SI NEWS BUSINESS BRIEFINGS

AI Governance for Corporate Boards

A Report on Stakeholder Value in the AI Paradigm

Brian R. Miller

A PUBLICATION OF SI News™

MAY 2026 · SI NEWS EDITORIAL · si-news.ai

SI NEWS BUSINESS BRIEFINGS

AI Governance for Corporate Boards

A board-level governance framework for the AI paradigm — written for the smart-but-non-technical director.

AUTHOR

Brian R. Miller — Chief Information Security Officer of a major Healthcare Payer/Insurer with two decades of board-facing technology and risk leadership across regulated industries.

EDITORIAL

Produced under the SI News Business Briefings report standard. The report is editorial work, not vendor-sponsored. No specific AI product, vendor, or consulting firm is recommended within these pages.

DISTRIBUTION

Published by SI News™ as a Business Briefing. Available to all readers under the SI News reader-tier model. May be cited with attribution. The report is a v0 first edition; reader comments and critique are welcomed and will inform subsequent versions.

LIMITATIONS

This report is not legal, accounting, or regulatory advice. Directors with specific compliance obligations should consult counsel in their jurisdiction. See **Appendix B** for the full editorial integrity statement, source list, and known limitations.

PUBLICATION CODE · EDITION

SI-NEWS-AIGOV-2026-01 · First Edition · May 2026 · Digital online edition

FOREWORD

A note from the author

I have spent the last twenty-plus years as a security and technology executive, much of that time reporting to corporate boards on technology and risk. I have watched four large technology paradigm shifts arrive in the corner office: the late-internet build-out, the mobile revolution, the cloud transition, and now generative and agentic AI. Each of those transitions changed what boards needed to know, what they needed to ask, and where the failures that mattered actually originated.

The first three followed a recognizable pattern. Management led the technical conversation. The board's job was to understand the strategy at a level deep enough to test it, to ensure the right people were doing the work, and to make sure the company did not lose more than it gained. The technology was a tool. The strategy was the strategy.

AI breaks that pattern. The technology *is* the strategy in a way the prior three transitions were not, because AI does not simply automate or accelerate existing work — it shifts what the work is, who does it, and what answers the company can give to its stakeholders. A board that treats AI as a procurement decision will fail. So will a board that treats it as a research project. The job sits somewhere between, and it sits in a place most boards do not yet have the muscle memory for.

This report is written for the board that knows the muscle has to grow and is looking for a place to start. It assumes a reader who is a smart generalist — accomplished in a domain, accountable to shareholders, attentive to the broader stakeholder picture, but not technically expert in AI. It defines its terms inline. It cites primary sources. It is willing to be wrong on the things that are still genuinely uncertain.

A word about the editorial commitment behind the writing. This report sits in a moral-realist tradition — the conviction that obligations to employees, customers, shareholders, and broader society are real and binding, not optional, and not a matter of branding. That tradition has names attached to it the reader will recognize: C.S. Lewis on the universal moral law, Francis

Schaeffer on the necessity of grounded presuppositions, Timothy Keller on cultural engagement that neither withdraws nor capitulates. The report does not preach those names. But it does carry the conviction that they ground: that authentic commitment to stakeholders is costly, that costless virtue is indistinguishable from marketing, and that boards who treat AI governance as a brand exercise will produce neither good governance nor good brand.

> *By aligning economic and social outcomes, a company can maximize value for all stakeholders. Driving that kind of broad value requires real strategy and governance work — not parroting the issue of the day.*
>
> — The thesis of this report

That thesis runs through every Part. The board's job in this paradigm is to build the serious governance machinery the alignment requires, and to keep pivoting that machinery as lessons land.

That is the work this report tries to make legible.

— Brian R. Miller
Synthetic Insights LLC · May 2026

ON METHODOLOGY

A note on how this report was produced, because the question is fair and the answer matters. This work draws on a body of research developed across Synthetic Insights over several years, with corroborating practitioner input from senior security and technology leaders. The substrate spans three knowledge bases that reinforce one another. Our **research-and-development repository** captures findings from sustained work in AI architecture, agentic systems, and applied ethics — including the patent-pending **Ethics as Infrastructure** approach, which treats moral reasoning as a foundational architectural layer rather than a compliance check added after the system is built. Our **editorial knowledge base** curates and

contextualizes primary research, regulatory documents, peer-reviewed work, and the analytical writing that informs how a governance question should be framed. Our **practitioner-voice knowledge base** captures the judgments and case-level pattern recognition of senior security and technology leaders who have governed through prior paradigm shifts. Research grounds the framing; editorial discipline curates and integrates it; practitioner voice tests it against the realities of board-level work.

The production process itself is **AI-augmented**. We use AI carefully and transparently to accelerate the work of pulling together source material, drafting structural scaffolds, and preparing the result for publication — a process that has compressed what would have been a months-long traditional research-and-writing cycle into weeks without sacrificing rigor. Every quantitative claim is cited to a primary source. Every chapter is reviewed and edited by humans whose judgment, voice, and accountability stand behind the final text. The technology is a force-multiplier on the production layer; the editorial, intellectual, and moral commitments behind the work remain ours. In short, we use the technology we cover — with the same governance discipline this report recommends to its readers.

Contents

What is inside

CLOSING

EXECUTIVE SUMMARY

Seven findings, seven actions, ninety days

This report is written for the board member who governs in the AI paradigm and is not yet a technical expert in AI. It treats AI as a paradigm shift in the Schumpeter / Carlota Perez sense — the latest after railroads, electricity, IT and the internet, and mobile and cloud — and reasons forward from what boards learned and failed to learn across those prior transitions.

4

GOVERNANCE FRAMEWORKS THE BOARD SHOULD TRIANGULATE: NIST AI RMF, EU AI ACT, ISO 42001, NACD

7%

EU AI ACT MAX PENALTY AS SHARE OF GLOBAL ANNUAL TURNOVER — LARGER THAN GDPR'S 4%

6

DOCUMENTED PUBLIC AI GOVERNANCE FAILURES REFERENCED AS CAUTIONARY CASE STUDIES

90 DAYS

FROM "AI ON THE AGENDA" TO "GOVERNANCE MACHINERY OPERATIONAL" PER THE CHAIR'S PLAYBOOK

The thesis

THE BINDING CONVICTION

By aligning economic and social outcomes, a company can maximize value for all stakeholders — **employees, customers, shareholders, and broader society**. The alignment requires real strategy and governance work, not parroting the issue of the day. The board's job in the AI paradigm is to build serious governance machinery that pivots as lessons land.

What this report finds

FINDING 1 · WRONG FRAME

The AI economic cycle is in the **late installation phase**. Boards governing AI as a procurement decision or as a research project are operating from the wrong frame for the moment. The deployment phase begins in earnest over the next three to five years.

FINDING 2 · CATEGORICAL FAILURES

The largest governance failures so far are categorical, not technical. The Air Canada chatbot, the Samsung data exfiltration, the Amazon recruiting tool, the Apple Card credit disparity, the UnitedHealth claims algorithm, the Uber autonomous-vehicle fatality — each was a failure of **operational discipline**, not a failure of available frameworks.

FINDING 3 · FRAMEWORKS COMPLEMENT

The four primary governance frameworks (NIST AI RMF, EU AI Act, ISO 42001, NACD guidance) are **complementary, not competing**. A serious program triangulates across all four. Programs that adopt only one are incomplete.

FINDING 4 · WORKFORCE IS THE TEST

Workforce navigation is where the stakeholder thesis is most directly tested. Companies running a **substitution-led** AI workforce transition face the largest medium- and long-run risk — talent pipeline collapse, reputational deterioration, regulatory and political backlash — despite the short-run margin lift.

FINDING 5 · STRUCTURE MATTERS

The board's structural choices about its own work are the leading indicator of governance success. Boards that have decisively allocated AI oversight to a committee, updated charters, developed director AI literacy, and required structured AI reporting from management are visibly ahead. The gap is widening.

FINDING 6 · THE COSTLESS-VIRTUE TRAP

Costless commitments are the dominant governance failure mode — and they are diagnosable. **The test:** ask of each AI-related stakeholder commitment, "what specific decision would the company make differently if this is real, and which of those decisions has already been made?" Commitments that fail the test are rhetoric.

FINDING 7 · DESIGN FOR PIVOT

The governance machinery built now **must pivot** as lessons land. Frameworks, metrics, committee structures, management mechanisms, literacy investments — all will need substantial revision by 2028 and again by 2031. Boards that design for permanence are designing for obsolescence.

What this report recommends

The hundred-plus specific recommendations in the body of the report reduce to seven the board should treat as non-negotiable.

#	Action	What it means in practice
1	Establish the AI inventory	Current, maintained record of every AI system the company develops, procures, or deploys — including third-party AI embedded in vendor products.
2	Approve the AI risk-appetite statement	Senior leadership's declaration of how much AI risk the company is willing to bear and on what dimensions. Required by NIST AI RMF GOVERN 2.3.
3	Resolve the committee-structure question	Decide whether AI oversight is distributed across existing committees, concentrated in a dedicated technology or AI committee, or integrated into the risk committee. Update charters.
4	Adopt the four-framework triangulation	NIST AI RMF as methodology, ISO 42001 as auditable architecture, EU AI Act as binding overlay, NACD as the board-process baseline.
5	Stand up the six management mechanisms	Inventory; appetite statement; governance committee with stop authority; pre-deployment review; post-deployment monitoring; vendor AI risk management.
6	Adopt a management report card	Quarterly or more frequent structured reporting on inventory, risk, incidents, productivity, capital, workforce, vendors, and assurance.
7	Operate the costless-virtue test	Continuously test every stakeholder-relevant commitment: would the company make any specific decision differently if it were real?

The remainder of the report develops the diagnosis behind these recommendations, the substrate behind the thesis, and the operational detail behind each action.

PART ONE

I

Diagnosis & Trajectory

Three chapters that orient the board on what AI actually is in 2026, where it is plausibly headed across one-, three-, five-, and ten-year horizons, and what the historical pattern of paradigm shifts teaches about how to govern this one.

CHAPTER 1

AI in 30 minutes

> *The board does not need to build it. The board needs to know enough to govern it: to ask the right questions, to recognize when management's confidence is and is not warranted, to detect the difference between a system that is genuinely understood and a system that is merely deployed.*
>
> — Brian R. Miller

What AI actually is in 2026

The term "AI" is used in 2026 to mean three different things, often in the same sentence. Separating them is the first move.

Machine learning is the broad family of techniques in which a computer system learns patterns from data rather than being explicitly programmed with rules. A spam filter that learns from examples is machine learning. A credit-scoring model trained on historical lending data is machine learning. This category is over fifty years old and is already embedded in nearly every large enterprise. Most of it is uncontroversial and most of it is governed under existing model-risk frameworks. Boards have been overseeing this for years even if they did not always know the term.

Deep learning is the subset of machine learning in which the patterns are learned through artificial neural networks — mathematical structures loosely inspired by neurons in the brain. Image recognition, voice transcription, and recommendation engines moved from acceptable to remarkable in the 2010s because of deep learning. This is also old enough to be governed by ordinary model-risk practices, with the caveat that the systems are harder to interpret and harder to audit than the techniques that preceded them.

Generative AI and **agentic AI** are the two categories driving the current paradigm shift, and they are what most readers mean when they say "AI" in 2026. Generative AI produces new content — text, images, code, audio, video —

in response to natural-language instructions. Large language models (LLMs) are the best-known examples. Agentic AI takes the next step: rather than producing a single output in response to a prompt, an agentic system plans a sequence of steps, calls tools and external services to execute those steps, observes the results, and iterates toward a goal it was given in natural language.

> WHY THE DISTINCTION MATTERS
>
> Classical machine learning fails statistically — its errors are predictable in distribution. Generative AI fails idiosyncratically and confidently — it will produce a wrong answer that looks completely indistinguishable from a right one. Agentic AI compounds the problem because the system can act in the world before any human has reviewed the output. The board's oversight question shifts from "is the model accurate" to "does the system have authority to act, and on what basis."

What today's AI does well

There are three categories of work where contemporary AI systems are reliably useful, and where deployment without elaborate justification is reasonable.

Pattern recognition at scale. Tasks that involve finding regularities in large bodies of data — classifying images, transcribing audio, detecting anomalies in transactions, summarizing structured data — are now genuinely solved. Where governance applies, it is the existing model-risk discipline applied with appropriate attention to bias, drift, and explainability.

Drafting and transformation of text. Producing a first draft of an email, a contract, a memo, a marketing brief; translating; summarizing; rewriting for tone; converting one document format to another; producing code from a specification — all are now reliable enough that the human's role shifts from author to editor. The productivity gain is real and measurable; the leading enterprise studies place it in the range of 25 to 60 percent for the affected workflows. The risk is also real: drafts that pass review when they should not have, hallucinated citations, plausible-sounding misstatements that survive into final documents because the editor trusts the draft.

Knowledge retrieval and authoring. Searching a corpus of documents in natural language, summarizing what was found, and drawing the user's attention to specific passages — the workflow commonly called "retrieval-augmented generation" — is the single most reliable use case for generative AI in the enterprise today. It works because the system is constrained to answer from a known document set, and because the source passages are produced alongside the response so a human can verify.

What today's AI does badly

The board's governance challenge is not to know what AI does well but to know — with confidence — what it does badly, so that the deployments which depend on it are bounded accordingly.

Hallucination

Generative AI produces fluent, plausible, confident output in domains where it has no factual grounding. It will invent citations, fabricate case law, attribute quotations to people who never said them, and produce statistics that sound right and are not.

> CASE STUDY · MOFFATT V. AIR CANADA (2024)
>
> A chatbot deployed on Air Canada's website misstated the policy for bereavement-fare refunds; a passenger acted on the misstatement; the British Columbia Civil Resolution Tribunal held that the airline owed him the refund and that the airline could not treat its chatbot as a separate legal entity from itself.
>
> The tribunal's reasoning matters more than the dollar amount: **a deployed AI system is the company speaking**, even when management would prefer it were not.
>
> Source: 2024 BCCRT 149 (British Columbia Civil Resolution Tribunal). Decision online at the tribunal's public docket.

Bias amplification

When an AI system is trained on historical data, it learns the patterns in that data, including the patterns we would not want it to learn. The 2018 Amazon hiring tool that systematically downgraded resumes from women is the textbook case — it had learned, accurately, the historical pattern in Amazon's engineering hires and projected that pattern forward as policy. The 2019 Apple Card investigation, the 2024 Workday hiring-screen lawsuit, and the growing body of fair-lending model-risk findings all share this structure. The system did what it was trained to do; the training data encoded the harm; the deployer inherited the legal exposure.

Brittleness at distribution shift

AI systems perform well on inputs that resemble the data they were trained on and degrade — sometimes silently — on inputs that do not. A fraud-detection model trained on pre-pandemic transaction patterns will not transparently report that it has become less accurate as patterns shift. A pricing model trained on one geography will not announce that it is overconfident when deployed in another. This is the heart of the model-risk discipline; AI extends rather than removes the requirement.

Reasoning gaps

Despite impressive demonstrations, contemporary AI systems do not reason the way humans do. They produce outputs that *look* like reasoning because the patterns of reasoning are well-represented in their training data. They fail predictably on novel multi-step problems, on logical constraints they have not seen before, and on questions where the correct answer requires understanding rather than pattern-matching. Boards should be skeptical of any deployment that depends on the system to "figure out" what to do in genuinely new situations.

Alignment

The most difficult failure mode is the one in which the system does exactly what it was instructed to do, and the instruction was subtly wrong. The phrase "alignment problem" in the technical literature refers to the difficulty of specifying goals in a way that captures what humans actually want rather than what the words literally say. For boards, the practical translation is that AI

systems will optimize the metric they were given, and the metric is rarely the same thing as the outcome the company wanted. Every deployment is in some measure an alignment problem.

What AI is NOT

This section exists to puncture overhype, because overhype is the single largest source of governance failure in the current moment. A board that believes too much will fail to ask the right questions; a board that believes too little will be overrun by a paradigm shift it could have governed.

NOT WHAT PEOPLE SAY IT IS

Not AGI. Artificial General Intelligence does not exist. Vendor pitches that use the term casually are a credibility signal.

Not sentient. AI does not have experiences, preferences, intentions, or moral status.

Not reliable without oversight. No commercially deployed AI system in 2026 operates at a quality level justifying removal of human review from consequential decisions.

ALSO NOT WHAT IT'S SOLD AS

Not values-free. Every system optimizes for something. Every dataset reflects choices. "Values-free AI" is a category mistake.

Not a strategy. AI is a capability. A company that announces it is "becoming an AI company" has produced a budget, not a strategy.

The economic engine

A board cannot govern AI well without understanding what AI costs and what produces those costs. The five inputs are **compute, data, talent, energy, and capital.**

Compute is the most concentrated of the five. As of 2026, Nvidia controls roughly 80 percent of the merchant market for AI training chips; the hyperscalers (Amazon, Microsoft, Google) are each building custom silicon to reduce that dependency, and that effort is partially succeeding. The compute market is supply-constrained, geopolitically contested, and the single biggest line-item driver of AI cost.

Data is the input most companies have undervalued. Distinctive proprietary data is durably scarce; generic web text is essentially commoditized. The litigation around training-data rights is unresolved and getting more expensive. A board should know what data the company owns, what rights it has to use that data for AI purposes, what data its AI vendors are using on its behalf, and what indemnification it has against training-data infringement claims.

Talent is in acute and unequal supply. Senior AI research talent commands compensation that is meaningfully above the rest of the engineering organization, which produces its own internal-equity governance problem. AI-fluent product, legal, and risk talent is scarcer than the AI-research talent and more important to deployment outcomes.

Energy has become a constraint material enough to be a board-level question. The largest training runs and inference clusters now negotiate directly with grid operators. For companies whose AI workloads are large enough to matter, energy strategy and AI strategy are now the same conversation.

Capital is the input the board sees most clearly. The trap is treating AI capital allocation as a single decision rather than as a portfolio of options under uncertainty. The companies producing the most durable returns from AI in 2026 are those that committed early to building organizational capability rather than to procuring specific AI products.

Board action items

1. **Catalog of deployed AI.** Does management maintain a current inventory of every AI system the company develops, procures, or deploys, including third-party systems embedded in products and services? If no inventory exists, that is the first task.
2. **Hallucination liability posture.** Is the company's customer-facing AI deployment structured such that the company is the speaker — and therefore liable for the speech — or has someone allowed the deployment without that recognition? *Moffatt v. Air Canada* is the case to reference.
3. **Training-data rights position.** What contractual indemnification does the company hold from its AI vendors against training-data infringement claims? Is the company itself using customer or employee data in ways that would survive scrutiny?

4. **Realistic frontier mapping.** For each material AI deployment, does management know — and can it articulate to the board — the boundary of the system's reliable competence, and the human-review machinery that catches outputs which fall outside that boundary?
5. **Vendor-pitch literacy.** Has the board's AI literacy reached a level where directors can detect the difference between a vendor whose system genuinely solves a problem in scope and a vendor whose system is being sold against a problem outside its frontier?

CHAPTER 2

The trajectory

> *The honest beginning of any forecast is to declare what is known, what is contested, and what is unknowable. The further out the horizon, the wider the uncertainty band — readers should not mistake confident prose for confident knowledge.*
>
> — Brian R. Miller

Why forecasts go wrong

The base rate of accurate ten-year technology forecasts is unfavorable. Roy Amara's observation that we overestimate technology in the short run and underestimate it in the long run is now itself a forty-year-old observation, and it has been confirmed repeatedly across the diffusion of personal computing, the internet, mobile, cloud, and crypto.

For AI specifically, the calibration challenge has three sources. First, the visible capability frontier is moving faster than the deployed capability frontier — a system can demonstrate a capability in a research setting two to three years before that capability becomes reliable enough for general enterprise use. Second, the bottleneck shifts unpredictably; in 2022 the bottleneck was data, in 2023 it was compute, in 2024 it was post-training methodology, in 2025 it was reasoning evaluation, and in 2026 it is arguably agentic-system reliability. Third, the regulatory and litigation environments are themselves moving variables; a capability that is technically possible may be commercially unavailable for reasons that do not appear in any technical roadmap.

What this chapter offers, accordingly, is not a prediction. It is a set of working scenarios annotated with the most important variables to watch. The board's task is not to bet on the central scenario; it is to know which evidence would shift the assessment, and to track that evidence.

The one-year view (2026 — mid-2027)

In the next twelve months, three movements are highly likely.

Foundation-model commoditization continues. Models in the GPT-4 / Claude 3.5 / Gemini 1.5 generation will continue to fall in price per token at the rate of approximately five to ten times per year that has held since 2023. By the end of 2026, what was a frontier model at the start of the year will be available at roughly one-tenth the cost — and increasingly available from open-weights providers without per-token fees at all.

The agentic-system reliability gap closes — incompletely. Agentic systems that plan, call tools, and act on the world have been deployed in 2025 and 2026 in narrow, well-defined domains. Their reliability outside well-defined domains is poor and improving slowly. The next twelve months will see the gap close further in specific domains but not generally. Boards should expect *useful* agentic deployments by mid-2027; they should not expect *general-purpose* agentic deployments.

Enterprise deployment shifts from experiment to operations. The 2024-2025 phase, in which most enterprise AI was funded out of innovation or experimentation budgets, is ending. Through 2026 and into 2027, the budget conversation moves to the operating line: the same scrutiny that applies to other operational technology applies to AI.

The three-year view (2026 — 2029)

At three years, the central scenario is straightforward to describe and the upside and downside scenarios are wider.

Central scenario. AI is deeply integrated into knowledge work — drafting, summarization, search, code generation, customer interaction, structured-data analysis. The integration is uneven across functions: legal, finance, sales operations, engineering, marketing, and customer support are heavily AI-augmented; supply-chain operations, HR, and corporate strategy are partially augmented; physical operations and skilled trades are minimally affected by AI directly but significantly affected indirectly through the economy. Productivity

gains are real and measurable but smaller than the most enthusiastic 2023 forecasts suggested. The companies that captured the gains are those that built organizational capability rather than those that procured specific AI products.

Upside scenario. Agentic systems achieve reliable multi-step execution in additional domains — financial analysis, customer relationship management, software operations, certain categories of legal and regulatory work. The productivity gain in those domains is meaningfully larger than in the central scenario.

Downside scenario. Three contributors push the central scenario lower. First, regulatory friction (most notably EU AI Act enforcement starting August 2026) raises the cost of deployment. Second, the post-training quality gains plateau. Third, public trust deteriorates from a high-profile failure (a fatality, a major election-interference event, a large-scale data exfiltration), changing the political and commercial environment within which AI is sold.

The five-year view (2026 — 2031)

At five years, paradigm-level effects become visible.

The deployment phase begins in earnest. The Carlota Perez framework that Chapter 3 develops distinguishes between the *installation phase* of a technological paradigm — when the infrastructure is being built, often in a speculative-finance bubble — and the *deployment phase* — when the infrastructure is mature, the bubble has corrected, and the productivity gains diffuse through the broader economy. For AI, the 2020-2025 period had clear installation-phase characteristics. The 2026-2031 period is likely to mark the transition to deployment.

Workforce adjustment becomes the central question. Five years is enough time for meaningful labor-market adjustment. The empirical evidence in 2026 already shows shifts in entry-level roles in software, marketing, and customer service; by 2031 these shifts will have compounded. The question is not whether displacement happens but whether the transition is managed in ways that preserve worker dignity and produce broadly shared gains, or whether it is allowed to run as pure cost optimization. Chapter 6 treats this at length.

The regulatory environment hardens. The EU AI Act is fully phased in by August 2027; the US regulatory environment is much less predictable but the direction of motion is toward more rather than less. By 2031, the cost of regulatory compliance for material AI deployments will be a significant operational line item, and the differential cost across jurisdictions will shape where AI capabilities are developed and deployed.

Foundation-model concentration partially erodes. The current concentration of frontier capability among a handful of US-based labs and a smaller set of Chinese labs is the most-cited example of AI-era platform concentration. By 2031 the open-weights ecosystem is mature enough that the dependency relationship is meaningfully different: for many enterprise uses, organizations can run capable models on their own infrastructure with no per-token vendor fee. The frontier remains expensive and concentrated; the merely-useful middle does not.

The ten-year view (2026 — 2036)

At ten years, the prudent posture is to widen the uncertainty bands and concentrate on identifying what boards should *plan for* versus what they should *hedge against*.

PLAN FOR (HIGH CONFIDENCE)

AI capabilities at the 2026 frontier — drafting, summarization, code generation, knowledge retrieval, single-domain agentic execution — will be commoditized.

20–50 percent of knowledge-work activities will be AI-augmented or AI-executed by 2036.

Regulation will be more elaborate, more globalized, and more sector-specific.

HEDGE AGAINST (UNCERTAIN)

Whether the foundation-model paradigm continues to improve through scaling, or plateaus and is succeeded by a different architecture.

Whether artificial general intelligence is achieved, approached, or remains out of reach.

Whether the geopolitical environment around AI produces a fragmented, concentrated, or intermediate landscape.

Honest uncertainty bands — what no one knows

The following list is what boards should treat as unknowable today, and what management presenting on the AI strategy should not claim to know:

- The timing and shape of the next foundation-model architectural shift.
- Whether the current rate of capability improvement continues, accelerates, or plateaus.
- The path of US federal AI legislation.
- The outcome of the foundational training-data and AI-liability cases.
- The reliability ceiling for general-purpose agentic systems.
- The terminal level of compute concentration vs. distribution.
- The labor-market adjustment dynamics — slow (1980s manufacturing) or fast (2000s media).
- The actions of state actors — China's AI policy, EU regulatory enforcement intensity, multilateral compute restrictions.

> A NOTE ON CONFIDENT ANSWERS
>
> A board that hears confident answers on these questions from a vendor or a consultant should treat the confidence itself as a signal of unreliability.

Board action items

1. **Adopt a scenario-based posture rather than a forecast-based one.** The board should ask management to present its AI strategy as performance across the central, upside, and downside scenarios — not as a single best-guess plan. Plans that work only in the central scenario are not strategies; they are bets.
2. **Identify and track the leading indicators.** Model-quality benchmarks, regulatory milestones, litigation outcomes, enterprise productivity studies, and competitor disclosure are the public signals that discriminate among scenarios. A board-level dashboard tracking three to five leading indicators is a small investment that pays large governance dividends.
3. **Treat the deployment phase as the strategic moment.** Companies that build organizational capability (data, governance, talent, change management)

during the installation-to-deployment transition consistently outperform those that procure point solutions.

4. **Plan governance machinery to be paradigm-robust.** The frameworks the company adopts now should accommodate further regulatory accretion without architectural rework.
5. **Reject false certainty.** A vendor or consultant who claims to know the ten-year trajectory should be evaluated on calibration, not capability. The board's most valuable advisors will be the ones who tell it what they do not know.

CHAPTER 3

Lessons from past disruptions

> *Boards have governed through technological paradigm shifts before. The shifts are rarer than the rhetoric suggests — perhaps five in the last two centuries qualify — but the pattern across them is consistent enough to be useful. Calibration, not nostalgia.*
>
> — Brian R. Miller

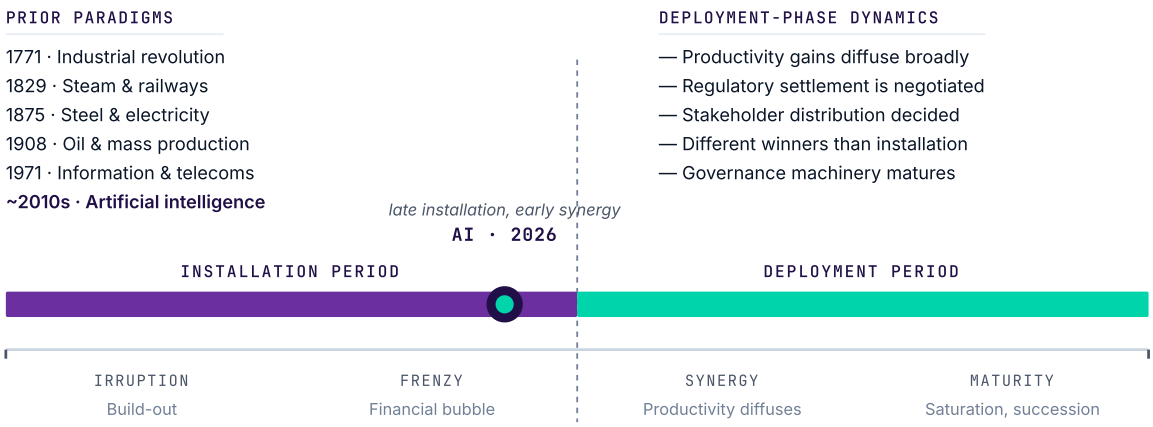
The Schumpeter / Perez lens

Joseph Schumpeter's mid-twentieth-century concept of *creative destruction* described capitalism as a process in which entrepreneurial innovation continually destroys the structures of the previous economic order even as it builds the next. Schumpeter's insight was that the productive churn was not an accident or a failure — it was the mechanism. The implication for boards was that incumbency was always temporary, that the value of a defensive moat decayed faster than its accounting depreciation suggested, and that the comfortable extrapolation of recent earnings was nearly always wrong at paradigm shifts.

Carlota Perez took Schumpeter's framework and operationalized it. Her 2002 book *Technological Revolutions and Financial Capital* identified a recurring four-phase pattern across the five great technology revolutions since 1771. The four phases — irruption, frenzy, synergy, and maturity — fall into two larger periods. The *installation period* (irruption + frenzy) is when the new infrastructure is built, often in a speculative-finance bubble that overshoots the underlying productive capacity. The *deployment period* (synergy + maturity) is when the infrastructure is mature, the bubble has corrected, and the productivity gains diffuse through the broader economy.

FIGURE 1 · THE PEREZ TECHNOLOGICAL-PARADIGM CYCLE

Where AI sits today across the two-period model that has governed every prior paradigm shift



Source: Adapted from Perez, C. — Technological Revolutions and Financial Capital (Edward Elgar, 2002).

For the purposes of board governance, the Perez framework produces three observations that recur across cases.

The financial bubble of the installation phase concentrates wealth among the infrastructure builders, not the eventual beneficiaries. This is the pattern in railway speculation, in dot-com speculation, and now in AI-infrastructure speculation. The companies that captured the largest share of the value created by the previous paradigm shifts were rarely the companies that built the infrastructure.

The largest productivity gains arrive in the deployment phase, not the installation phase. Electricity was invented in the 1880s; the productivity gains from electrification diffused through US manufacturing in the 1910s and 1920s — a 30-to-40-year lag. The internet was commercialized in 1995; the productivity gains diffused most broadly through the 2010s and 2020s. The implication is that boards who allocate capital under the assumption that the installation-phase players will dominate the deployment phase are usually wrong.

The transition between installation and deployment is governed as much as it is invented. Each prior paradigm produced a wave of governance institutions — antitrust law for railways, labor and safety law for mass production, financial regulation after the 1929 crash, broadcast and telecommunications regulation in their respective eras. The shape of those institutions strongly affected which companies and which stakeholders captured the gains.

Five paradigms in pattern form

Railroads (1840s — 1870s)

The British and American railway booms of the mid-nineteenth century are the first case where the Perez pattern is unmistakable. Capital flowed into railway-company stocks at a pace that produced a textbook financial bubble (the British "Railway Mania" of 1844-1846 saw at least 263 acts of Parliament authorizing new railway construction). The bubble corrected painfully — by 1850 the index of British railway-company share prices was approximately 65 percent below its 1845 peak.

What boards got right and wrong:

- **What was built survived; what was overcapitalized did not.** The physical infrastructure became the backbone of industrial economies for the next century. The corporate vehicles that overpaid for the construction did not.
- **The companies that captured the largest share of railway-era value were not the railroads.** They were the companies that arose to *use* the railroad — Sears Roebuck, Standard Oil, Carnegie Steel — and the financial institutions that intermediated the capital flows.
- **Governance arrived late and was reactive.** The Interstate Commerce Act of 1887 and the antitrust framework that culminated in the Sherman Act of 1890 came after the worst abuses.

Electricity (1880s — 1920s)

The electrification of industry is the most-studied case of a long productivity lag between invention and impact. Edison's commercial generating station at Pearl Street opened in 1882. Electric power was widely available in US industrial

centers by 1900. But manufacturing productivity gains from electrification did not show up in the aggregate data until roughly 1915-1920, and the largest gains came in the 1920s.

The cause was not insufficient technology. It was institutional. Factories built before electrification were laid out around central steam-driven shafts that transmitted power via belt and pulley. Retrofitting them with electric motors that simply drove the existing shaft produced only modest gains. The real productivity gain required redesigning the factory entirely — distributed electric motors at each machine, flexible layouts, single-floor construction, electric lighting permitting a second and third shift, electric materials handling. That redesign took thirty years and required not only the technology but a generation of managers and engineers who could think in the new paradigm.

> *The most important investments may not be in the AI itself but in the workflow, organizational, and human-capability redesign that lets the AI produce outsized returns. The lag is not avoidable; it can only be managed.*
>
> — The electrification lesson, applied

IT and the internet (1960s — 2010)

The information-technology paradigm covers a longer span than the prior two and is the one most boards have lived through. The financial bubble of 1999-2000 is the textbook example of installation-phase frenzy. The Nasdaq peaked in March 2000 at roughly 5,000 and declined to roughly 1,100 by late 2002 — a 78 percent decline in a single index over thirty months. The companies that survived the crash and went on to dominate the deployment phase — Amazon, eBay, Google, Salesforce — looked at the trough of 2002 like much smaller versions of themselves.

Boards that mistook the bubble for the trend lost the most. The companies that levered up to participate in the late stages of the dot-com run were destroyed in the correction. **Boards that mistook the trough for the end of the trend also lost.** Several mainline retailers and media companies treated the 2002-2004 trough as evidence that the internet was a failed paradigm and disinvested. They were wrong; the deployment phase was just beginning. **The most**

consequential governance failures were in financial reporting, audit, and acquisition integrity. Enron, WorldCom, and Tyco produced the Sarbanes-Oxley Act of 2002.

Mobile and cloud (2007 — 2020)

The iPhone shipped in 2007. AWS launched its first commercial services in 2006. The two together became the defining computational paradigm of the 2010s. The decade in which user-data collection scaled — 2008 to 2018 — was largely a decade in which boards did not interrogate data practices closely. The eventual regulatory response (GDPR in 2018, CCPA in 2020, the patchwork of US state privacy laws since) was both predictable and avoidable; boards that did not push management to anticipate it inherited an expensive remediation. The 2013-2017 wave of large-scale breaches — Target, Anthem, Equifax, Marriott — was a function of cloud-era data concentration outpacing the security and governance machinery around it.

Patterns that repeat

Five patterns recur across the cases.

> PATTERN 1 · THE BUBBLE LOOKS LIKE THE TREND
>
> Late-installation asset prices substantially overshoot underlying productive capacity. Boards that treat late-bubble valuations as the basis for capital allocation systematically destroy value. Boards that treat the post-correction trough as the start of the deployment phase systematically create it.

> PATTERN 2 · CAPABILITY IS NOT THE SAME AS BENEFIT
>
> The largest beneficiaries of each prior paradigm shift were not the companies that built the infrastructure but the companies that used the infrastructure to do something previously impossible.

PATTERN 3 · THE PRODUCTIVITY LAG IS INSTITUTIONAL, NOT TECHNOLOGICAL

The dynamo did not produce a productivity gain until the factory was redesigned. The internet did not produce a productivity gain until the workflows were redesigned. AI will not produce its full gain until organizations have rebuilt the work, the roles, the data infrastructure, and the management practices around it.

PATTERN 4 · GOVERNANCE ARRIVES

The question is whether the board anticipates it. Antitrust, labor law, financial regulation, data-privacy regulation — each prior paradigm produced a substantial governance settlement. Boards that anticipated the settlement consistently outperformed boards that resisted.

PATTERN 5 · STAKEHOLDER DISTRIBUTION IS DETERMINED IN THE DEPLOYMENT PHASE

The installation phase concentrates wealth among the infrastructure builders. The deployment phase determines whether the productivity gains are broadly shared or narrowly captured. The institutions, labor settlements, regulatory frameworks, and corporate practices established during the transition govern that distribution.

Where we are in the AI paradigm cycle

By the criteria in the Perez framework, AI in 2026 is in the late installation phase / early synergy transition. The infrastructure build-out is well advanced but not complete. The financial-capital concentration in installation-phase players is at or near a peak. The institutional governance settlement is being actively negotiated. The deployment-phase productivity gains are visible in narrow domains and not yet broadly diffused.

A board reasoning by historical analogy should expect: a correction in the installation-phase asset prices within the next several years; a meaningful broadening of the productivity-gain distribution as deployment matures; an accelerated regulatory environment that will impose costs but will also clarify

the rules of competition; and a stakeholder-distribution settlement that will be either captured by a few or shared more broadly depending substantially on the governance choices boards make in 2026-2030.

Board action items

1. **Adopt the installation-vs-deployment frame in capital allocation.** Distinguish between investments that depend on installation-phase asset prices remaining elevated and investments that produce value in the deployment phase. The latter should dominate.
2. **Plan for the productivity lag.** The largest gains from AI will not materialize on a two-year ROI horizon. Boards that demand short-horizon AI ROI will starve the deeper investments that produce the larger gains.
3. **Anticipate the governance settlement.** Build governance machinery now that absorbs further regulatory accretion without rework.
4. **Take the stakeholder-distribution question seriously.** The board's choices over the next five years materially affect whether the AI paradigm's gains are broadly shared.
5. **Distinguish picks-and-shovels companies from deployment beneficiaries.** Today's foundation-model providers are necessary inputs to the company's AI strategy; they are unlikely to be its largest sources of advantage.

PART TWO

II

The Stakeholder-Value Thesis

The conviction that drives every Part that follows. Economic and social outcomes are not on a fixed-pie trade-off. The board's job is to be intentional about which path the company is on.

CHAPTER 4 · THE THESIS IN FULL

CHAPTER 4

The thesis in full

> *The thesis is not ESG-lite. It is a fiduciary-and-governance enterprise. Its claim is that the duties a board already owes — care, informed judgment, monitoring — extend, in the AI paradigm, to a wider set of consequences than the duties were historically read to cover.*
>
> — Brian R. Miller

The thesis stated in full

THE THESIS

By aligning economic and social outcomes, a company can maximize value for all stakeholders — employees, customers, shareholders, and broader society. Driving that kind of broad value requires real strategy and governance work, not parroting the issue of the day. The board's job in the AI paradigm is to build serious governance machinery that pivots as lessons land.

The thesis has four moving parts, and each is defended separately below.

First, *alignment is possible.* Economic and social outcomes are not on a fixed-pie trade-off such that one stakeholder's gain requires another's loss. There are paths through the AI paradigm in which all four stakeholder groups gain together, and there are paths in which only one or two do. The board's job is to be intentional about which path the company is on.

Second, *the alignment requires real work.* It is not a default outcome of pursuing shareholder return alone, and it is not a default outcome of declaring stakeholder principles. It is a function of decisions actually made — about which deployments to fund, which to delay, which to reject; about how the workforce transition is managed; about which jurisdictions to engage and on what terms; about which vendors to choose and which to walk away from.

Third, *parroting the issue of the day is the failure mode.* The largest reputational risk in the AI paradigm is not that a company refuses to engage with stakeholder concerns; it is that it engages performatively, produces statements without substance, and is caught when the substance is asked for.

Fourth, *the work pivots as lessons land.* The governance machinery boards build in 2026 will need substantial revision by 2028, and again by 2031. Boards that design for permanence are designing for obsolescence. Boards that design for pivot are designing for what the paradigm actually requires.

This is not ESG-lite

A board reader is rightly skeptical of any framing that sounds like a repackaging of the environmental, social, and governance (ESG) conversation of the late 2010s. The ESG conversation produced real movement on some questions (climate-risk disclosure has materially advanced; board-level diversity questions are now table stakes) and produced backlash on others, with the backlash itself becoming a significant political and commercial dynamic by the mid-2020s. The thesis here is not a recycling of that conversation. The differences matter.

ESG, as it became operationalized in capital-markets practice, was primarily a *scoring* enterprise. Ratings agencies produced ESG scores; asset managers used those scores in portfolio construction; companies optimized for the scores. The mechanism connecting the score to underlying performance was weak; the mechanism connecting the score to stakeholder outcomes was weaker still. The system was vulnerable to the criticism that companies were optimizing for the appearance of stakeholder responsibility without producing stakeholder benefit — a vulnerability that became the political handle for the backlash.

The thesis here is not a scoring enterprise. It is a fiduciary-and-governance enterprise. Its claim is that the duties a board already owes — to act with care, to inform itself adequately, to monitor management's exercise of the company's power — extend, in the AI paradigm, to a wider set of consequences than the duties were historically read to cover. The board cannot govern the AI deployments well without taking seriously their effects on employees displaced from jobs, on customers exposed to algorithmic decisions, on communities where data is collected, on the broader information environment within which the company sells.

The costless-virtue test

The single most useful diagnostic for distinguishing authentic stakeholder commitment from its imitation is the **costless-virtue test.** The test is straightforward: a stated commitment that costs the company nothing — no revenue forgone, no efficiency given up, no convenience surrendered, no shareholder return reduced — is indistinguishable from marketing.

The pattern is recognizable across cases.

> PATTERN A · AI ETHICS PRINCIPLES WITHOUT CONSEQUENCE
>
> A company adopts a set of principles — fairness, transparency, accountability, beneficence — and embeds them in marketing materials. The principles are not connected to deployment-gate decisions, to compensation, to override authority, or to public reporting. The principles produce no behavior the company would not have produced without them. The principles are costless. They will not survive scrutiny.

> PATTERN B · THE "RESPONSIBLE AI" TEAM WITHOUT AUTHORITY
>
> A company stands up a responsible-AI function and publicizes it. The function reports up through a product organization and lacks an independent reporting line to the audit committee, the board, or any party outside management. The function can identify problems but cannot stop deployments. The function is costless to the deployment velocity.

> PATTERN C · THE "FAIRNESS AUDIT" THAT DOES NOT GATE
>
> A company commissions a third-party algorithmic-fairness audit of a deployed system. The audit identifies disparate impact. The deployment continues unchanged. The audit produces a published statement that an audit was conducted. The audit is costless to the deployment. It is, in fact, worse than no audit — it produces documentary evidence that the company was aware of disparate impact and proceeded.

The corollary of the test is more demanding: authentic stakeholder commitment in the AI paradigm *will* cost the company something. Specific deployments will be slowed or stopped. Specific revenue will be forgone

because the customer relationship would require a deployment the company will not stand behind. Specific competitors will gain in the short run by deploying what this company refuses to. Boards that flinch at these costs will produce costless virtue.

The accountability-deflection patterns to watch for

Closely related to costless virtue is the set of rhetorical structures that deflect accountability when stakeholder commitments fail to materialize. Four patterns recur frequently enough to deserve named recognition.

Pattern	The deflection	Where it fails
Denial	"The system did not do that. The output was not from us."	A deployed AI system is the company speaking (the Air Canada tribunal's reasoning).
Boundary-evasion	"That is not our domain. That is the responsibility of [policy / the user / the regulator]."	When the harm is foreseeable and the company has the leverage to prevent or remedy it.
Virtuous framing	"This is the privacy-preserving option. This is the safe option. We did this for the user's protection."	When the principle is applied selectively, only where it limits company liability.
Complexity	"This is too technical to explain. The model is a black box."	The complexity is a description of the original deployment decision the board should have governed.

Why this is hard-headed strategy, not signaling

Three arguments support the conclusion that taking the stakeholder thesis seriously is *strategically superior* — not merely ethically defensible — in the AI paradigm.

Argument from regulatory inevitability

The governance frameworks treated in Chapter 5 (NIST, EU AI Act, ISO 42001, NACD guidance) all build in directions that increase the cost of stakeholder externalities. The companies that anticipate the direction of regulatory motion

and build governance machinery accordingly have lower compliance cost in 2028 than the companies that resist and retrofit. The EU AI Act's 7-percent-of-global-revenue maximum penalty is large enough that boards do not have the luxury of treating governance as separate from strategy.

Argument from talent

AI capability is built and operated by people, and the most capable AI people in the world are unusually attentive to the stakeholder consequences of the systems they build. The waves of departures from leading AI labs over safety and alignment concerns are evidence that the talent pool the company needs has preferences about the work it does. Companies that take the stakeholder questions seriously recruit and retain better. Companies that do not lose the marginal hire who would have most improved the deployment.

Argument from the deployment phase

The Perez pattern predicts that the deployment phase of a paradigm is governed by a different set of rules than the installation phase. The rules of the deployment phase favor companies whose products and services produce broadly distributed value, are operated within the regulatory settlement, and are trusted by the constituencies they serve. The installation-phase companies that mistreated those constituencies on the way up — through monopolistic pricing, through data extraction without consent, through opaque decision-making in consequential contexts — face a steeper climb in the deployment phase than the companies that built trust as they built capability.

Objections answered

Three objections recur whenever the thesis is presented. Each deserves a direct answer.

Objection: "Our duty is to shareholders. The thesis sounds like it dilutes that duty."

The reply is that fiduciary duty has never been a license to ignore the legal, regulatory, reputational, and operational consequences of decisions. A board that approved a deployment that produced 10 percent short-run revenue lift and 25 percent loss of long-run franchise value would be breaching fiduciary duty

even though its motive was shareholder service. The thesis is not asking boards to subordinate shareholder duty to other constituencies; it is asking boards to recognize that the shareholder interest in 2030 is meaningfully affected by the company's stakeholder posture in 2026. The two interests are not on a tradeoff axis in the AI paradigm; they are coupled.

Objection: "Our competitors will move faster if they ignore these concerns. We will lose."

The reply has two parts. First, the comparison is to a counterfactual that rarely materializes — the competitor who ignores stakeholder concerns and faces no consequences typically does so in markets where consequences arrive slowly. The history of platform-era reputational meltdowns suggests that the consequences arrive faster than this objection assumes. Second, where the objection is correct, the board's question is not "should we match" but "should we operate in a market where matching is the price of competition." Some markets the company should not be in.

Objection: "These are political questions. They are not for the board."

The reply is that the questions the thesis raises — what happens to employees displaced by AI, what privacy posture the company adopts, what disclosure it makes about algorithmic decisions, what jurisdictions it operates in and on what terms — are not political in the partisan sense. They are governance questions in the historical sense, and they have been part of corporate-board work for the entire history of corporate boards. The novelty of AI does not make them new; it makes them urgent.

Board action items

1. **Adopt the costless-virtue test as a discussion discipline.** Whenever management presents a stakeholder-relevant commitment, the board's first question should be "what specific decisions would the company make differently if this commitment is real, and which of those decisions have already been made?"
2. **Surface the deflection patterns when they appear.** Train the board to recognize the four patterns — denial, boundary-evasion, virtuous framing, complexity — and to name them when they appear in management presentation.

3. **Resist scoring as governance.** The board should not allow stakeholder posture to be evaluated primarily through ESG-style scores or vendor-produced ratings. Scores are useful inputs; they are not governance.
4. **Couple the stakeholder posture to capital allocation.** A stakeholder commitment that does not appear in the capital plan is not a commitment.
5. **Build the governance machinery to pivot.** The board's job is to ensure the machinery exists; the machinery's job is to surface the lessons that require pivoting.

PART THREE

III

Governance Frameworks

Four primary frameworks — NIST AI RMF, the EU AI Act, ISO 42001, and NACD guidance — in plain English, with the board's triangulation strategy across them.

CHAPTER 5 · NIST · EU · ISO · NACD

CHAPTER 5

NIST • EU • ISO • NACD

> *The four frameworks are not substitutes for one another. NIST AI RMF provides methodology. ISO 42001 provides auditable management-system architecture. The EU AI Act provides binding legal requirements. NACD guidance provides the board's operating spec. A serious program references all four.*
>
> — Brian R. Miller

AI governance vs. corporate governance

Before the frameworks themselves, the distinction the rest of this chapter will turn on: *AI governance* and *corporate governance* are related but separable functions.

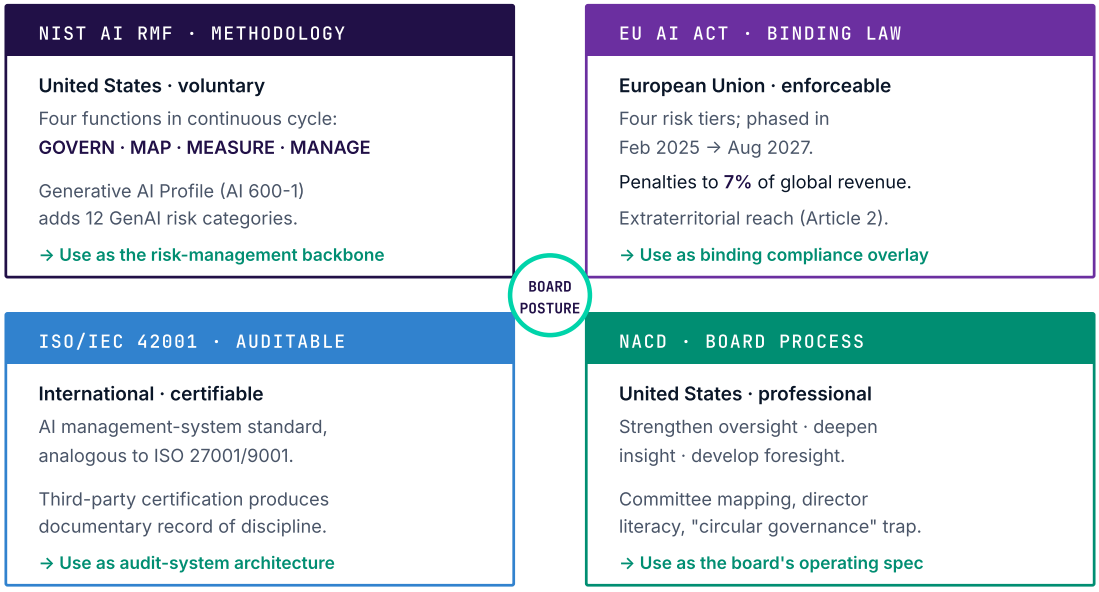
AI governance is the operational function within management that manages AI-specific risks — model risk, bias, hallucination, privacy, security, vendor management, deployment authorization, post-deployment monitoring. It sits inside management. It is staffed by people with AI-specific competence, often coordinated by a chief AI officer, chief data officer, or risk function. The frameworks treated below — NIST, EU AI Act, ISO 42001 — are largely AI-governance frameworks. They specify what the AI-governance function should do.

Corporate governance is the board-level function that oversees management, including management's exercise of the AI-governance function. It is staffed by directors with general business and oversight competence, supplemented in the AI paradigm by directors with sufficient AI literacy to interrogate management's claims credibly. NACD guidance is largely corporate-governance guidance. It specifies what the board should do.

Conflating the two produces predictable failures. A board that takes on AI governance directly — trying to specify the model risk procedures, debating bias thresholds — has stepped outside its competence and into management's. A board that delegates corporate governance entirely — accepting management's statements that "AI governance is handled" without independent verification — has abdicated its own function.

FIGURE 2 · FOUR-FRAMEWORK TRIANGULATION

Each framework serves a function; the board's posture is to triangulate across all four



Source: Author analysis of NIST AI 100-1, EU Regulation 2024/1689, ISO/IEC 42001:2023, NACD Blue Ribbon Commission 2024.

NIST AI Risk Management Framework

The NIST AI RMF was published in January 2023 as a voluntary framework for managing AI risks across the lifecycle of AI systems. It remains the most widely-referenced US framework. The core framework is still at version 1.0; the most consequential extension is **NIST AI 600-1: Generative AI Profile**, published July 2024, which adapts the framework specifically for generative AI.

The framework organizes risk management around four functions — **Govern, Map, Measure, Manage** — designed to operate as a continuous cycle rather than a sequential checklist.

The four functions in plain English

Govern establishes the organizational machinery for AI risk: who is accountable, what risk tolerance the organization has declared, what policies apply, how the AI inventory is maintained, what training the affected workforce has. The most directly board-relevant provision in the entire framework is subcategory **GOVERN 2.3**, which assigns risk-tolerance declaration explicitly to senior leadership. A board that has not formally approved an AI risk-appetite statement has not satisfied the framework's foundational expectation.

Map is the function that produces situational awareness about each AI system: its intended purpose, the populations affected, the foreseeable harms (including harms to third parties), the inputs and outputs, the data flows. Map's most useful output for boards is the *go/no-go decision* — the analytic basis for deciding whether a system should be developed, deployed, or continued.

Measure selects and applies metrics for AI risk: reliability, safety, security, bias, explainability, privacy, environmental footprint. Its most useful provision for boards is the requirement that measurement methodology be subjected to *independent review*. A board should ask, for each high-risk deployment, who reviewed the measurement methodology and what their independence consisted of.

Manage translates measured risk into prioritized action: risk treatment plans, incident response procedures, post-deployment monitoring, and — crucially — disengagement procedures for systems that have failed. *Systems whose disengagement procedure is "we will look into it" should not be deployed.*

> THE GENAI PROFILE (NIST AI 600-1)
>
> The GenAI Profile adds twelve risk categories specific to generative AI — including hallucination ("confabulation" in the framework's term), harmful bias from training data, lowered barriers to cyber-attack, environmental impact, and privacy violations through inference. The profile maps over 200 recommended mitigations across the four functions. Companies deploying generative AI without reference to the GenAI Profile face an increasing documentation gap as regulatory and litigation scrutiny intensifies.
>
> Source: NIST AI 600-1, July 2024. Available at nvlpubs.nist.gov.

The EU AI Act

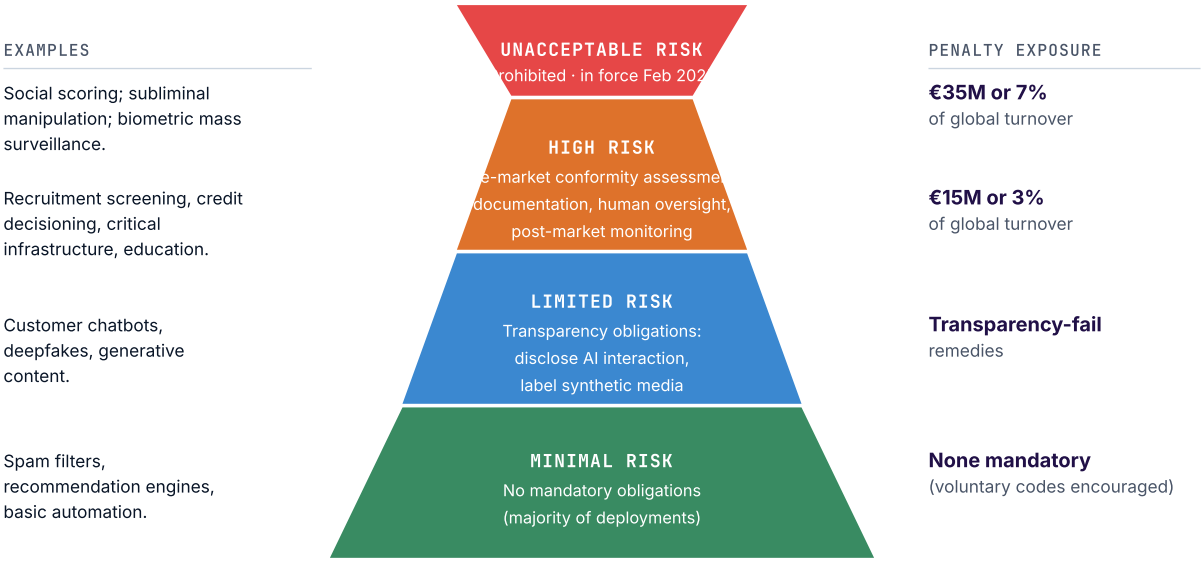
The EU Artificial Intelligence Act, formally Regulation (EU) 2024/1689, was adopted in June 2024 and entered into force August 1, 2024. It is the world's most comprehensive AI-specific law to date. Its provisions are phased in across a multi-year schedule.

Date	What activates
2 February 2025	Prohibitions on unacceptable-risk practices. AI literacy obligations.
2 August 2025	General-purpose AI (GPAI) model obligations. Governance institutions operational.
2 August 2026	Full general regime applies. High-risk system enforcement begins. Article 50 transparency obligations.
2 August 2027	High-risk AI embedded in regulated physical products (medical devices, machinery, certain toys).

The Act's central structure is a **four-tier risk classification.**

FIGURE 3 · EU AI ACT RISK-TIER STRUCTURE

Obligations scale by use-case risk; penalties rise to 7% of global annual turnover



Source: Regulation (EU) 2024/1689, Articles 5, 6, 50, 99. Penalty figures Article 99 § 1–5.

Three features consequential for US-headquartered companies

Extraterritorial reach. The Act applies to providers established outside the EU when "the output produced by the AI system is used in the Union." A US company that does not sell to EU customers can fall within scope if its AI outputs are consumed within the EU through cloud services, licensed models, or intermediary platforms.

GPAI obligations. Providers of general-purpose AI models — most frontier foundation models — face documentation, copyright-compliance-policy, and downstream-disclosure obligations regardless of headquarters. Models meeting the systemic-risk threshold (currently presumed for training compute exceeding 10^{25} FLOPs) face additional notification, risk-assessment, and incident-reporting obligations.

Quality-management-system requirement (Article 17). Providers of high-risk systems must operate a quality management system that includes "an accountability framework setting up the responsibilities of management and

other staff." This is the EU AI Act provision that most directly contemplates board-level accountability for AI governance. A board that has not satisfied this provision is carrying unquantified legal exposure under a regime whose maximum penalty exceeds the GDPR's.

ISO/IEC 42001 — the auditable management system

ISO/IEC 42001, published in December 2023, is the international standard for AI Management Systems (AIMS). It serves a different function than the prior two frameworks: where NIST AI RMF provides risk-management methodology and the EU AI Act establishes binding legal requirements, ISO 42001 specifies an auditable management-system architecture. Organizations can be certified to ISO 42001 by accredited third-party bodies, in the same way that organizations are certified to ISO 27001 (information security) or ISO 9001 (quality management).

The certifiability matters. A certified management system produces a documentary record that is recognizable to regulators, insurers, customers, and litigation counterparties. Companies that operate in regulated sectors, that contract with governments, or that face cross-border AI deployments increasingly find that the ISO 42001 certification is functionally required.

NIST has published a formal crosswalk between AI RMF 1.0 and ISO 42001, enabling integrated implementation. For boards, the practical question is whether the company's AI governance program is auditable in the ISO sense — whether an outside party with the relevant standard in hand could examine the company's documentation, observe the operating practices, and reach a conclusion. *Programs that exist primarily in slide decks and that produce no audit trail will not certify and should not satisfy the board.*

NACD guidance — the board-facing standard

The National Association of Corporate Directors has produced the most comprehensive board-facing AI governance literature in the US, anchored by its 2024 Blue Ribbon Commission report *Technology Leadership in the Boardroom* and the accompanying *Director Essentials: Implementing AI Governance*.

NACD's structural recommendation organizes the board's work around three imperatives.

> **STRENGTHEN OVERSIGHT**
>
> Upgrade committee charters; define decision-making authorities between board and management; ensure trustworthy AI use aligned with organizational purpose and values. Audit committee owns AI risk integration into ERM. Compensation committee owns workforce and incentive design. Nominating-and-governance owns director skills and succession.

> **DEEPEN INSIGHT**
>
> Establish metrics for technology oversight, request briefings at a cadence that matches the pace of AI development (monthly or bi-weekly for highest-risk deployments rather than quarterly), build structured channels to internal technology leaders.

> **DEVELOP FORESIGHT**
>
> Recognize AI as a core element of long-term strategy, redesign the board calendar for forward-looking technology discussions, reassess the company's AI posture regularly as the environment evolves.

NACD's 2025 Public Company Board Practices and Oversight Survey provides the data: more than 62 percent of directors now allocate full-board agenda time for AI; approximately 40 percent of companies have charged at least one board-level committee specifically with AI oversight (up from 11 percent in 2024); close to 50 percent of S&P 500 companies now mention AI in their descriptions of director qualifications (up from 26 percent in 2024). The directional movement is unmistakable. The lagging companies are now visibly behind their peers.

NACD identifies six risk categories that boards should treat as standing agenda items: **bias in AI outputs, explainability deficits, privacy exposures, security vulnerabilities, potential for misuse, and model drift.** NACD warns specifically against *"circular governance"* — the condition in which the entities being

overseen effectively dictate the terms of oversight — a failure mode with particular salience when management controls the AI systems the board is charged with evaluating.

Triangulation — where the frameworks agree and diverge

The four primary frameworks **agree** on more than they disagree on. The points of convergence are the operational core of any serious AI governance program:

- **AI inventory.** All four require, implicitly or explicitly, that the organization knows which AI systems it operates. The inventory is the foundational artifact.
- **Risk classification by use case.** All four categorize risk by what the system is used for rather than by the underlying technology.
- **Human oversight.** All four require that humans remain in a position to review, override, and disengage AI systems whose outputs are consequential.
- **Documentation and traceability.** All four require records sufficient for an outside party to reconstruct what the system did and why.
- **Bias, fairness, and disparate impact.** All four treat algorithmic bias as a first-order risk.

The points of **divergence** are where the board's triangulation work matters:

- **Voluntary vs. binding.** NIST AI RMF and NACD guidance are voluntary; ISO 42001 is certifiable but voluntary; the EU AI Act is binding law.
- **Pre-market vs. post-market.** The EU AI Act includes a substantial pre-market regime (conformity assessment, registration, CE marking) that has no analog in the US frameworks.
- **Penalty exposure.** Only the EU AI Act carries direct financial penalties.

Cautionary cases — what governance failure looks like

Six public cases illustrate the failure modes the frameworks are designed to prevent.

Case	What happened	Framework function that failed
Moffatt v. Air Canada (2024)	Chatbot misstated bereavement policy; airline tried to disclaim liability.	NIST MANAGE; EU AI Act Article 50 transparency.
Samsung ChatGPT data leakage (2023)	Engineers uploaded proprietary code to consumer AI within weeks of policy reversal.	NIST GOVERN. Communication without enforceable policy or technical control.
Amazon recruiting tool (2014–2018)	Algorithm penalized women candidates; internal awareness for years before discontinuance.	NIST MAP (intended use, disparate-impact analysis); MEASURE (bias monitoring).
Apple Card credit disparity (2019)	Lower credit limits for women with equivalent credit profiles; no recourse mechanism.	Explainability and transparency obligations recurring across all four frameworks.
UnitedHealth / nH Predict (2023—)	Algorithm allegedly denied Medicare Advantage claims at ~90% error rate; physicians overridden.	EU AI Act human-oversight obligation; NIST MANAGE; NACD oversight of safety-critical AI.
Uber autonomous-vehicle fatality (2018)	Safety oversight body met twice; emergency braking deactivated; pedestrian killed.	All four frameworks. Board-designated oversight existed but was not operated.

In each case, the framework provisions existed (or would have existed had the framework been in force) that, if implemented, would have prevented the harm. The failure was not the absence of guidance; it was the absence of operational discipline tied to the guidance.

Board action items

1. **Adopt a framework triangulation strategy.** NIST AI RMF as methodology, ISO 42001 as management-system architecture, EU AI Act as binding-compliance overlay where relevant, NACD as board-process baseline. A program that picks only one is incomplete.
2. **Approve a formal AI risk-appetite statement.** Per NIST GOVERN 2.3, this is senior leadership's responsibility.
3. **Require an AI inventory at the next meeting.** If management cannot produce one, the inventory is the first deliverable.

4. **Resolve the committee-structure question.** Distributed, dedicated, or risk-integrated. Decide and update charters.
5. **Assess director AI literacy.** Identify at least one director with deep AI knowledge (background or formal advisory relationship). Build a development plan for the rest of the board.
6. **Eliminate "circular governance."** Verify that the parties briefing the board on AI risk are not the same parties whose deployments would be slowed by negative findings.

PART FOUR

IV

The Work

The chapters that distill the report into action. Workforce navigation is where the stakeholder thesis is most directly tested. The board's playbook is what the diagnosis is for.

CHAPTER 6

Workforce navigation

> *This is where the stakeholder thesis meets its hardest test. The case for taking employees seriously as stakeholders is easy when the AI deployment expands the workforce. The case becomes load-bearing when the deployment contracts it.*
>
> — Brian R. Miller

What the evidence shows so far

The pre-2026 forecasts of AI's labor-market effects ranged widely. The Goldman Sachs March 2023 estimate placed up to 300 million jobs globally as "exposed" to AI automation. The OECD 2023 *Employment Outlook* placed roughly 27 percent of jobs in OECD member countries in the highest-automation-risk category. The IMF's 2024 staff discussion note placed roughly 40 percent of global employment in the AI-exposure category.

The 2025-2026 realized data is meaningfully different from the most enthusiastic forecasts. Three observations are now well-established.

First, exposure is not displacement. The Goldman-style headline numbers measure how many jobs contain at least one task that AI can perform; they do not measure how many jobs disappear. The historical pattern across automation waves is that exposed jobs are reshaped more often than they are eliminated.

Second, the displacement is concentrated at the entry of the career ladder. The 2024-2026 hiring data shows a pronounced softening at junior roles in software engineering, legal practice, customer service, content production, and certain categories of marketing. The data does not yet show meaningful displacement at senior levels of those professions. The long-run consequence for talent supply is meaningfully different and worse — **the senior practitioners of 2040 are not being trained in the early-career roles of 2026.**

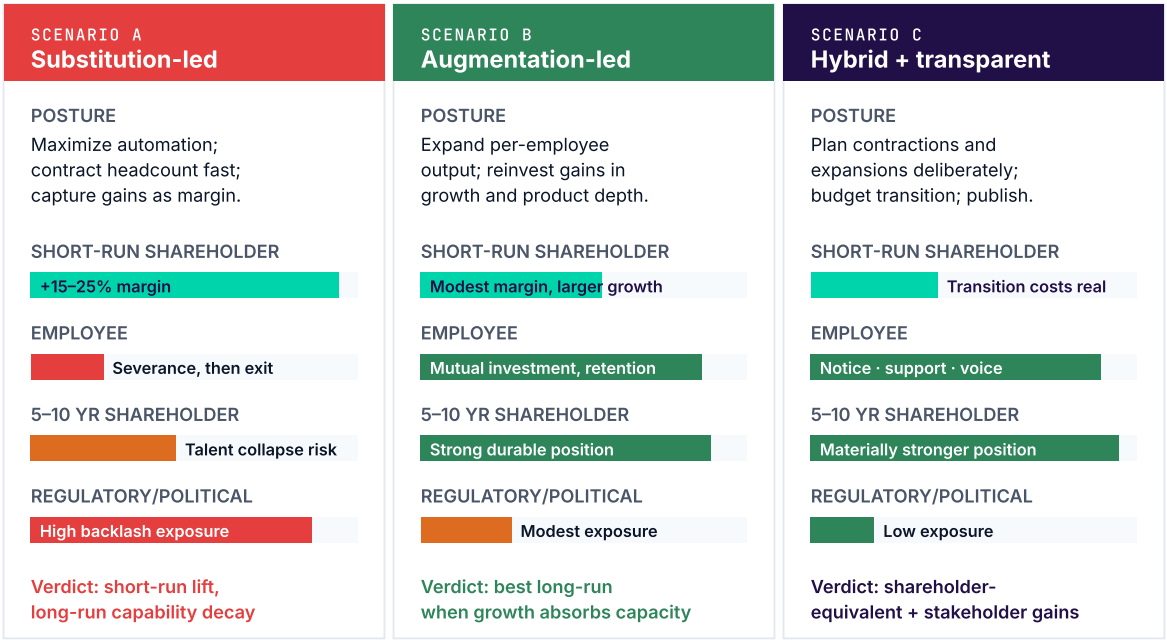
Third, the productivity gains are real but smaller than the most optimistic forecasts. Enterprise studies through 2026 consistently find productivity gains in the 20-60 percent range for tasks within AI's frontier of competence; the gains drop substantially when the same workflows are run across the full range of tasks knowledge workers actually face.

Three transition scenarios

The board's planning posture should distinguish three transition scenarios. They are not predictions; they are the meaningful range of outcomes against which the company's workforce plan should be tested.

FIGURE 4 · THREE WORKFORCE-TRANSITION SCENARIOS

Stakeholder math across short-run and five-to-ten-year horizons



Source: Author analysis. Margin figures from MIT-BCG enterprise AI productivity studies 2023–2026 (composite range).

The three scenarios are not equivalent under the stakeholder thesis. Scenario A maximizes a single stakeholder on a short-run horizon. Scenarios B and C produce shareholder outcomes that are *equivalent or better* on a five-to-ten-year horizon, while producing materially better employee, customer, and

community outcomes. The thesis defended in Chapter 4 claims that the apparent tradeoff between shareholder and other-stakeholder interests is illusory under the right horizon; the workforce question is the case where the claim is most directly tested.

Reskilling that works vs. reskilling that performs

A substantial share of the corporate reskilling commitments announced in 2023-2026 are at risk of being identified as performance rather than substance. The pattern is recognizable.

RESKILLING THAT PERFORMS

Announced rather than budgeted. Dollar amounts small relative to the workforce changes it is meant to address. Produces certificates rather than placements. Certificates in skills with weak labor-market demand. Completion rate unmeasured. Participants predominantly those whose roles were not at risk. Run by communications functions, not talent functions.

RESKILLING THAT WORKS

Funded at a level that approaches the cost of the workforce reductions it is intended to enable. Targets skills with documented internal demand or external placement-rate evidence. Produces measured placement outcomes — internal redeployment, external placement, salary preservation. Operated by talent and workforce-development functions with internal authority. Reported to the board on the same cadence as financial results.

The board's diagnostic for reskilling is straightforward: ask management to report, by cohort, *what happened to the employees in scope* one and two years after the program was offered. If the answer is not available, the program is not measured. If the answer is available and is poor, the program is not working. If the answer is available and is strong, the program is a model worth replicating.

Displacement done with dignity

There will be cohorts of employees whose roles cannot be successfully reshaped or redeployed. For those cohorts, the question shifts from reskilling to transition — and the manner of transition becomes the principal stakeholder question.

The moral-realist tradition the report draws upon treats this directly. The conviction that every employee bears inherent dignity — as a person, not as a productivity unit — produces a specific decision rule. The rule has three operational components.

Notice

Employees are entitled to know what is happening to their roles, on what timeline, and what the company's plan is for them, with enough time to act on the information. Notice periods that produce maximum surprise to the employee and minimum legal exposure to the company are not consistent with the dignity framing.

Support

Employees in transition are entitled to material support — severance proportional to tenure, healthcare continuity, outplacement, retraining funding, retention bonuses where required to complete the transition, and the network of internal connections that helps with external placement. The support is costly, and *the cost is the point* — costless severance is the costless-virtue test in its workforce form.

Voice

Employees and their representatives are entitled to a voice in how the transition is conducted. Where representation does not exist, the company can stand up employee advisory mechanisms specifically for the transition — focus groups, listening sessions, an ombudsperson.

> *The three components — notice, support, voice — are exactly the practices that produce the best outcomes on the conventional metrics. The dignity framing and the strategic framing converge.*
>
> — The convergence argument

Pace and transparency

Two operational variables shape every workforce transition: how fast it happens and how openly it is communicated.

On **pace**, the temptation is to move as fast as the automation allows. This is the wrong reference. The right reference is the pace at which the company can actually execute the redeployment and transition machinery. Companies that have run AI-era workforce transitions at the speed automation allows but faster than the support machinery can absorb have produced consistently poor outcomes — both for the employees affected and for the companies' subsequent operating performance.

On **transparency**, the temptation is to communicate as late as possible and as narrowly as possible. The empirical pattern is that transitions communicated late produce *more* rather than less disruption (rumor fills the vacuum), *worse* rather than better retention of the targeted institutional knowledge (departing employees disengage as soon as they recognize the pattern), and *more* rather than less political and reputational cost (the late communication is read as deception).

Build vs. buy vs. partner — the workforce lens

The technical question of whether to build AI capability internally, buy AI products from vendors, or partner with specialized firms has a workforce dimension that boards underweight.

Build preserves and develops internal capability. The employees who learn to design, train, deploy, and operate AI systems become the next generation of the company's technical leadership. **Buy** transfers capability development to vendors; time-to-deployment is faster but the company's longer-run capability depends on whether the buy decision is paired with sufficient internal expertise. **Partner** combines elements of both, with workforce implications that depend on the contractual structure — partnerships that develop joint capability produce different outcomes than partnerships that outsource AI capability indefinitely.

Board action items

1. **Require a workforce-transition plan with each major AI deployment.** No deployment that materially changes the staffing profile of any function should be approved without an accompanying transition plan covering pace, reskilling investment, support for non-redeployable cohorts, and measurement.
2. **Set the reskilling-investment expectation.** Reskilling that performs is dilutive of the very stakeholders it is meant to serve.
3. **Require placement-outcome reporting.** For each reskilling cohort, management should report what happened to participants twelve and twenty-four months after the program.
4. **Adopt notice / support / voice as the transition discipline.** The dignity framing's three components should be required, costed, and reported.
5. **Interrogate build / buy / partner decisions through the workforce lens.** The board's review of AI strategy should include explicit consideration of how each major choice shapes the company's internal capability.

CHAPTER 7

A board's playbook

> *Diagnosis was the first six chapters. This is action: the decisions the board should make, the order in which to make them, and the operating disciplines that should persist once the initial decisions are taken. Read this chapter as a checklist a chair can use at the next board meeting.*
>
> — Brian R. Miller

Where AI belongs on the agenda

The first decision is structural: where in the board's regular operating cadence does AI live? Three structural options are available and each is defensible in the right circumstance.

Option	Structure	Best fit	Risk
A · Distributed	Existing committees take AI per domain: audit (risk / cyber / privacy), comp (workforce), nom-gov (director skills).	AI integrated across multiple functions; existing charters updatable.	Fragmentation if charters not updated with precision.
B · Dedicated	Standing technology or AI committee with primary responsibility.	Extensive or high-risk AI deployments (financial services, healthcare, autonomous systems).	Becomes isolated specialist function disconnected from rest of board.
C · Risk-integrated	AI oversight folded into existing risk committee with education and escalation protocol.	Mature risk-management discipline already in place.	Risk committee overloaded; AI-specific issues lose visibility.

The wrong answer — applicable to every company — is ambiguity. Boards that have not decided which committee owns what will produce predictable failures: items that fall between committees and are governed by none, briefings that duplicate without aggregating, accountability that diffuses. *The decision is less important than the decisiveness.*

The management report card

For any committee or full-board cadence to function, management must produce reporting that is fit for the board's purpose. The following items should appear in the AI management report, on a cadence appropriate to the company's exposure.

Reporting item	Cadence	Function
Current AI system inventory incl. 3rd-party-embedded	Quarterly minimum	Foundational artifact for everything else
Each deployment's risk-tier classification	At deployment; on change	Regulatory readiness, board visibility
Active AI risks above appetite, with owner	Quarterly	Risk-appetite operationalization
AI-related incidents (bias, hallucination, security)	Within days, aggregated quarterly	Incident pattern recognition
AI productivity metrics by function	Quarterly	Connection between investment and value capture
AI capital and operating spend vs. plan	Quarterly	Standard capital-allocation oversight
Workforce transition status by affected cohort	Quarterly during transition	Stakeholder-thesis operationalization
Regulatory and litigation environment changes	Monthly during active phases	Environmental scanning
Vendor portfolio status — concentration, terms	Annual minimum	Vendor-risk management

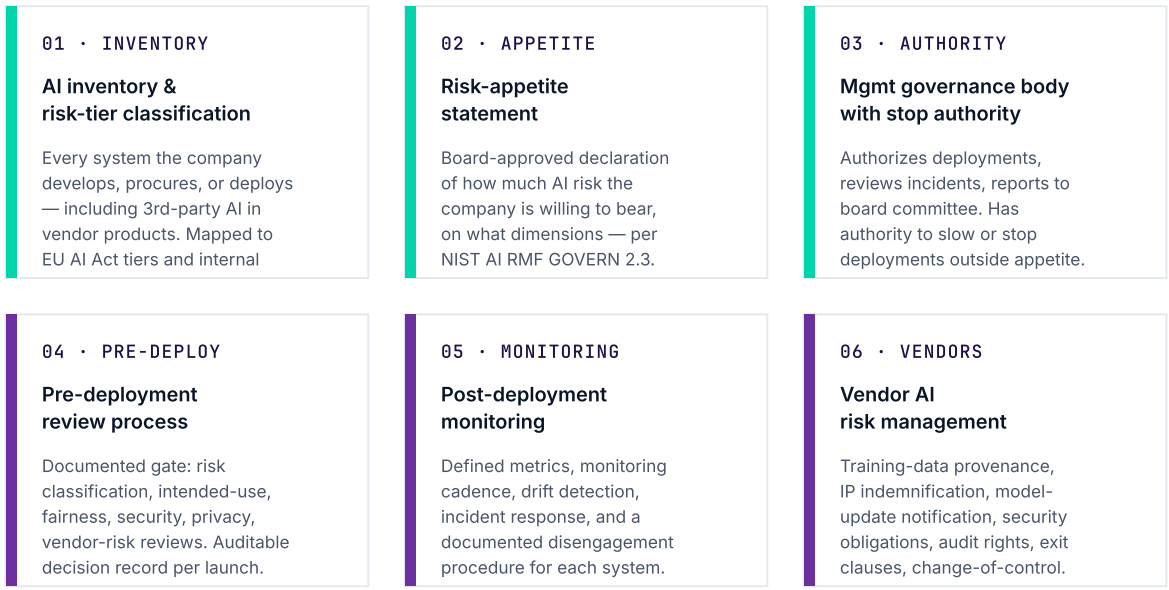
Reporting item	Cadence	Function
External assurance status (ISO 42001, etc.)	Annual	Independent assurance

The six management mechanisms

Six operating mechanisms should be in place — or scheduled — by the end of the current fiscal year. The board's job is not to build them; it is to require management to build them and to verify that they exist and function.

FIGURE 5 · THE SIX MANAGEMENT MECHANISMS

Operating machinery to stand up by end of current fiscal year; board verifies existence and function



Source: Author synthesis. Mechanisms 01–03 (teal): foundational. Mechanisms 04–06 (purple): operational discipline.

A board that, twelve months from now, can confirm that all six mechanisms exist and operate — with documentation, owners, and reporting — has built the floor. The ceiling is higher; the floor is the precondition for further work.

Pivot triggers

The thesis claims that the governance machinery must pivot as lessons land. Six signals should trigger a pivot.

1. **Regulatory environment change** — EU AI Act enforcement intensity over 2026-2028, US federal legislation if it materializes, state laws, industry-specific rules.
2. **Incident** — an AI-related event at the company or a publicly-reported event at a peer that reveals a failure mode the company's machinery is not currently designed to catch.
3. **Vendor-portfolio change** — pricing change, change of control, public failure, new entrant.
4. **Capability shift** — a new model class, a new agentic architecture, a meaningful reasoning advance that changes what the company could deploy.
5. **Workforce signal** — disproportionate departures, recruitment challenges, internal feedback indicating poor transition experience.
6. **Stakeholder signal** — customer, investor, regulator, or community feedback that the company's AI posture is producing outcomes that do not match its stated commitments.

The board should require management to monitor each of these signals and to report — proactively — when a trigger has fired. A board that learns about a pivot trigger from the financial press has lost a step.

Director literacy

A board does not need to make every director an AI expert. It does need to develop a literacy floor that allows the board to interrogate management's claims credibly.

Every director should understand: the categories of AI and what each does well and badly; the principal frameworks (NIST, EU AI Act, ISO 42001, NACD) at a working level; the company's own AI inventory and its highest-risk deployments; the regulatory environment most directly affecting the company. The investment is on the order of 10-20 hours per year.

At least one director should have deeper AI competence — either through direct background or through a formal advisory relationship with a credible external expert. This director is the board's first call when a deployment, an incident, or a regulatory question requires informed judgment.

The board collectively should have access to *independent* AI advisory capacity — retained advisors, internal-audit support, or formal third-party reviews — that does not report through management. A board whose only AI advisors report through the management it is overseeing is in the "circular governance" condition.

The first 90 days — a chair's checklist

The following sequence is intended to be usable directly. A chair who works through it will have moved the board materially forward.

FIGURE 6 · 90-DAY CHAIR'S PLAYBOOK

From "AI is on the agenda" to "AI governance machinery is operational" in three sprints

Source: Author. Adapted from NACD 2024 Blue Ribbon Commission and NIST AI RMF GOVERN function (subcategory 2.3 risk-appetite declarat

A chair who completes this list in 90 days has moved the board from “the AI conversation is on the agenda” to “AI governance machinery is in place.” The work that remains is the work the machinery now exists to execute.

The irreducible seven

The action items at the close of each of the prior six chapters produced approximately thirty individual recommendations. This Section reduces them to the seven the board should treat as non-negotiable.

THE IRREDUCIBLE SEVEN

1. **Establish the AI inventory.** Without it, no other action is grounded.
2. **Approve the AI risk-appetite statement.**
3. **Resolve the committee-structure question and update charters.**
4. **Adopt the four-framework triangulation** (NIST · ISO 42001 · EU AI Act · NACD).
5. **Stand up the six management mechanisms** with owners and dates.
6. **Adopt the management report card** and the reporting cadence.
7. **Operate the costless-virtue test continuously.**

The seven recommendations are the operating discipline of a board that has taken the rest of the report seriously. They will not be sufficient — the AI paradigm will produce questions and incidents that no preset framework anticipates. But they will produce a board posture that can govern under whatever shape the paradigm takes next, and a company posture that produces broadly distributed stakeholder value as it does so.

CLOSING

Conclusion

The diagnosis. The framework. The thesis. The frameworks. The playbook. What remains is the work.

The report's **diagnosis** was that AI is a paradigm shift of the same magnitude as electrification and the internet, and that boards governing it as a procurement decision or as a research project are operating from the wrong frame.

The report's **framework** was Carlota Perez's installation-vs-deployment distinction. The installation phase concentrates wealth among the infrastructure builders, often in a speculative-finance bubble. The deployment phase is where the productivity gains diffuse, where the regulatory and governance settlement is negotiated, and where the long-run distribution of stakeholder value is decided. AI in 2026 is in the late installation phase. The deployment phase begins in earnest over the next three to five years.

The report's **thesis** was that boards can position their companies to win in the deployment phase by building governance machinery that produces broadly distributed value rather than narrowly extracted value — and that the apparent tradeoff between shareholder and other-stakeholder interests is illusory under the right time horizon. The argument was supported on three legs: the regulatory environment is moving in the direction the thesis anticipates; the talent pool prefers employers whose stakeholder posture is substantive; and the historical pattern of paradigm shifts shows that the companies which captured the durable returns were the ones that produced broadly distributed value rather than extracted it.

The report's **framework guidance** was to triangulate across NIST AI RMF (methodology), ISO 42001 (auditable management-system architecture), EU AI Act (binding-compliance overlay where relevant), and NACD guidance (board process). Programs that adopt only one of the four are not yet at the standard.

The report's **playbook** was structural — committee ownership, risk-appetite statement, management report card, the six governance mechanisms — and disciplinary, with the costless-virtue test as the continuous operating discipline.

What remains is the work. The board that completes the playbook in Chapter 7 over the next ninety days will have done the structural part. The board that operates the costless-virtue test continuously over the next five years will have done the disciplinary part. The board that pivots when the lessons land — when the regulatory environment shifts, when an incident reveals a failure mode, when a vendor relationship changes, when a workforce signal emerges, when a stakeholder constituency tells the company that its substance and its statements do not match — will have done the part that matters most.

> *The board that does all three will, with high probability, govern a company through the AI paradigm that produces broadly distributed value, earns the trust of the constituencies it serves, operates within the regulatory settlement as it forms, and produces the kind of long-run shareholder return that comes from durable competitive position rather than from short-run extraction.*
>
> — The closing claim

The conviction that animates the report is that the question is not whether such a company is possible. The question is whether the board chooses to govern in a way that produces it. The choice is now, and the cost of the choice is exactly what gives the commitment its weight.

That is the work. There is no shortcut around it, and no need for one.

— Brian R. Miller, Synthetic Insights LLC, May 2026

Appendix A

Glossary

Board-friendly definitions for terms used throughout the report.

Agentic AI
An AI system that plans a sequence of steps, calls tools or services to execute those steps, observes the results, and iterates toward a goal it was given in natural language. Distinguished from generative AI by the system's authority to act on the world rather than only to produce content for human use.

AI governance
The operational function within management that manages AI-specific risks — model risk, bias, hallucination, privacy, security, vendor management, deployment authorization, post-deployment monitoring. Distinguished from *corporate governance*, which is the board's oversight of management's AI-governance function.

AI inventory
A current, maintained record of every AI system the organization develops, procures, or deploys, including third-party systems embedded in vendor products.

Alignment
The technical-literature term for the difficulty of specifying an AI system's goals in a way that captures what humans actually want rather than what the words literally say.

Bias (algorithmic)
Systematic skew in an AI system's outputs that disadvantages or favors a particular group. Often inherited from training data that encoded historical patterns.

CE marking
The conformity mark applied to products that satisfy EU regulatory requirements for placement on the European Economic Area market. Under the EU AI Act, high-risk AI systems require CE marking before deployment in the EU.

Caremark doctrine
Delaware corporate-law standard under which directors can be held liable for failing to implement adequate monitoring systems for known and foreseeable risks. AI risk is increasingly understood to fall within the doctrine's scope.

Confabulation / hallucination

The phenomenon in which a generative AI system produces fluent, plausible, confident output that is not factually correct. NIST AI 600-1 uses "confabulation."

Conformity assessment

The EU AI Act process by which a provider of a high-risk AI system demonstrates that the system meets the regulation's requirements.

Deep learning

A subset of machine learning in which patterns are learned through artificial neural networks. Underlies image recognition, voice transcription, and most modern generative AI.

Deployer

Under the EU AI Act, an organization that puts an AI system into operational use. Distinguished from a *provider*.

Deployment phase

The second of the two periods in Carlota Perez's framework for technology paradigms. Characterized by maturation, correction of installation-phase financial excess, and broad diffusion of productivity gains.

Disparate impact

A legal and regulatory standard under which an outcome that disproportionately affects a protected class can produce liability even without demonstrated discriminatory intent.

Drift

The phenomenon by which an AI system's performance degrades over time as the data it processes diverges from its training data.

Explainability

The capacity of an AI system (or its operators) to articulate the reasons for a specific output in terms a human can interpret.

Foundation model

A large-scale AI model trained on broad data and intended to be adapted to many downstream tasks. Leading 2026 foundation models include the GPT, Claude, Gemini, Llama, Mistral, and DeepSeek series.

Generative AI

AI that produces new content — text, images, code, audio, video — in response to natural-language instructions.

GPAI — general-purpose AI

The EU AI Act's term for a foundation model that can perform a wide range of tasks. Subject to a specific regulatory regime separate from the high-risk system regime.

Human-in-the-loop / human oversight

The design principle under which AI outputs are reviewed by qualified humans before — or, in some cases, shortly after — taking effect.

Installation phase

The first of the two periods in Carlota Perez's framework. Characterized by infrastructure build-out, financial-capital concentration, often a speculative bubble.

ISO/IEC 42001

The international standard for AI Management Systems (AIMS), published December 2023. Provides an auditable management-system architecture, certifiable by accredited third-party bodies.

Large language model (LLM)

A neural-network AI model trained on large quantities of text to predict the next token in a sequence.

Machine learning (ML)

The broad family of techniques in which a computer system learns patterns from data rather than being explicitly programmed.

Model risk

The risk that an AI or other quantitative model produces incorrect or unintended outcomes, leading to financial loss, regulatory exposure, or reputational harm.

NACD

The National Association of Corporate Directors. Producer of the most-cited US board-facing AI governance literature.

NIST AI RMF

The NIST AI Risk Management Framework. Voluntary US framework published January 2023, extended by the Generative AI Profile (AI 600-1) published July 2024.

Notified body

An EU-designated third-party conformity-assessment organization, authorized to assess specific categories of high-risk AI systems for EU AI Act compliance.

Open weights

AI models whose trained parameters are released publicly, typically under a license that permits commercial use. The Llama, Mistral, and DeepSeek model families are

open-weights as of 2026.

Provider

Under the EU AI Act, an organization that develops an AI system or has it developed, and places it on the EU market under its own name.

Retrieval-augmented generation (RAG)

A pattern in which a generative AI system is constrained to answer from a defined corpus of documents that are retrieved at query time, with citations to the source passages.

Risk appetite

A formal declaration by an organization's senior leadership of the level and nature of risk the organization is willing to accept. NIST AI RMF GOVERN 2.3 makes this a board-level responsibility for AI risk.

Schumpeterian creative destruction

Joseph Schumpeter's mid-twentieth-century concept describing capitalism as a process in which entrepreneurial innovation continually destroys the structures of the previous economic order even as it builds the next.

Systemic risk (GPAI)

Under the EU AI Act, the classification applied to general-purpose AI models with capabilities that may pose risks at the EU level. Currently presumed for models trained using computational resources exceeding 10^{25} floating-point operations.

Token

The unit of text a language model processes. AI compute and pricing are commonly denominated per token.

Trustworthy AI

The umbrella term used across multilateral AI frameworks (OECD, EU, NIST, ISO) for AI systems that are safe, secure, fair, transparent, accountable, and respectful of fundamental rights.

Appendix B

Sources & methodology

Editorial integrity statement

This report:

- Cites primary sources for every quantitative claim. Where a claim could not be sourced to a primary record, it is qualified or omitted.
- Names specific cases rather than generalizing from anonymous incidents.
- Distinguishes well-established findings from contested ones in the body text.
- Is vendor-neutral by editorial commitment. No specific AI vendor, product, or consulting firm is recommended.
- Is not legal, accounting, regulatory, or specific compliance advice.

Method

The report was developed by drawing on the author's twenty-plus years of experience as a security and technology executive; surveying internal Synthetic Insights research on AI ethics, agentic systems, and governance; reviewing primary published sources for the four governance frameworks treated in Chapter 5; surveying contemporary corporate-AI failure cases as documented in legal proceedings, regulatory investigations, and first-tier journalism; and triangulating against the Schumpeter / Carlota Perez literature on technological paradigms and the recent academic and policy literature on AI labor effects.

Primary sources

Governance frameworks (Chapter 5)

- NIST AI 100-1, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, January 2023.
- NIST AI 600-1, *AI Risk Management Framework: Generative Artificial Intelligence Profile*, July 2024.

- Regulation (EU) 2024/1689 (EU AI Act), Official Journal of the European Union.
- ISO/IEC 42001:2023, *Artificial intelligence — Management system*, December 2023.
- OECD, *AI Principles*, May 2024 update.
- White House OSTP, *Blueprint for an AI Bill of Rights*, October 2022.
- Executive Order 14110, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*, October 2023.
- NACD, *2024 Blue Ribbon Commission Report: Technology Leadership in the Boardroom*, October 2024.
- NACD, *Director Essentials: Implementing AI Governance*, 2024-2025.
- NACD, *2025 Public Company Board Practices and Oversight Survey: AI Analysis*.

Historical paradigms (Chapter 3)

- Schumpeter, J., *Capitalism, Socialism and Democracy*, Harper & Brothers, 1942.
- Perez, C., *Technological Revolutions and Financial Capital*, Edward Elgar, 2002.
- David, P.A., "The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox," *American Economic Review* 80(2), 1990.
- Odlyzko, A., "Collective hallucinations and inefficient markets: The British Railway Mania of the 1840s," 2010.
- Brynjolfsson, E., and McAfee, A., *The Second Machine Age*, W.W. Norton, 2014.

Workforce & economic (Chapters 2 and 6)

- Briggs, J., and Kodnani, D., "The Potentially Large Effects of Artificial Intelligence on Economic Growth," Goldman Sachs Global Economics, March 2023.
- Cazzaniga, M., et al., "Gen-AI: Artificial Intelligence and the Future of Work," IMF Staff Discussion Note SDN/2024/001, January 2024.
- OECD, *Employment Outlook 2023: Artificial Intelligence and the Labour Market*.
- Acemoglu, D., and Restrepo, P., "Automation and New Tasks," *Journal of Economic Perspectives* 33(2), 2019.
- Autor, D., Mindell, D., and Reynolds, E., *The Work of the Future*, MIT Press, 2022.
- Dell'Acqua, F. et al., "Navigating the Jagged Technological Frontier," HBS Working Paper 24-013, September 2023.

Cautionary cases (referenced across chapters)

- *Moffatt v. Air Canada*, 2024 BCCRT 149.
- Samsung ChatGPT data exfiltration, Bloomberg / Gizmodo reporting, April 2023.
- Amazon recruiting tool, Reuters reporting, October 2018.
- Apple Card credit-limit disparity, NYDFS investigation, 2019-2021.
- UnitedHealth / Optum nH Predict class-action complaint, November 2023.
- NTSB Investigation Report, Uber autonomous-vehicle fatality, Tempe, Arizona, 2018.
- *New York Times v. OpenAI*, S.D.N.Y., 2023-ongoing.
- *Mobley v. Workday*, N.D. Cal., 2024-ongoing.

Limitations

Time-bounded. Current as of May 2026. AI capability, regulation, and corporate governance practice are all moving variables. Sections most exposed to staleness are Chapter 2 (trajectory), Chapter 5 (governance frameworks, particularly EU AI Act phasing dates and US federal policy posture), and the case studies (where ongoing litigation may change the legal posture).

US-centric framing. While the report addresses the EU AI Act in depth and references OECD multilateral norms, its primary audience is US-based corporate directors. Boards operating from non-US jurisdictions should supplement with their local regulatory and governance literature.

Industry-agnostic. The report is written for boards generally. Specific industries (financial services, healthcare, defense, education) have AI governance considerations the report does not develop.

Vendor-neutral. The report deliberately does not evaluate specific AI products, vendors, or consulting firms.

First draft. This is a v0 first edition. The author will iterate; subsequent versions may incorporate new evidence, sharpen positions on contested questions, and expand the case studies.

Appendix C

Board discussion questions

Forty questions organized by committee likely to own the relevant work. Insert directly into board packages. A board that works through these across two to three meetings will have substantially exercised the governance discipline the report recommends.

FOR THE FULL BOARD

1 Does our company have a current, maintained inventory of every AI system we develop, procure, or deploy — including third-party AI embedded in vendor products?

2 Have we formally approved an AI risk-appetite statement? When was it last reviewed?

3 Which committee of this board has primary ownership of AI oversight? Is it documented in the committee's charter? Are there ownership gaps at the seams?

4 At least one director should have deep AI knowledge or a formal advisory relationship. Who is that director, and when was the relationship last refreshed?

5 What is the board's AI literacy development plan for the directors who are not the AI-specialist director?

6 When was the last time the company's AI strategy was presented as a substantive — not informational — item? What decisions did that presentation produce?

7 Are we positioned for the EU AI Act enforcement timeline? Specifically, have we assessed whether any of our AI systems' outputs are used in the European Union in ways that bring us within Article 2 jurisdiction?

8 Does the company have a documented exit / disengagement procedure for every material AI deployment? For deployments where the answer is "no," should the deployment continue?

FOR THE AUDIT COMMITTEE

1. How does AI risk integrate into the enterprise risk management framework?
2. What independent assurance exists over the company's AI deployments?
3. Have we received an assessment of the company's training-data IP exposure, including indemnification from AI vendors?
4. What is the company's incident-response framework for AI-related incidents? When was it last exercised?
5. Are the AI productivity metrics management reports reconcilable to financial results?
6. Does the company carry adequate insurance for AI-related liability?
7. What is the company's exposure to NIST AI RMF GOVERN 2.3 (risk-tolerance declaration)?
8. Has the company conducted a structured assessment against NIST AI 600-1 (Generative AI Profile)?

FOR THE COMPENSATION COMMITTEE

1. Are management's incentive structures aligned with the company's stakeholder posture on AI?
2. Is the cost of the AI-related workforce transition budgeted proportional to the workforce changes it is meant to enable?
3. What is the company's policy on retention of the senior practitioners whose roles are being reshaped by AI?
4. How is the AI-specialist talent paid relative to the rest of the organization?

FOR THE NOMINATING AND GOVERNANCE COMMITTEE

1. Does the board's skills matrix include AI literacy as a competency?

2 What is the board's plan for refreshing AI literacy over the next three to five years?

3 Are committee charters current with the AI oversight responsibilities the board has assigned?

4 Is there a structured channel by which the board hears from internal technology leaders directly?

FOR THE RISK COMMITTEE

1 What is the company's exposure under the EU AI Act? Have we engaged outside EU regulatory counsel?

2 Have we conducted vendor-concentration analysis for our AI vendor portfolio?

3 What is the post-deployment monitoring discipline for the company's highest-risk AI deployments?

4 Are the company's AI-related cybersecurity controls current with the threat environment?

FOR AN AI OR TECHNOLOGY COMMITTEE

1 What deployments has the AI governance committee authorized in the last quarter? Were any rejected or slowed?

2 What are the three to five highest-risk AI deployments in the company today? When did the board last review each?

3 What is the company's posture on the build-vs-buy-vs-partner question for AI capability?

4 Has the company implemented ISO/IEC 42001 certification? If not, what is the rationale?

ON THE STAKEHOLDER THESIS

1 For each AI-related stakeholder commitment, what specific decision has the company made differently because the commitment is in place?

2 Are there AI-related stakeholder commitments the company has made that have not yet produced any specific operational decision? Should those commitments be reframed or withdrawn before they fail the costless-virtue test publicly?

3 What is the company's posture on AI-related stakeholder constituencies whose interests we have not yet heard from?

ON THE WORKFORCE QUESTION

1 For each material AI deployment that affects headcount, is there a workforce-transition plan?

2 For each reskilling cohort completed in the last twelve months, what placement outcomes can management report?

3 Have we adopted notice / support / voice as the discipline for workforce transitions whose cohorts cannot be redeployed?

PIVOT TRIGGERS

1 What pivot trigger has fired in the last six months that we have not acted on? If none, are we confident we are looking for triggers, or only confident that we have not noticed any?

2 What is the cadence at which the board reviews whether the AI governance machinery itself needs to pivot?

The questions above are starting points, not a complete inquiry. A board that has answered them well will have substantially advanced the governance discipline; a board that has answered them poorly should treat the gaps as the agenda for the next twelve months.

FROM THE REPORT

The board's job in the AI paradigm is to build **serious governance machinery** *that pivots as lessons land — not to parrot the issue of the day.*

Most board-level AI guidance is either technologist-written and assumes a fluency boards do not yet have, or consultant-written and frames AI as a checklist of capabilities. This report sits in a different place. It treats AI as a paradigm shift in the Schumpeter / Carlota Perez sense — the latest after railroads, electricity, IT, and mobile — and reasons forward from what boards learned (and failed to learn) across those prior transitions.

Inside: an AI primer for non-technical directors; honest forecasts at one-, three-, five-, and ten-year horizons; the disruption-history pattern that recurs at every paradigm; a full articulation of the stakeholder-value thesis; plain-English treatment of NIST AI RMF, the EU AI Act, ISO 42001, and NACD; a workforce-navigation chapter; and a 90-day playbook a chair can use at the next board meeting.

SI NEWS™

BRIAN R. MILLER · SYNTHETIC INSIGHTS LLC

EDITORIAL: SI NEWS™

si-news.ai

REPORT CODE

SI-NEWS-AIGOV-2026-01

EDITION

SI News v0

MAY 2026