WEST YOST

AI GUARDRAILS #5

/// BOARDS IN THE LOOP

A GOVERNANCE FIELD GUIDE FOR AGENTIC AI IN WATER OPERATIONS

Andy Bochman | Resilience Strategic Lead | West Yost

As agentic AI moves into operations, here's how utility Boards of Directors (and City Councils exercising utility oversight responsibilities) should come to understand their governance obligations. The previous brief in this series closed in a highly cautionary fashion, in essence saying only the most well-resourced water utilities should even consider experimenting with agents right now. Another said economic and other pressures would compel utilities to move faster and accept more risk than they otherwise might. Some will and some won't deploy, and hopefully many will wait, watch, and learn. But whichever camp your organization finds itself in, and especially if you're moving forward or planning to move forward with agentic AI soon, this brief focuses on a central governance question: What constitutes responsible board-level oversight?

INTRODUCTION

The Chief AI Officer at the SANS (SysAdmin, Audit, Network, and Security) Institute recently shared that detecting intrinsic agentic risks is not a cybersecurity function, that it "may fall to another job function entirely, maybe one that doesn't exist yet." As a Board member, your fiduciary role suggests it's not just a new job function at the operational level; it's a new oversight obligation at the governance level. It's one you should not simply delegate downwards and assume it's handled.

But where to start if you want to get up to speed? Fortunately, there are a number of authoritative sources, both from within and outside the water community that can get you what you need. The AI field is evolving rapidly so it's going to require a commitment to continual learning to maintain at least a minimal degree of fluency, and this brief will point you to resources to help you do just that.

On top of emerging and varying state-specific rules, expert guidance is emerging from three primary sources of which utility Board members should be cognizant: legal, federal, and corporate, and their guidance is beginning to converge.

I. THE LEGAL LOOK AT RISK MONITORING

Delaware's foundational Caremark doctrine establishes that Boards have a duty to make a good-faith effort to build and maintain a board-level system of monitoring and reporting critical operational risks. Increasingly, that obligation plausibly extends to AI deployment and oversight. As a recent Stanford Law article observes, "the question is not whether the Board understood the risk. It is whether the Board ensured it would be told about it"[1]. For water utility Boards, the more recent Marchand "mission-critical" line of cases may be especially important because it emphasizes heightened board oversight of risks central to public safety and core operations. The same reasoning Delaware courts applied to aircraft safety at Boeing could translate directly to water utility operations. [2]

II. THE USG AND AUSTRALIA'S FEDERAL GOVERNMENT WEIGH IN ON AGENTIC AI AND CRITICAL INFRASTRUCTURE.

The Cybersecurity and Infrastructure Security Agency (CISA) and the Australian Cyber Security Centre recently published guidance on the secure adoption of agentic AI. The guidance warns that, as these systems assume larger operational roles in critical infrastructure, they can expand attack surfaces, gain broader system access than intended, behave unpredictably, and make operational decisions harder to trace and verify. The guidance therefore recommends deploying agentic AI incrementally, beginning with clearly defined low-risk tasks and continuously assessing systems against evolving threat models. This is essentially Cyber-informed Engineering (CIE) logic applied to AI agents.[3]

III. THE CORPORATE GOVERNANCE ESTABLISHMENT WEIGHS IN.

KPMG launched its global AI Board Governance Principles in April 2026, noting that its Global AI Pulse Survey found that nearly three quarters of Boards have only limited AI expertise. Fortunately, in the same month, the World Economic Forum (WEF) published what many consider to be the single best piece on this topic: a board-level playbook for governing agentic AI that, more than any other source, identifies specific failure mechanisms. [4]

It includes a call to "design for legible friction and internalize the liability perimeter." In other words, slow things down and insist on fully understanding the stakes of failures, especially for high-consequence implementation and configuration decisions touching operations. You cannot provide responsible oversight if you can't explain how decisions were made both about and by agents. One way to exercise this role is to formally defend a single high-stakes decision made by an autonomous agent as if under legal scrutiny.
If leadership cannot clearly trace the decision back to defined objectives and human intent, the system should not be operating. You cannot buy an indemnity clause for a synthetic actor acting on your behalf. If the agent acts, a named executive must own the consequences.

CROSS-SECTOR RECOMMENDATIONS FOR YOUR AWARENESS

Members of the WEF drafted the above-mentioned playbook around a core argument:

"The modern boardroom is ... reallocating decision rights to autonomous systems such as AI agents, while retaining governance models built for human judgement."

That mismatch is the governance problem in a nutshell. The following statements from the playbook are particularly apt for the governance challenges you face, segmenting the problem into several constituent parts:

- ***"The true failure mode of an autonomous agent is rarely a breakdown. It is hyper-competence applied to a flawed metric."*** This is the single most important sentence for you and your fellow utility Board members to internalize. Our first "I am the Guardrails" brief touched on this topic, which we're calling intrinsic risk, noting that agentic AI systems in utility operations "… introduce authorized systems that can drift, hallucinate, degrade, and optimize toward unintended consequences — none of which register as anomalies under traditional cybersecurity monitoring."
- ***"Relying on static audits for an autonomous agent is like analyzing the trajectory of a bullet after it has struck the wall."*** What might have worked in the past for deterministic software applications or machine learning systems is no longer sufficient. If you decide to deploy agents, your utility's operations will require a different kind of monitoring, one that leverages process engineering skills and continuously monitors for issues.
- ***"You can outsource execution to a synthetic system, but not fiduciary duty."*** This one is self-explanatory, and a point that needs to be addressed well before putting agents anywhere near utility operations.

PRACTICAL ADVICE FOR WATER BOARD MEMBERS JUST GETTING STARTED

What follows is a maturity-based set of board obligations with respect to agentic AI systems in water utilities:

1. DUTY OF CARE NOW EXTENDS TO AI AGENTS IN OPERATIONS.

Water utility treatment integrity, distribution system pressure management, and SCADA security are unambiguously mission-critical. In August 2026, the EU AI Act requires deployers of high-risk AI systems to provide transparent disclosure to workers before implementation, meaningful human oversight with authority to intervene, continuous monitoring for discriminatory impacts, and detailed logging of AI decisions for at least six months.[5] It is very likely that global industrial systems software vendors will implement these capabilities in all their product offerings rather than maintaining separate versions for Europe and the rest of the world.

2. YOU DON'T NEED AI SKILLS; YOU DO NEED A BASIC UNDERSTANDING OF CAPABILITIES AND RISKS, AND RECURRING UPDATES ON BOTH.

The duty of care requires that directors be sufficiently informed about AI risks and exercise sound business judgment. It does not require technical expertise, but it does require, at a minimum, access to adequate information about how AI systems are being deployed and what controls exist. The Caremark obligation isn't to understand the model; it's to ensure you'll be told immediately when something significant occurs.[6]

3. ANTICIPATION IS KEY.

CISA's agentic AI guidance notes that organizations must anticipate what could go wrong, assess how agentic AI risk scenarios might affect operations, and establish ongoing visibility and assurance.[7] In the Operational Technology (OT) systems that run treatment plants and pump stations, the failure modes touch pressure, flow, chemical dosing, and public health. This is where CIE can play a very helpful role as a risk mitigator with or without agents in the mix.

4. CONSIDER EMPLOYING THESE WEF DIRECTIVES FOR YOUR BOARD:

- Force a Directors and Officers (D&O) liability stress test. Review D&O insurance coverage for exposure to agentic AI negligence, because many policies do not account for agent-driven risk and most Boards are likely materially underinsured.
- Execute a "synthetic subpoena" or a "synthetic administrative order" drill by requiring management to formally defend a single high-stakes decision made by an autonomous agent, as if under legal or regulatory scrutiny.

CONCLUSION

With the multitude of challenges already facing your utilities, you have more than enough on your plate already. However, we would be remiss if we did not place these emerging obligations regarding agentic AI on your radar. International standards organizations recommend keeping a skilled and attentive human engineer in the loop whenever agents are part of your operational decision-making process[8]. We think it best that you and your fellow Board members are firmly in the governance loop as well.

Andy Bochman is Resilience Strategic Lead at West Yost. He previously served as Senior Grid Strategist at Idaho National Laboratory, where he co-developed the Cyber-informed Engineering (CIE) methodology. He is the co-author of Countering Cyber Sabotage (CRC Press, 2021).

Comments and questions: resilience@westyost.com

Links

1. https://law.stanford.edu/2026/03/17/the-ungovernable-machine/
2. https://www.akingump.com/en/insights/articles/does-ai-care-about-caremark-applying-the-core-principles-of-corporate-governance-to-artificial-intelligence-integration
3. https://www.westyost.com/wp-content/uploads/2026/03/AI-Guardrails-3-Andy-Bochman-Engineering-the-Guardrails-West-Yost-26.pdf
4. https://www.weforum.org/stories/2026/04/board-playbook-governing-agentic-ai/
5. https://www.internationalsos.com/insights/redefining-corporate-responsibility-in-the-age-of-ai
6. https://agility-at-scale.com/ai/governance/board-oversight-of-ai-governance/
7. https://www.cyber.gov.au/business-government/secure-design/artificial-intelligence/careful-adoption-of-agentic-ai-services
8. https://www.automation.com/article/briefing-ai-risks-critical-infrastructure