Artificial Intelligence, Human Accountability

What Boards and Executives Must Oversee as AI Takes On the Work

Cliff Justice, Founder, Catalyq · 10 November 2026

Key takeaways

1. **The operating model is inverting.** Agents do more of the work and people govern it. Most agents are still in pilots, so this is the time to set the standard.
2. **Agents change the economics of the business.** Costs shift from people to consumption, and the savings go to whoever the process design, the contracts and the prices favor. Management should show where they land.
3. **Companies have governed delegated work before.** The outsourcing discipline applies: keep what sets the company apart, define roles and service levels, document the work, and keep a tested way to take it back.
4. **Accountability follows control.** One executive owns AI, one board committee oversees it, and every agent has a named business owner.
5. **Watch the money as closely as the risk.** AI spending needs one owner and a full view, and contracts should not lock in today's prices.
6. **Keep the people who can check the work.** Expertise fades when automation does the work. Keep skills on purpose, and reserve for people the decisions that need judgment.

From people enabled by technology to technology governed by people

Most operating models a board oversees today were built on the assumption that people do the work and technology helps them. Controls, management structures and job design all follow from it. A person performs each step, another checks it, and a manager answers for the result.

AI agents reverse that arrangement. An agent can reconcile accounts, triage claims, update records and draft filings across many steps without a person touching each one. People set the goals and limits, check the output, and step in when something goes wrong.

An agent still acts within limits people define: the tools it may use, the data it may see, the actions it may take without approval, and the point at which it stops. Because people set those limits, people can govern the agent, and the company has to design that governance deliberately.

	People enabled by technology	**Technology governed by people**
Who does the work	People, with software as a tool	Agents, with people setting goals and limits
What managers supervise	People and their output	Agents, their limits and their results
How control works	One person does, another checks	Rules enforced where agents act, records of every action, tests after every change
What the workforce needs	Skill at doing the task	Judgment about the task, and skill at checking and redesigning it
Where failures come from	Individual error, usually contained	One rule or model change, repeated across thousands of transactions

Adoption is running ahead of governance. Forrester finds that most enterprises are pursuing agents while few have reached production at scale, and Gartner predicts that more than 40% of agentic AI projects will be canceled by the end of 2027, citing rising costs, unclear value and weak risk controls. With most agents still in pilots, boards have a window to set the standard.

The last row of the table is the one boards should weigh most. A person's mistake usually affects the cases that person handled. An agent that applies the wrong rule can repeat the mistake on every transaction until someone notices.

Oversight starts with strategy. The first question for a board is which of the company's competitive advantages AI is likely to erode first, as AI-native competitors and even the model makers themselves move into its markets, and what that means for its business model and operating model. The next section looks at how agents change the company's economics. The rest of the paper is about governing the agents a company deploys in response, and the closing section returns to strategy.

How agents change the economics of the business

Agents change what the business costs to run, and how that cost behaves.

Costs move from people to consumption. A process run by people has a cost that moves in steps: the company hires, trains and pays a team whether volume is high or low. A process run by agents has a cost that moves with use. Model providers charge by the token, the units of text a model reads and writes, so every task carries its own price. The unit that matters becomes the cost per completed task.

Prices fall while bills can rise. According to Epoch AI, the cost of a given level of AI performance has fallen about 13 times a year since 2023, with wide variation by task. Usage can grow faster. Anthropic reports that its agents use about four times the tokens of a chat, and its multi-agent research system about fifteen times. As agents take on longer, multi-step work, a company can pay far less per token and still see its AI bill climb, so the way the work is designed sets much of what AI costs.

The savings come from redesign. Automating a few steps of a process built around people can leave its cost where it was. Work waits at the handoffs between agents and people, and the team that remains has less to do while the cost holds. The companies that capture the savings redesign the whole process around what agents do well and what people must still decide.

Competitors start from a different cost base. AI-native competitors build their operations around agents from the start and can price below incumbents that carry more people for the same work. Providers face the same pressure: Persistent Systems' chief executive, Sandeep Kalra, told Reuters that clients now want the same work for 25% to 30% less. A company's prices, margins and service levels will be measured against those rivals.

Someone keeps the savings. The gain from agents can go to the company, to its customers through lower prices, to its providers through contracts priced for people, or to model vendors through consumption, and some of it can fund growth. Which of these happens is a strategic choice, and it should be made on purpose.

The business case is a moving target. What agents can do is changing faster than most planning cycles. METR, an independent research group, estimates that the length of software tasks AI agents can complete about half the time has doubled about every four months since 2023. Each gain changes the case. Work that was out of reach six months earlier becomes possible, companies hand agents longer tasks and run several together, and token volume rises with them.

Three numbers move at once: what agents can do, what each token costs, and how many tokens the work consumes. A case built on today's capabilities may understate what is possible next year, and a case built on today's prices may understate next year's bill. Vendors also change how they charge, moving between tokens, seats, credits and outcomes, and they retire models on their own timetable. The business case should therefore use ranges, release investment in stages as results arrive, avoid commitments that cannot be repriced, and be refreshed every quarter against measured cost per task.

The board should ask management for an economic view to go with the governance view. For the main workflows, it should show today's cost per completed task, the target after redesign, how much of the saving reaches the company, and where it will be reinvested.

Companies have done this before

The companies that stayed in control of their outsourced work kept the processes that set them apart, signed agreements with clear roles, service levels and documented processes, negotiated transition terms that made it safe to bring work back or move it, often spread work across several providers, and kept a team of their own to govern the relationship. The companies that got locked in had let that discipline lapse.

AI agents are the next transfer of work, and the same discipline applies.

What kept outsourcing clients in control	What it means for agents
Keeping the processes that set the company apart	Deciding which work agents may do and which decisions stay with people, and keeping the company's knowledge of how the work is done
Clear roles and responsibilities	A named business owner for every agent, with a written scope and decision rights
Agreed service levels	Measures for every agent: cost per completed task, quality, the business outcome, and the incident threshold that reaches the board
Clear rules and well-documented processes	Rules enforced at the moment the agent acts, and a written account of how the work is actually done, without which no one can check an agent or replace it
Strong transition terms	A tested way to take work back from an agent, to people or to another model or vendor
More than one provider	More than one model and vendor, held to one standard the company sets
A governance capability over the retained team and the providers	A team of the company's own people who govern the agents and the providers

The comparison breaks down in three places, and those are where the new risks sit:

- **An agent signs no contract and cannot be held liable.** The company answers for what it does, as the Air Canada case below shows.

- **An agent can change overnight.** A model update can alter its behavior with no renegotiation, so testing after every change is needed alongside change-control clauses.
- **Delegation now has two layers.** Providers increasingly run agents of their own on the company's work, so the company delegates to a provider that delegates to software.

What the board needs to see: one standard for every agent

Many companies already depend on agents they did not build and cannot fully see. They come in three kinds, and governance usually breaks in a different place for each.

Kind of agent	Example	Where governance usually breaks
Built by the company	An invoice-matching agent built by IT on a cloud platform	The controls live inside the vendor's platform and are lost when the company changes vendors
Built or used by employees	Assistants in office software, agents assembled in low-code tools, personal AI accounts	No inventory exists, and company data can leave through personal accounts
Run by providers	Claims intake performed by an outsourcing provider's agents	The company sees only what the contract requires the provider to report

The board should expect one standard across all three kinds, covering what an agent may do, the records of what it did, and the right to stop it and reverse its actions. Otherwise each platform may govern its own agents well while no one oversees them all.

Five things show the standard is in place:

- **An inventory.** Every agent in production, its business owner, what it may do, the systems and data it touches, and the model it uses.
- **Risk tiers.** Controls scaled to the consequence of an error. An agent that recommends a refund needs lighter controls than one that issues it.
- **Rules enforced where the agent acts.** A check at the moment an agent calls a system can refuse a payment above a limit. A rule written only into the agent's instructions can be ignored or argued around.
- **Records and reversal.** A record of every action, and a tested way to suspend an agent and undo what it did.

- **Decision rights.** A named person who approves a new agent, wider authority or a change of model, and who signs off the risk.

Accountability and money

Accountability for an agent's actions rests with a person and an organization. The working rule, used in outsourcing for decades, is that accountability follows control. Whoever has the right and the practical ability to make a decision or prevent a failure answers for the result. The rule decides who carries the cost between a company and its providers. To customers, regulators and courts, the company answers for what is done in its name, whatever its contracts say.

[embedded content: Accountability map · board, committees, executive owner, process and platform owners, internal audit, providers]

Every agent traces up this map to a named person: its business process owner, then the executive owner of AI, who reports to the chief executive and briefs the board committee that oversees AI. Internal audit tests the controls independently and reports to the audit committee. The executive owner should be one person with a budget and accountability for results, since AI run by a committee across IT, operations and risk leaves no one accountable.

AI spending needs the same oversight as any material cost. AI costs are spread across model charges, software platforms, providers' fees and the time people spend checking agents' work, so no single budget line shows the total. The board should expect one owner for the AI budget, a full view of the cost, and value reported by workflow. Long commitments to a model or platform vendor deserve the audit committee's attention.

Do not lock in today's prices. Outsourcing clients often failed to account for technology getting cheaper over the life of a contract, so the benefit went to the provider. AI prices are falling faster still. Costs can also rise. Anthropic's IPO filing warns that compute costs could rise significantly if demand accelerates, and the supply of new capacity is under pressure: in the November 2026 elections voters approved 38 local measures restricting data centers, and Goldman Sachs estimates that up to about 45% of planned U.S. data center expansion could face temporary halts. Long commitments should name the cost of AI as a pricing assumption, so the price falls when costs fall and is reviewed if they rise. Providers will price that protection in, so it is worth negotiating.

[image: The same answer, 725 times cheaper in 18 months. Cost per question to score 75% on PhD-level science questions fell from about 30 cents with OpenAI o3 in January 2025 to $0.0004 with GPT-5.6 Luna in mid-2026. Source: Epoch AI, September 2026.]

Legal and disclosure exposure

The board's duty does not move to a vendor or a model. Under Delaware's *Caremark* doctrine, directors can be liable if they utterly fail to put any reporting system in place, or if they consciously ignore the red flags such a system raises. In *Marchand v. Barnhill* and later cases, Delaware courts stressed oversight of risks that are "mission critical" to the business, and in 2023 the Court of Chancery held in the *McDonald's* case that officers owe a similar duty within their areas of responsibility. Commentators such as Pierluigi Matera argue that AI leaves the standard in place and raises the evidence a board needs to show good faith. Companies incorporated outside Delaware should take advice on their own state's standard. In an EY review of Fortune 100 disclosures, about 40% assign AI oversight to a board committee, up from 11% in 2024, most often the audit committee. Whichever committee takes it on should be clearly responsible for the whole.

The rules are still moving. The EU AI Act's 2026 amendments kept its AI literacy duty and moved most high-risk obligations to December 2027, or August 2028 for AI built into regulated products. Colorado's replacement AI law takes effect on January 1, 2027, with rights to notice, an explanation and correction when automated systems help make consequential decisions about consumers, and a growing number of state and city laws cover AI in hiring and in chatbots. Federal policy is also open. A Justice Department task force created in December 2025 is challenging state AI laws, and the bipartisan FRONTIER Act, which would set a national standard for developers of advanced models, awaits the Congress elected in November 2026. The voluntary NIST AI Risk Management Framework and ISO/IEC 42001 are the common reference points for the controls. Counsel should map which rules apply.

Case: Air Canada (2024). Air Canada's website chatbot told a grieving passenger he could claim a bereavement fare after he had traveled, against the airline's own policy. Air Canada argued that the chatbot was a separate legal entity responsible for its own actions. The tribunal rejected that argument and held the airline responsible for what its chatbot said. An agent that issues refunds or posts journal entries raises the same question with more at stake.

Agents that touch the books can be part of financial reporting controls. An agent that reconciles accounts or posts journal entries may be performing a key control that management certifies under the Sarbanes-Oxley Act and, where the external auditor attests to internal control, that the auditor tests. Changes to that agent's instructions or model then belong under the same change management as any key control. The audit committee should hear from the external auditor how it tests agent-performed controls.

What the company says about AI is a disclosure question. Risk factors and management's discussion should describe AI risks the company actually faces. The SEC has charged firms with "AI washing",

overstating their use of AI, and in January 2025 settled charges against Presto Automation, a public company, over claims about its AI product. An agent incident that exposes data may also be a cybersecurity incident with its own disclosure rules.

Agents widen the attack surface, and they test the insurance. An agent that holds passwords and access rights can be manipulated by instructions hidden in the documents, emails or web pages it reads, so agents belong in the company's cyber risk program. The board should also ask whether the company's directors and officers policy and its cyber policy cover losses caused by an agent's actions.

Most agent errors stay with management and are fixed in the normal course of work. The board's concern is the error that repeats at scale, misstates the accounts, exposes data or reaches regulators and the press. The board should agree with management in advance on the threshold that sends an incident to it, and the same threshold should route the incident to the disclosure committee. A surprisingly good result deserves the same review as a failure: an agent that approves more claims raises throughput and may also raise loss exposure.

Case: Klarna (2024 to 2025). In 2024, Klarna said its AI assistant handled two-thirds of customer service requests, the work of 700 full-time agents. In May 2025 its chief executive said cost had been "a too predominant evaluation factor" and quality had suffered, and Klarna began hiring people again so that customers could always reach a human.

The skills to keep up

Directors already know they are behind. In PwC's 2026 survey of public company directors, 71% named AI as the skill their board most needs to strengthen, more than twice the next answer, and 82% said the information they get on AI's link to risk and performance is fair, poor or unavailable.

As agents do more of the work, judgment becomes the scarce skill. A survey of 319 knowledge workers by researchers at Microsoft and Carnegie Mellon, presented at CHI 2025, found that people who trusted AI more reported applying less critical thinking to its output. The data is self-reported, but it matches a common experience: people check work well when they know how to do it themselves.

Some roles will be made redundant, at least for a time. No law requires a board to weigh AI's effect on the workforce, but governance practice treats human capital as a board matter: management runs the change, and the board should confirm that cuts follow a measured strategy with a fair plan for the people affected. Lower costs can also raise demand, as the economist William Stanley Jevons observed when more efficient use of coal raised total consumption, and that demand may in time need more people than before, in different roles with different skills. Many companies are cutting ahead of the evidence, and Forrester expects half of AI-attributed layoffs to be quietly reversed, with the jobs often

returning offshore or at lower wages. Workforce decisions carry obligations, from WARN Act notice in the United States, where New York now asks employers to disclose cuts driven by AI or automation, to works council consultation in Europe. The public is watching too: in a late-September AP-NORC poll, eight in 10 U.S. adults said government should prioritize protecting workers and keeping AI under human control. The committee that oversees human capital should see the plan before agents scale.

When AI knows more than the people overseeing it

Governing technology with people assumes the people can tell when the technology is wrong. That assumption weakens as agents improve. If an agent does the work better and faster than the staff who used to do it, those staff stop doing it, and over time they lose the expertise to check it.

Judgment takes more than knowledge. Consequential decisions are made well when the data is weighed against the company's priorities, its tolerance for risk and the interests of customers, employees and investors. AI can inform those decisions, and the people who make them remain answerable for them.

Aviation shows where the line falls. The automation on a modern airliner can fly most of a flight, including the landing, and large passenger aircraft still carry two pilots. Pilots weigh what the instruments show against their experience and their tolerance for risk, and build in margin when in doubt. They also share the fate of their passengers. Executives and directors have stakes of their own, in their reputation, their pay and, for directors, a duty of oversight that a court can test. An agent has none of these, so decision rights belong with named people. Let agents act alone where a wrong decision is cheap, and keep an experienced person in command where it is not.

In 1983 the human-factors researcher Lisanne Bainbridge described the ironies of automation in the journal *Automatica*: automating most of a task leaves operators little chance to practice, their skills fade, and yet they are expected to take over in the rare, abnormal situations where skill matters most. Two fields have since found evidence of it.

- **Aviation.** In 2013 the FAA warned airlines, in Safety Alert for Operators 13002, that continuous use of automation could weaken pilots' ability to recover an aircraft from an unexpected state, and told them to fly by hand on purpose to keep the skill.
- **Medicine.** A study at four Polish endoscopy centers, published in *The Lancet Gastroenterology & Hepatology*, found that after doctors began using AI to spot polyps, their detection rate on procedures without AI fell from 28.4% to 22.4% within months. The study is observational and does not prove cause.

Finance, claims, procurement and legal work face the same risk as agents take them over. Boards should expect management to plan for it.

1. **Keep proficiency on purpose.** Like pilots who fly by hand to keep the skill, some staff should do the work without the agent on a schedule. Junior staff need rotations and supervised practice too, because junior work is how people become experts.
2. **Verify agents with more than human review.** Where people can no longer check every output, use tests with known answers, a second model or method that checks the first, statistical sampling, and outcome measures that reveal drift. Write down how the work is done and keep it current, so agents can be tested against it and the knowledge moves with the company when it changes vendors.
3. **Reserve some decisions for people.** Decisions that weigh risk, growth and the interests of customers, employees and investors stay with people, with the agent recommending. So do consequential decisions where the company cannot verify the agent's reasoning. Decision rights should say which decisions these are.

Management should be able to show the board, process by process, who could step in if an agent failed.

Agents the company does not run

Providers already do a growing share of their clients' work with agents, in finance, customer service, claims and IT. The same standard applies, but the company's levers are the contract and the existing vendor risk program, which should cover provider agents too. The board should know which services contracts already involve agents, whether those contracts say who answers for an agent's errors, what records the company can see, and who keeps the savings when fewer people do the work.

The skills question crosses the boundary too. When providers rebadged a client's staff, boards learned to ask who was left to manage the provider. A provider may now take on the people and then replace them with agents, so the company's knowledge of the work ends up in the provider's configurations. The contract should confirm that the company keeps its data, its decision rules and the records of what agents did if the relationship ends.

Forward deployed engineers are the new onsite team. Vendors increasingly send their own engineers to build agents inside the client, and Gartner expects roughly 70% of these vendor-led engagements to be abandoned by 2028, warning of premium fees and of clients that never build skills of their own.

Plan for a model going dark. Most companies depend on a handful of model providers, and access can change overnight. In June 2026, Anthropic suspended access to two of its newest models for nearly three weeks to comply with U.S. export controls, restoring it when the controls were lifted. A company whose critical processes run on one model should know what happens if that model is withdrawn, repriced or restricted, and should have a tested fallback: another model, another provider, or people.

Governance as an advantage

Good governance also lets a company move faster. A company that knows which agents it runs, can stop and reverse them, and owns the know-how behind them can give agents more of the work with confidence, switch to cheaper or better models as prices fall, and renegotiate with providers from strength. That brings the paper back to the strategy question it opened with: a company that can trust its agents with more of the work is better placed to rebuild the advantages AI erodes.

Questions for management, and what to ask for within 90 days

1. **Which of our competitive advantages is AI likely to erode first, and what is management doing about it?** A view of how AI-native competitors and model makers could enter our markets, and what it means for our business model and operating model.
2. **Who owns AI, and who oversees it?** A named executive owner, and a proposal for which committee of this board oversees AI, written into its charter.
3. **Have we reset our risk appetite for how fast AI is moving into our sector?** An updated risk appetite statement that covers agents, approved by the committee that oversees AI.
4. **Which agents and models do we depend on, and what would happen if one were shut down?** One inventory across IT, business teams and providers, with a business owner and a risk tier for each, and a tested fallback for every critical process that relies on a single model.
5. **Can we stop an agent and undo what it did?** Evidence that the company can suspend an agent and reverse its actions, and a proposed threshold for reporting agent incidents to the board and the disclosure committee.
6. **What does AI cost us, what are we getting for it, and do our incentives reward the right results?** A full view of AI spending with one budget owner, cost per completed task for the main workflows, and how executive and team incentives reflect the value AI creates and the risks it carries.

7. **Who could step in if an agent stopped tomorrow, and who is left to manage our providers?** Confirmation that, for the three most critical processes run by agents, named people can do or check the work without them, a plan to keep that skill, and the retained team that manages each major provider.
8. **Do we keep the right to use our data, decision rules and agent records if we fall out with a provider, including vendors whose engineers build inside the company?** A list of the services contracts where a provider uses agents or plans to, with renewal dates, what each contract says about these rights on exit, and a plan to review them.
9. **When was our AI governance last independently assured, and who reviews what we say about AI in our filings?** A review by internal audit or an outside party, reported to the audit committee, and confirmation that the disclosure committee reviews every public claim about the company's use of AI.
10. **Does this board have the AI literacy its fiduciary role requires?** A schedule of AI briefings for the board, including at least one from outside the company, and a candid self-assessment of the board's own expertise.

Three companion papers in the *Human Judgment at Machine Scale* series set out how management can deliver these: *Own the Harness: An operating guide to the controls, owners and costs of AI agents*; *The New Onsite Team: What Boards Should Know About Forward Deployed Engineering*; and *Contracting with the Agentic Workforce: Who keeps the savings when your provider's AI agents do the work*.

Legal references. *In re Caremark International Inc. Derivative Litigation*, 698 A.2d 959 (Del. Ch. 1996); *Stone v. Ritter*, 911 A.2d 362 (Del. 2006); *Marchand v. Barnhill*, 212 A.3d 805 (Del. 2019); *In re McDonald's Corporation Stockholder Derivative Litigation*, 289 A.3d 343 (Del. Ch. 2023); *Moffatt v. Air Canada*, 2024 BCCRT 149.

About the author. Cliff Justice founded and led KPMG's Shared Services and Outsourcing Advisory practice. He was then KPMG's head of AI and Intelligent Automation, and later U.S. Leader of Enterprise Innovation, where he founded KPMG Studio and KPMG Ventures. He now advises boards and leadership teams on AI and is the founder of Catalyq, which reviews services contracts and offers executive briefings on these topics. The analysis of outsourcing in this paper draws on his experience advising on dozens of complex outsourcing and insourcing transactions with a combined value of more than $6 billion.

About Catalyq. Catalyq reviews services contracts and offers executive briefings on the topics in this paper, and readers should weigh its recommendations with that interest in mind.

This paper is for general information and is not legal advice. Boards should take counsel on their own duties.