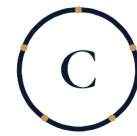


PUBLISHED BY



IN COLLABORATION WITH



Center for
AI Oversight

— THE ESSENTIALS OF AI OVERSIGHT, IN ONE SITTING

The Director's Briefing

*A Companion to AI Oversight:
The Private Director's Body of Knowledge*

Published by the **Private Directors Association**, developed in collaboration with the **Center for AI Oversight**.

FIRST EDITION, V.1.0 | JULY 2026

© 2026 Private Directors Association and the Center for AI Oversight

This briefing distills the essentials of **AI Oversight: The Private Director's Body of Knowledge** into a document a director can read in one sitting. It is a companion, not a substitute. The full Body of Knowledge carries the complete framework, the case analysis, the Director's Toolkit, and the reference appendices. Cross-references throughout this briefing point to the Sections, Tools, and Appendices where each topic is developed in full.

THE CASE FOR ACTION

I. Why This Cannot Wait

A Program, Not a Committee

The central argument of the Body of Knowledge is that effective AI oversight is not a committee. It is a program. A program is a formal management system through which the company directs, manages, and monitors its AI activities. The committee is the oversight mechanism; the program is what the committee oversees. The charter, policies, risk methodology, reporting protocols, escalation triggers, and monitoring mechanisms are the program. The question courts, investors, and insurers will increasingly ask is not whether the board has a committee, but whether a functioning system exists and whether the board is actively monitoring it.

The Velocity Gap

The urgency comes from what the Body of Knowledge calls the *Velocity Gap* (Section 1.2). AI capability develops on a curve that keeps accelerating. Board oversight, even at its best, is linear. The gap widens on its own, and a board that established adequate oversight two years ago may be structurally behind today without having made any errors. What closes the gap is not more frequent meetings. It is a change in the nature of oversight: pre-mapped decision frameworks that allow the board to respond with appropriate speed and judgment when a decision arrives.

Three Predictable Patterns

When oversight cannot keep pace, three predictable patterns emerge. The first is *Missed Opportunity*: the board, lacking a framework to evaluate AI initiatives quickly and confidently, defaults to caution, and the company falls behind one deferred decision at a time. The second is *Reckless Haste*: the board pressures management to move quickly without the governance infrastructure to manage the risks being introduced. The third is *Paralyzed Inaction*: the board defers every significant AI decision until someone else (regulators, industry bodies, advisors, or peers) defines the path first, mistaking the absence of an external answer for the absence of a decision to make. All three share a common root cause: the absence of a functioning oversight program. Governance is the accelerator. It is what allows the board to authorize management to move quickly, because the controls, reporting mechanisms, and escalation triggers are in place to support confident decision-making in both directions.

None of This Is Hypothetical

None of this is hypothetical. In November 2023, the board of the most consequential AI company in the world fired its chief executive and lost control of its own succession process within seventy-two hours, a velocity failure examined in Section 1.2. In February 2024, a Canadian tribunal held an airline liable for its chatbot's misstatements and rejected the argument that the AI was a separate legal entity, a precedent examined in Section 5.4. Federal enforcement against algorithmic discrimination in hiring is already producing consent decrees, examined in Section 5.3.

The Legal Foundation

The legal foundation is direct (Section 1.4). The Caremark line of Delaware cases establishes that directors have an affirmative duty to ensure that information and reporting systems exist for mission-critical risks, and *Marchand v. Barnhill* extended that standard in 2019 to a private, family-controlled company whose board had no reporting system for the operational risk that destroyed it. The standard has two prongs: directors must ensure that a system exists, and they must actively monitor it. As AI becomes central to how a private company operates, competes, and creates value, it is becoming a mission-critical risk. The absence of any oversight program is itself the failure.

This standard reaches private companies even where Caremark does not formally bind them. Investors conducting due diligence, lenders evaluating credit risk, insurers underwriting D&O coverage, and acquirers assessing valuation will all apply the same test. The practical question is not whether Caremark technically applies, but whether the board's oversight would withstand scrutiny from any of these audiences.

Builders and Buyers

One distinction shapes where oversight attention goes (Section 1.3). *Builders* develop proprietary AI models and face the full spectrum of development lifecycle risk. *Buyers*, the majority of private companies, purchase AI embedded in third-party software and platforms. Buyer boards are not overseeing model development. They are overseeing vendor selection, contract protections, data exposure, concentration risk, and business continuity. The scope of the duty is the same for both; the concentration of attention differs.

2. What the Program Must Contain

The most common governance failure is not the absence of a committee. It is the assumption that forming a committee constitutes governance, what the Body of Knowledge calls the *Committee Fallacy* (Section 3.1). A committee is a meeting. A program is a management system. If the board cannot answer who identifies issues, who decides what action is required, and who has the authority and resources to implement that action, the program does not have authority. It has ambiguity.

The Governance Charter

The charter is the program's constitutional document and its license to operate (Section 3.2). It establishes scope (which AI activities fall within the program's purview and what thresholds trigger formal review), authority (the mandate to require information from management, set tolerances, and escalate directly to the board), accountability (who owns the program at the management level and who stands in front of the board to answer for it), the expectation that initiatives are evaluated through both a risk lens and an opportunity lens, and the cross-functional coordination that AI risk demands. Without a documented mandate, authority is assumed rather than granted.

Tolerances and Decision Architecture

The director's role is not to approve individual AI projects. It is to ensure the program defines the tolerances within which AI decisions can be made and the criteria for when decisions must be escalated (Section 3.3). A tolerance is not a no. It is a decision point. Management moves quickly within defined parameters; activities outside those parameters are escalated to the appropriate decision-maker in a programmatic way. The Body of Knowledge calls this a decision architecture, and it is what allows the organization to move with confidence rather than in spite of governance.

Risk Appetite in Concrete Terms

Vague aspirations will not guide management; concrete tolerances will (Section 4.1). The risk appetite statement should address financial, operational, reputational, and legal dimensions in terms specific enough that management knows when to escalate and the board knows when to intervene: a defined ceiling on revenue exposure to any single AI dependency, a requirement that customer-data systems pass vendor due diligence before deployment, a standard for when customer-facing AI requires human oversight. Strategic risk belongs in the same statement, because in some circumstances the greater risk is not implementing AI too quickly, but too slowly.

The Crown Jewels Assessment

Not all AI activities carry equal risk. The program should map the company's most valuable assets (intellectual property, customer data, brand, key processes, competitive advantages) against the AI systems that interact with them (Section 4.2). The Crown Jewels that depend on AI are the mission-critical risks the Caremark framework describes. If the program cannot produce this mapping, the assessment has not been done.

Risk Intelligence

The AI landscape generates more noise than signal. Risk Intelligence is the defined process that filters the external landscape and surfaces what requires board awareness, with a named owner, documented escalation criteria, and a record of what surfaced and what the board did with it (Section 4.3). Each escalated item must produce a documented outcome, even if that outcome is a conscious decision not to act. Designed this way, Risk Intelligence is also the mechanism by which the board demonstrates the active monitoring that Caremark's second prong requires.

Auditing the Program, Not Just the AI

A governance structure on paper that nobody follows provides no protection; the Caremark line's later cases make the point explicitly (Section 3.4). The board should require periodic, independent evaluation of whether the program itself is functioning as designed, distinct from any technical audit of AI systems.

Escalation on Defined Triggers

The most common oversight failure is information known to management that never reaches the board (Section 7.1). Escalation cannot depend on management's judgment about what the board needs to know. It must be triggered by defined criteria: a risk outside a defined tolerance, a data breach involving AI systems or vendors, outputs producing material loss or regulatory exposure, discovery of a significant system deployed without governance review, a vendor failure affecting mission-critical operations, or a strategic opportunity requiring board-level evaluation. Opportunities belong in the escalation framework alongside risks. Defined protocols protect management as much as the board: they ensure no individual is later exposed for having sat on material information when a clear protocol existed for surfacing it.

Materiality and Disclosure

Private companies do not have SEC filing obligations, but they face functionally equivalent requirements from the parties that fund, insure, and ultimately acquire them: LP and investor agreements, credit covenants, sector regulation, and insurance notice provisions (Section 7.2). A material AI event that triggers an insurance notification but is not reported may void coverage precisely when coverage is needed most. Escalation thresholds should be calibrated to these existing obligations.

Crisis Readiness and the Kill Switch

Plans that have never been tested are assumptions (Section 7.3). The board should require evidence that management has practiced its response to an AI failure event through tabletop exercises, and that for every mission-critical AI system and autonomous agent, a tested kill switch exists: a specific person has the authority to shut it down, the technical means exists and has been tested, and the downstream consequences of the shutdown are understood. A kill switch that requires coordinating with a vendor under a support contract is not a kill switch.

3. What Directors Verify

The second half of the program is verification: confirming that the AI systems and vendors the company relies on are visible, trustworthy, secure, and economically sound. The director does not perform the verification. The director ensures the program does, and that its findings reach the board.

The AI Registry

You cannot govern what you cannot see. The Registry is a centralized inventory of every AI system in use across the organization, whether purchased, built, or adopted informally (Section 5.2). Shadow AI, the unauthorized tools employees adopt without governance review, is a primary source of unmeasured risk in private companies, and the Registry is the instrument that answers the question every board should already have asked: what AI is operating in this business that has not been formally reviewed? Many vendors are not the creators of the models they provide, so the Registry should also surface upstream dependencies. A company can believe it has diversified its AI ecosystem while multiple vendors rely on the same underlying provider.

Data Provenance and Exit Readiness

Who owns the data, and who owns the model, are valuation questions, not technical ones (Section 5.1). Undocumented provenance creates valuation haircuts in diligence, and the absence of governance artifacts is itself a finding that becomes a price negotiation. Vendor agreements should explicitly address data ownership, portability, deletion on termination, and whether the vendor can use the company's data to train its models. Change-of-control provisions in AI vendor contracts deserve specific attention well before a transaction is contemplated.

Human Oversight and Fairness

When an AI system makes a consequential mistake, the question will not simply be whether a human was present, but whether the organization had established the criteria for when human involvement was required and whether those humans were equipped to exercise meaningful judgment (Section 5.3). Oversight effectiveness degrades as automation becomes reliable, so the program should periodically confirm that the humans in the loop remain capable of challenging and overriding AI outputs. Systems that affect decisions about individuals carry the additional obligation of not producing discriminatory outcomes; enforcement is active, and the defense that the company merely used a vendor's tool is not protective.

The Agentic Frontier

Autonomous agents break the assumption that a human reviews every output before action (Section 5.4). An agent cannot be accountable; a person must be. The program should establish four conditions before agents are deployed at all: decision rights that classify each agent as authorized to act, to recommend, or to escalate; named ownership, with a business owner accountable for outcomes and a technical owner accountable for behavior; hard boundaries, the Agency Limits and Circuit Breakers enforced technically rather than documented in policy; and auditability, so that every consequential action can be reconstructed after the fact. These four conditions apply at every operating mode.

AI Security

AI introduces attack surfaces distinct from traditional IT: prompt injection, model manipulation, and data extraction through crafted queries (Section 5.5). For Buyer companies the critical question is whether the vendor's AI can be manipulated to expose the proprietary data entrusted to it. A vendor's AI security posture is the company's AI security posture if the company's data is in their system.

Vendor Concentration and Resilience

A vendor relationship is a commercial arrangement; a dependency is a structural vulnerability (Section 6.1). If three business functions depend on the same AI vendor, the company has a concentration risk that is invisible when each function is evaluated independently. Contracts should address data ownership, indemnification for AI-generated outputs, service levels that reflect AI-specific failure modes, the vendor's right to modify or retrain models, and liability allocation. And the resilience question must be answered before it is asked by events (Section 6.3): if the most critical AI vendor went down tomorrow, can the business continue to operate, and has that scenario actually been tested? If the answer is no, or unknown, that is itself a governance finding.

Unit Economics

AI initiatives should be held to the same financial discipline as any other investment (Section 6.2). The board should require management to demonstrate measurable business impact against the fully loaded cost of ownership, and to answer the forward-looking question as well: not just whether current AI investment is profitable, but whether the company is investing enough, in the right places, to maintain competitive position.

4. Right-Sizing the Program

The Accordion Principle

The framework is rigorous, not rigid. The Body of Knowledge calls this the *Accordion Principle* (Section 2.1): the program expands or contracts based on the organization's AI risk exposure, size, and complexity. What remains constant are the oversight functions. What varies is the formality, depth, and frequency with which they are executed. A five-person company and a five-hundred-person company need different levels of formality, but both need a functioning system.

Three Operating Modes

Three operating modes describe most private companies (Section 2.2). *Founder-led* companies, where authority is concentrated and AI exposure is limited, may need only a documented policy, an inventory of AI tools, and a quarterly conversation at the board level. *Growth-stage* companies, where AI use is expanding and the board is separating from day-to-day management, need a governance charter, defined risk appetite, and standing board reporting. *Institutional* companies, where the AI portfolio is complex and regulatory exposure is significant, need a full program with committee oversight, independent assurance, and formal escalation protocols.

Mismatches Are the Finding

Most organizations will not fit neatly into a single mode. A company may have Institutional-level AI exposure but a Founder-led governance structure. That mismatch is itself a finding, and closing it is the purpose of the work the Body of Knowledge describes. The Governance Maturity Diagnostic (Tool B) locates the organization across five dimensions in roughly fifteen minutes; the mismatches across dimensions are the most important output, not the absolute mode.

5. The Fifteen Questions

The following questions are reproduced in full from Tool A of the Body of Knowledge, *Board-Level AI Oversight Questions*. They are designed for the next board meeting. They are not a test of technical knowledge. They are a test of whether the oversight program is functioning. A board that cannot answer them should treat the gap as the most important finding from the conversation. The questions are grouped by the Section of the Body of Knowledge they draw on, so directors can sequence them across multiple meetings. Three to five at a time is enough to surface whether the program is real.

SECTION 1. THE FIDUCIARY PIVOT

- 1 Does the board have a credible view of the company's AI strategy, the AI footprint that already exists in the business, and the gap between what the company faces and what the board currently sees?
- 2 Are we a Builder, a Buyer, or a hybrid, and is our oversight attention allocated accordingly?
- 3 If a Caremark-style inquiry happened tomorrow, could we produce evidence that a functioning oversight system exists and that we are actively monitoring it?

SECTION 3. PROGRAM AUTHORITY AND ACCOUNTABILITY

- 1 Does the governance charter specifically mention AI risk, and can it be produced on demand?
- 2 Who is the named individual accountable for the program's effectiveness, and does every director know who that person is?
- 3 When was the last independent evaluation of whether the program is functioning as designed, distinct from any technical audit of AI systems?

SECTION 4. RISK APPETITE AND TOLERANCES

- 1 What is our risk appetite for AI, expressed in terms specific enough that management knows when to escalate and we know when to intervene?
- 2 Which of the company's Crown Jewels depend on AI, and is the oversight proportionate to what's at stake if any of them fail?
- 3 Who owns Risk Intelligence in our program, and what are the pre-defined triggers that bring something to the board rather than leaving it in management's inbox?

SECTION 5. ASSET INTEGRITY

- 1 Does a comprehensive AI Registry exist, is it current, and does it include the systems employees adopted without going through governance?
- 2 For autonomous agents operating in our business, do we have a named business owner and a named technical owner for each one, and can every consequential action be reconstructed after the fact?
- 3 Where our data sits in a vendor's system, do we understand what security testing has been done, what the vendor's own AI governance looks like, and what happens to our data if the relationship ends?

SECTION 6. VENDOR RISK AND STRATEGIC SOUNDNESS

- 1 Where do we have vendor concentration we may not see, and do the contracts with those vendors address AI-specific risks?
- 2 Are our AI investments generating measurable return, and are we investing enough in the right places to maintain competitive position?

SECTION 7. ESCALATION AND CRISIS READINESS

- 1 For every mission-critical AI system and autonomous agent, does a tested kill switch exist, and does a named person have the authority and the means to use it?

6. Where to Begin, and Where to Go Deeper

Where to Begin

Three actions, in order, will start closing the gap between most private companies and the standard the Body of Knowledge describes. First, complete the Governance Maturity Diagnostic (Tool B). It takes fifteen minutes and locates the organization honestly. Second, identify who at the management level is accountable for the AI oversight program; if the answer is unclear, the program does not have authority, and the Board Mandate Matrix (Tool C) resolves the ambiguity. Third, put AI oversight on the next board meeting agenda as a substantive item, not an information update, using the fifteen questions above to structure the discussion, and document the discussion in the minutes. That single action begins building the record that Caremark requires.

Where to Go Deeper

The full Body of Knowledge develops everything this briefing has compressed. Part I (The Board Briefing) carries the complete case for action, including the Velocity Gap, the Builder and Buyer agendas, the Caremark case line, and the Case in Point studies of OpenAI, iTutorGroup, and Air Canada. Part II (The Oversight Blueprint) details the charter, policy framework, risk appetite, Crown Jewels Assessment, and Risk Intelligence design. Part III (Asset Integrity and Operations) covers data governance, the AI Registry, human oversight, the agentic frontier, AI security, vendor risk, unit economics, and operational resilience. Part IV (Escalation and Disclosure) sets out escalation protocols, materiality, and crisis readiness. The closing statement, The Standard, states the director's responsibility in a single page.

The Director's Toolkit (Part V) contains the four working artifacts: the fifteen questions reproduced above (Tool A), the Governance Maturity Diagnostic (Tool B), the Board Mandate Matrix (Tool C), and the Vendor AI Due Diligence Questionnaire (Tool D). Appendix A provides AI fundamentals for directors without a technical background. Appendix B presents the AI Oversight Program, the illustrative operating model developed and maintained by the Center for AI Oversight, with a concordance mapping each Section to the Program's five Pillars and fourteen Domains. Appendix C provides third-party contract considerations, and the Glossary defines the working vocabulary with citations to authoritative sources.

The complete *AI Oversight: The Private Director's Body of Knowledge* is available from the Private Directors Association (privatedirectorsassociation.org). The AI Oversight Program presented in its Appendix B is maintained by the Center for AI Oversight and is updated regularly to reflect changes in regulation, technical standards, and case law; the current version and supporting materials are available at cfaio.org.

Notice and Disclaimer

This briefing is published by the Private Directors Association in collaboration with the Center for AI Oversight. It is intended as a strategic guide for private company directors. It is not, and should not be construed as, legal, accounting, tax, investment, or professional advice.

Directors, officers, and other readers who face specific questions about their fiduciary duties, regulatory obligations, or the application of AI to their organization's particular circumstances should consult qualified counsel and other appropriate professional advisors. Nothing in this document creates an attorney-client relationship between the reader and any contributor.

References to case law, statutes, regulations, technical standards, and industry frameworks are accurate to the best of the authors' knowledge as of the publication date. Law, regulation, and the AI landscape continue to evolve. Readers should verify current authority before relying on any specific reference for action.

The illustrative AI Oversight Program referenced in this briefing is proprietary intellectual property of the Center for AI Oversight and is presented, with its terms of use, in Appendix B of the Body of Knowledge.

Errors and omissions are the responsibility of the authors. The Private Directors Association and the Center for AI Oversight disclaim any liability for actions taken or not taken in reliance on the content of this document. Comments, corrections, and suggestions for future editions are welcome and should be directed to the PDA.

Lead Author: Brian Allen, Executive Director, Center for AI Oversight.

This briefing is derived from the First Edition (V.1.0, July 2026) of the Body of Knowledge and is updated on the same cycle. © 2026 Private Directors Association and the Center for AI Oversight.