costs When You Skip Governance: The Stakes Matter

In our world, skipping governance isn't just an administrative oversight—it’s a liability. We need to look at this through the lens of escalation logic.

The "small" cost is the time and effort it takes to do things right now: sitting down with your engineers to map out data flows, defining who has the authority to override an AI's decision, and documenting every step of the adoption. It feels slow. It’s uncomfortable because it requires admitting that we don't have all the answers yet.

The "large" cost is what happens when those guardrails aren't there. In high-stakes environments like healthcare manufacturing or critical infrastructure, a governance failure isn't just a line item on a scrap report. It’s a safety incident. It’s a recall that costs millions in logistics and destroys a brand’s reputation overnight.

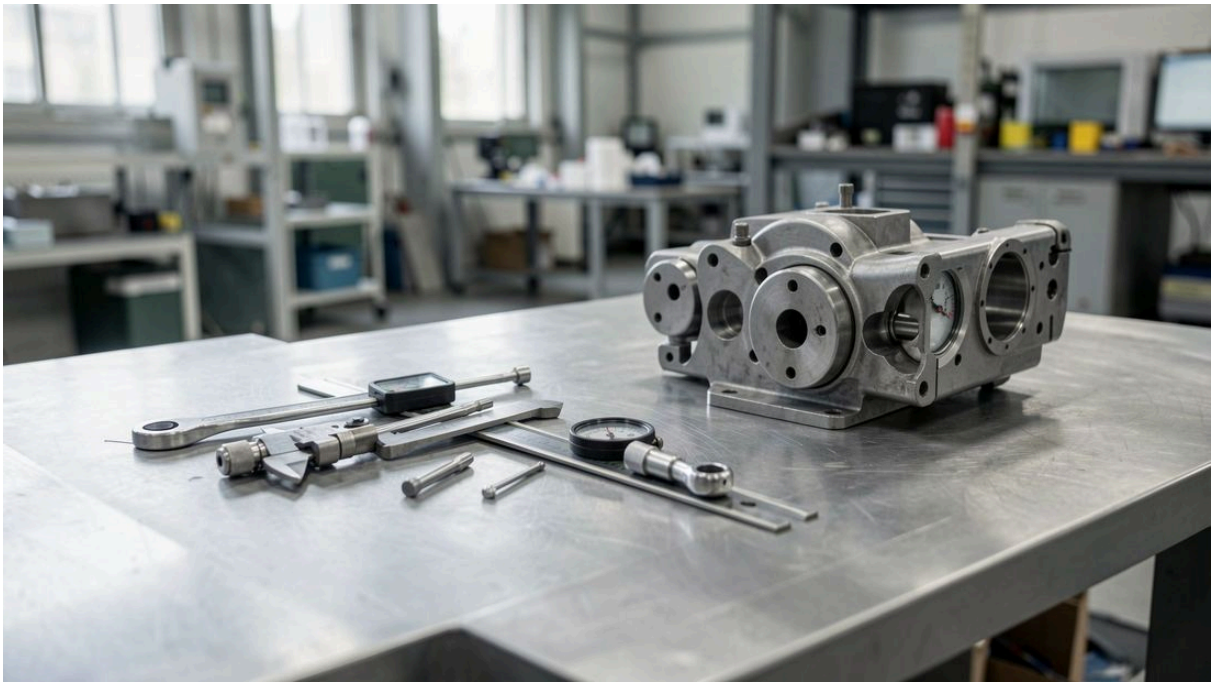
The avoided conversation today—the one about who is responsible when the AI fails—is uncomfortable. The delayed consequence of that silence is usually expensive, public, and much harder to recover from than a slightly slower pilot program. We choose between the friction of discipline now or the catastrophe of chaos later. I know which one I’d rather manage on the floor.

The Framework: Operationalizing AI Through Six Process Pillars

To move past the hype, we need a disciplined adoption framework. This isn't about "AI strategy"—it’s about operational pillars that must be satisfied before any model goes live.

1. **Data Integrity and Lineage:** You have to know exactly where your data comes from and who has touched it. If you can't trace the input, you can't trust the output. This means physical audits of the sensors and software feeds that feed the "brain."
2. **Model Transparency (Explainability):** If a machine flags a part as a reject, an operator should be able to see *why*. We need to move away from black boxes toward systems where we can point to the specific variable—torque, temperature, or time—that triggered the alert.
3. **Human-in-the-Loop Protocols:** This is your reaction plan. You must define exactly when a human takes over. What are the "red line" triggers? Who has the key to override the system, and what is the standard work for that override?
4. **Risk Stratification:** Not every AI application carries the same weight. A tool that optimizes the warehouse layout is a different beast than one that monitors a pressure valve on a volatile chemical line. Categorize your projects by risk before you invest a dime in development.
5. **Drift Monitoring:** Models "drift" just like machines do. They lose calibration as real-world conditions change. You need a scheduled maintenance plan for the model—periodic checks to ensure its outputs still align with your quality standards.
6. **Auditability and Traceability:** Can you show an auditor, three months from now, exactly what logic was being used on a specific Tuesday? If you can't produce that record, you haven't governed the process; you've just let it happen.

Making It Stick: Three Operational Checks You Can Run Tomorrow



You don’t need to wait for a committee meeting to start this work. You can begin assessing your AI readiness on your next Gemba walk with these three concrete checks:

1. The Paper Trail Trace

Pick one piece of data that feeds into an automated system. Physically follow it from the sensor or entry point to where it hits the screen. Is there a gap? Does someone have to manually "clean" that data before the AI sees it? If you find a manual workaround, that is your first governance failure.

2. The "What If" Drill

Find an operator who uses an automated tool or dashboard. Ask them: *"If this screen tells you something that looks wrong to your eyes, what do you do?"* If the answer is "I don't know" or "I have to call a supervisor," then your human-in-the-loop protocol isn't embedded in the work—it’s just an idea in someone’s head.

3. The Input Drift Check

Look at the hardware providing the data for your AI models (cameras, pressure gauges, flow meters). Are they calibrated? Is the lens clean? AI is only as good as its "eyes." If you haven't included these physical components in your standard maintenance schedule, your AI governance is incomplete. You can't have a high-tech solution built on low-tech neglect.

Call to Action

What are the highest-stakes decisions currently made by human judgment on your floor? Share what you think, or send us questions at newsletter@sixsigmakaizen.com.

Newsletter replies: newsletter@sixsigmakaizen.com · X: @kaizen_6sigma

Stay curious, stay improving,

David Rodgers

SixSigmaKaizen.com

References: Deming’s System of Profound Knowledge; DMAIC Methodology; Juran’s Quality Trilogy