

— White paper · October 2026

Agentic AI, read for the board

What eight 2026 reports measured, what none of them tells a board, and what to govern

Sam Banerjee

AI Decoded

PhD candidate, Human-Centred AI

University of Technology Sydney

SUMMARY

Executive summary

Eight reports, one reading: what they agree on, where they differ, and the three questions they leave to the board

ABSTRACT

Agentic AI, software given a goal that takes actions inside an organisation's systems, is the organising frame of the major 2026 industry reports on enterprise AI. This paper reads eight of them in full: Accenture, BCG, Deloitte, Anthropic, McKinsey, Stanford HAI, Microsoft and IBM, published between December 2025 and May 2026, with more than 25,000 respondents between them. They share four conclusions, disagree on the bottleneck, measure adoption with eight indicators that do not measure the same thing, several of them not scale at all, and leave the three questions a board must answer, appetite, authority and evidence, unaddressed. On one reader's grading of each report's leading claim as asserted, referenced or demonstrated, one of the eight is demonstrated. The paper sets out what the findings mean for boards in three situations, offers a position on what boards should govern and what they should leave to management, and describes the next stage of the research. An appendix collects the 38 questions the eight reports arm a board to ask.

SUMMARY

Agentic AI is where enterprise AI is moving, and eight of the most prominent 2026 reports say so, with more than 25,000 respondents between them. Read in full, they agree on four things, disagree on the bottleneck, measure adoption with eight indicators that do not measure the same thing, and leave the three questions a board must answer unaddressed.

Graded on the scale the author's doctoral work applies to governance claims generally, asserted, referenced or demonstrated, one of the eight headline claims is demonstrated, four are referenced and three are asserted. The word appetite does not appear in any of the eight. Two design decision rights; six describe accountability without saying who holds it. Reported AI incidents rose from 233 to 362; in a different survey the same year, executive concern about control fell seven points.

The executive layer is running on expectation. That is the reports' subject rather than their flaw. It becomes a board's flaw when the same expectations arrive in a funding paper and are approved as if they were evidence.

Disclosure. The author is the founder of AI Decoded, an AI governance advisory practice. This paper is produced as part of doctoral research at the University of Technology Sydney, supervised by Jianlong Zhou, and does not promote any product or service. Section O7 describes a method under development: a working method, not a validated instrument. Section O6 is the author's position and is not attributed to any of the sources. Every figure is as stated in the document cited; the references record what a reader should know before repeating each one.

Where the world is moving

The question is no longer what AI can tell you. It is what it can do.

For three years the question put to enterprise AI was what it could tell you. The question now is what it can do. An agent, in the sense the eight reports use, is software given a goal that takes actions inside an organisation's systems to reach it, with a person reviewing all, some or none of those actions. Every one of the eight is organised around that shift, including Stanford's measurement.

The reports say the shift is under way rather than coming. 23% of organisations use agents at least moderately today and 74% expect to within two years [3]. Active agents inside Microsoft 365 grew fifteenfold in a year [7]. The share of business requests to one major model that hand a task over entirely rose from 27% to 39% in eight months, and 86% of surveyed technical leaders deploy coding agents for production code [4]. More than 30% of 2026 AI investment is earmarked for agents [2]; 55% of operations executives are developing or deploying an agentic operating model [8]; 94% of executives say agents require rethinking the platform strategy [1].

What changes for a board is not the technology but where decisions happen. An agent moves a decision out of an approval chain and into real time. 58% of leading organisations expect decision rights to change as a result [2], and few have written down how. A board that approved a tool has, without being asked, approved a decision-maker. That is the problem this paper reads the eight reports against.

What the eight reports measured

Six executive surveys, one academic index, one practitioner article, and four things they agree on

Six are executive surveys, one is an academic index and one is a practitioner article. Only one ties its value claim to reported financial results rather than to what respondents say.

HOW THEY WERE READ

Selection. The eight are prominent 2026 reports on enterprise and agentic AI that travel together: the set was taken from a summary graphic circulating among directors in October 2026 and verified against the publishers' own sites. It is a purposive sample, not a count of what boards cite, and the respondent totals are context rather than a pooled sample; the populations and instruments differ. Two of the titles in circulation do not exist as named; they were traced to the documents they refer to, IBM's *The blueprint for agentic operations* [8] and Accenture's December 2025 platform-strategy report [1]. Two are not survey reports: McKinsey's is a ten-page article adapted from a book chapter [5] and BCG's is a web article with a slide deck [2]. One is a vendor's survey of its own customers [4].

Reading. Each document was read in full on 17 October 2026, from the published PDF where one exists and otherwise from the publisher's landing page, executive summary and chapter PDFs, including method notes, sample descriptions and footnotes. Third-party coverage was used only to confirm publication dates. Every figure in this paper is as stated in the document cited, and the reference list records what a reader should know before repeating each one.

Grading. Each report's headline claim was graded on a three-level scale, by one rule per level. **Asserted:** the claim is about an outcome, a return or the future, and the only thing behind it is the statement of the people asked, however many of them there are. **Referenced:** the claim is a measurement of a state, such as access, adoption or incidents, and the measurement is cited. **Demonstrated:** the claim is an outcome shown against a record of what happened, with the design of the comparison stated. The same survey can produce claims at two levels: Deloitte's access figures measure a state and are referenced [3]; Anthropic's returns figure reports an outcome with no definition or record behind it and is asserted [4]. The claim graded is the one each publisher puts first, on its landing page or in its executive summary, quoted in Table 3; the grade applies to that claim alone, not to every figure in the report. The claims differ in subject, from revenue growth to incident counts; the grade is of their evidential form, not their subject, and the tally compares forms. The grading was applied by one reader and has not been checked by a second.

FIGURE 1 · THE EIGHT LEADING CLAIMS ON THE SCALE; DETAIL IN TABLE 3

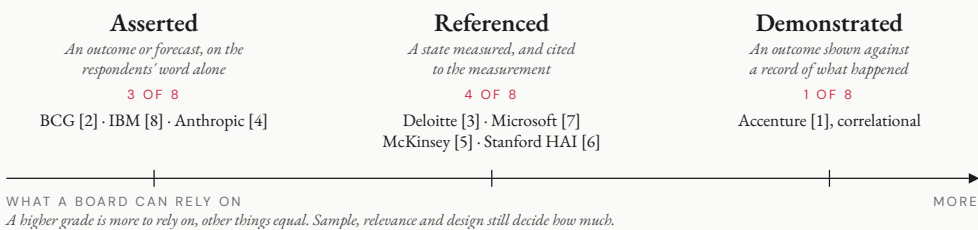


TABLE 1 · THE EIGHT REPORTS, IN PUBLICATION ORDER

#	PUBLISHER AND TITLE	PUBLISHED	SAMPLE AND FORM	WHAT IT MEASURED
1	Accenture <i>The new rules of platform strategy in the age of agentic AI</i>	18 Dec 2025	1,031 executives, 12 countries including Australia; regression on 602 companies' financials	Whether agents sit inside, across or instead of the application estate; value follows alignment of AI, platform and business strategy
2	BCG <i>AI Radar 2026: As AI investments surge, CEOs take the lead</i>	15 Jan 2026	2,360 executives including 640 CEOs, 16 markets; web article and 29-slide deck	CEO ownership, spend as a share of revenue, agent budgets; every return figure an expectation
3	Deloitte <i>State of AI in the Enterprise: The untapped edge</i>	21 Jan 2026	3,235 senior leaders, 24 countries including Australia; fieldwork August to September 2025	Access versus use, job redesign, the pilot trap, agent governance maturity, sovereign AI
4	Anthropic <i>The 2026 State of AI Agents Report</i>	c. Feb to Mar 2026	500+ US technical leaders, late 2025; the vendor's own adopters	Coding agents in production, delegation rising, self-reported returns; nothing on safeguards
5	McKinsey <i>Building the foundations for agentic AI at scale</i>	2 Apr 2026	Ten-page article from <i>Rewired</i> , second edition; adoption figures borrowed from its 2025 survey	Data architecture as the blocker; a control layer and a federated governance split
6	Stanford HAI <i>The 2026 AI Index Report</i>	13 Apr 2026	423 pages aggregating third-party data	Capability, investment, adoption, agent benchmarks, incidents, transparency, governance ownership
7	Microsoft <i>2026 Work Trend Index Annual Report</i>	5 May 2026	20,000 AI-using workers, 10 countries including Australia; Microsoft 365 telemetry	Human agency: the organisation and its managers, not individuals, determine the result
8	IBM Institute for Business Value <i>The blueprint for agentic operations</i>	12 May 2026	2,000+ operations executives, 16 countries including Australia	Decision rights in three tiers, 21 workflow blueprints; every value figure an expectation to 2027

WHAT THEY SHARE

Stripped of their branding, the eight agree on four things. First, everyone is experimenting and few have scaled: under 10% [5], 17% [1], 25% [3] and single digits [6], each by its own definition. Second, in all seven that name a blocker it sits inside the organisation rather than in the model: data, integration, jobs, silos, managers, alignment, conviction; capability is accelerating [6] and adoption of it is not. Third, governance is named in all eight and measured in two: 21% of organisations have a mature governance model for agents [3], and dedicated AI governance roles exist in 17% [6]; Microsoft measured governance maturity and did not publish the result [7]. Fourth, every value claim rests on expectation or on what respondents say of themselves, and the claims are not of the same thing. About 90% of executives expect a return from agents this year [2]; 45% efficiency gains are expected by 2027 [8]; 80% of one vendor's customers say they are seeing a return, undefined [4]; 20% report having achieved revenue growth, against the 74% who hoped to [3]. Accenture's 2.2x is the only figure tied to reported financials, and it is correlational [1].

Where they differ

Seven bottlenecks, eight indicators that do not measure the same thing, and a round number that appears in none of them

Seven of the eight name a bottleneck, each a different one, and in most cases it is the thing its publisher sells. Each also measures adoption with a different indicator, and several of the indicators are not measures of scale at all, which is why the figures cannot be averaged.

That is also why a round number borrowed from none of them, the 20% succeeding against 80% stuck in pilots that travels with these reports, has circulated so easily. It appears in none of the eight. Table 2 sets the bottlenecks and the indicators side by side.

TABLE 2 · THE BOTTLENECK EACH REPORT NAMES, AND ITS DEFINITION OF SCALE

PUBLISHER	BOTTLENECK NAMED	THE FIGURE BEHIND IT	ADOPTION OR SCALE INDICATOR USED
McKinsey [5]	Data architecture	Eight in ten cite data limitations as the roadblock to scaling agents	Under 10% have scaled agents to tangible value
Deloitte [3]	Job redesign	Access to sanctioned tools rose from under 40% to under 60% of workers; use among those with access did not move; 84% have not redesigned a job	25% have 40% or more of experiments in production
Accenture [1]	Strategy alignment across the application estate	18% have AI, platform and business strategy aligned; 57% rank integration with existing systems as the top risk	17% have AI embedded deeply in core processes
IBM [8]	Functional silos	82% say silos block the value autonomy delivers; 31% of proprietary data is reachable by AI and 13% used	55% developing or deploying an agentic operating model: intent, not scale
Microsoft [7]	Managers and culture	Organisational factors explain 67% of reported impact; 26% of workers see their leadership aligned on AI	16% of workers are frontier professionals: people, not firms
BCG [2]	CEO conviction and budget	72% of CEOs now call themselves the main AI decision maker, double the prior year	15% of CEOs are trailblazers: conviction, not results
Anthropic [4]	Integration and cost	46% and 43% of technical leaders name them as the top barriers	57% run agents on multi-stage workflows; 16% across functions
Stanford [6]	None named	88% of organisations use AI somewhere; measured productivity gains are modest; the explanation is left to the reader	Single digits have agents at scale in any function

The one-line summaries that travel with these reports fold the seven diagnoses into a single phrase, the operating model. One word for seven diagnoses is a filing system, not a finding. A paper that names the operating model as the problem has not told the board which of the seven problems it is being asked to fund.

What none of them tells a board

Appetite, authority and evidence: the three questions the reports stop before

This paper reads the eight in the order a risk committee uses for anything new it is asked to accept: where it sits in the risk framework, how much of it the organisation is willing to carry, what keeps it inside that limit, and the evidence that those controls are working. The eight cover the first step. Two design part of the third, two count whether governance roles and maturity models exist, none shows a control operating, and none reaches the second.

FIGURE 2 · THE ORDER THIS PAPER READS IN, AND WHAT THE EIGHT REPORTS COVER AT EACH STEP

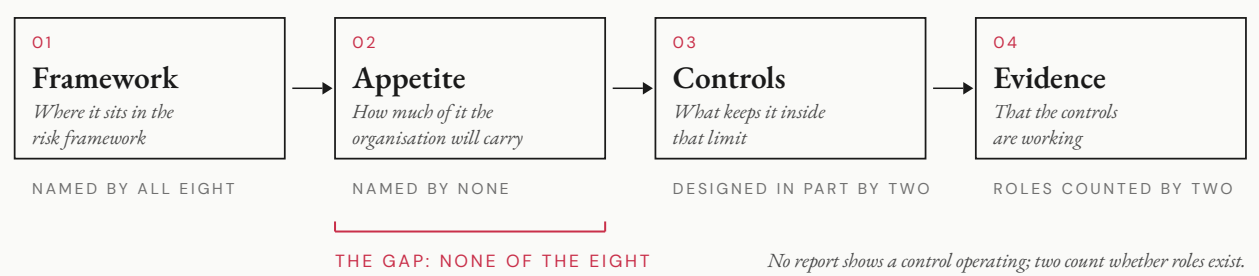


TABLE 3 · THE LEADING CLAIM OF EACH REPORT, AS THE PUBLISHER STATES IT, WHAT IT RESTS ON, AND THE GRADE

REPORT	THE CLAIM GRADED	WHAT IT RESTS ON	GRADE
Accenture [1]	The 18% with AI, platform and business strategy aligned grow revenue 2.2 times faster than their peers	A survey, plus a regression on 602 companies' reported 2024 outcomes	Demonstrated, correlational
Stanford HAI [6]	Reported AI incidents rose to 362 in 2025, from 233 in 2024	Third-party incident counts; the report's benchmark results on agents are a separate claim, and demonstrated	Referenced
Microsoft [7]	Organisational factors explain 67% of reported AI impact; individual factors 32%	A survey plus telemetry; its causal claims are associations	Referenced
Deloitte [3]	Access to sanctioned AI rose from under 40% to under 60% of workers; use among those with access did not move	A survey of current state, with one hoped-for versus achieved pair	Referenced
McKinsey [5]	Fewer than 10% have scaled agents to tangible value	A figure borrowed from McKinsey's 2025 survey; the article's prescription is a separate claim, and position	Referenced
BCG [2]	About 90% of executives expect agent ROI this year; spend to double to 1.7% of revenue	CEO intentions and expectations	Asserted
IBM [8]	Executives expect 45% efficiency gains from agentic operations by 2027	Executive expectations to 2027	Asserted

REPORT	THE CLAIM GRADED	WHAT IT RESTS ON	GRADE
Anthropic [4]	80% of surveyed leaders report returns from agents	Self-reported returns from one vendor's own customers; returns undefined	Asserted

EVIDENCE

One demonstrated, four referenced, three asserted. A board would not approve a capital model on a survey of other CFOs' optimism; it would ask for the backtest. Evidence that controls operate is a separate matter again, and the Australian example is recent. The Auditor-General's September 2026 audit of the Department of Health, Disability and Ageing found an AI framework, an audit and risk committee in active oversight and 55 AI uses on the register, and a Medicare fraud-detection system that had run an AI model on personal data from July 2024 to December 2025 not on it [9]. The framework existed. The register did not show the system that mattered. It is not an agent case, and nothing here claims the method would have caught it; it shows what a framework and an engaged committee can miss.

APPETITE

The word does not appear in any of the eight, and the concept is absent in other words too: none sets out how much of the risk a board should accept, in which decisions, or what evidence would show that a limit is holding. The two that prescribe decision rights [8] [5] prescribe a control, not a limit. Two further signals sit beside that absence, from different samples that prove nothing about each other. Reported AI incidents rose from 233 in 2024 to 362 in 2025, transparency scores for the major model developers fell from 58 to 40, and agents fail roughly one structured task in three on the benchmarks Stanford tracks [6]. Executive concern about lacking control or understanding of AI decisions fell seven points to 41%, while spend is set to double to about 1.7% of revenue [2]. The gap is in these eight reports rather than in the field. NIST's AI Risk Management Framework asks an organisation to set its risk tolerance and to name who is accountable for AI risk [10]; none of the eight connects its findings to that framework or to any equivalent.

AUTHORITY

Two of the eight design decision rights: IBM's three tiers, which send routine data-based decisions to agents, ambiguous cases to cross-domain teams and high-impact trade-offs to named workflow managers [8], and McKinsey's split between business domains, a central guardrail team and IT [5]. The other six describe accountability without saying who holds it. Where Stanford looks, AI governance sits with information security in 21% of organisations, with a dedicated role in 17%, and with no one in 5% [6].

What it means for boards

Pilots, production, strategy: three situations that coexist, and a different set of questions for each

Most boards are not yet looking at a funding paper. They are in one or more of three situations, which coexist in most organisations, and the figures and the questions differ for each. The rules below are the board's own practice, not instructions to management.

SITUATION 01 · RUNNING PILOTS

Which is most organisations

The trap. Two-thirds experiment and fewer than 10% scale [5]. 25% have moved even 40% of experiments into production, and three-month pilots stretch to 18 [3]. Pilots run on small teams, cleansed data and isolated environments; production does not.

The questions. Of the AI initiatives the board has approved, which are in production, which were stopped, and which are neither? Which of the pilot's conditions does the production case assume will hold?

The rule. Every AI approval carries a date on which it returns as in production or stopped; nothing the board approved stays a pilot by default. Above a threshold the board sets, the register of pilots, with start dates and owners, is a standing paper. No scale-up until governance for the production system exists.

SITUATION 02 · AGENTS IN PRODUCTION

Agents already acting inside the organisation's systems

The trap. Full delegation rose from 27% to 39% in eight months [4]; incidents rose from 233 to 362 [6]; 21% have mature governance for agents [3], and in 5% of organisations nobody holds it [6].

The questions. Who is accountable for the decisions this system makes without a person? Which decisions do agents take alone today, and who set that boundary? Could the board reconstruct one of those decisions after the fact?

The rule. A decision-rights map for every system in production above the board's materiality threshold: alone, with review, never, signed by the executive accountable for the top tier. An AI line in the incident report the committee already receives. The evidence the committee relies on is the record of controls operating, not the policy that says they will.

Spend is doubling before appetite is set

The trap. Spend is doubling and 94% will invest regardless of payoff [2]. 18% have AI, platform and business strategy aligned [1]; 26% of workers see their leaders aligned at all [7]. 83% rate sovereign AI important and 77% weigh a vendor's country of origin [3].

The questions. What is the organisation's AI spend as a share of revenue, and what does the board expect for it? Is the strategy tied to the business strategy and the platform estate, or a separate document? Which of the seven bottlenecks is it funding?

The rule. Appetite is set before spend: how much AI risk, in which decisions, with what limits. The strategy names its bottleneck and what fixing that one costs. One claim graded demonstrated by a milestone the board sets for the use case, twelve months being a reasonable default, or the board records that it has approved on expectation.

What to govern, and what to leave to management

Five things for the board, four for management, and why the boundary matters more than the model

Governing agentic AI is not a question of how much to trust the model. It is a question of which decisions the organisation has handed to software, who answers for them, and what record exists when one goes wrong.

My doctoral work grades the governance claims organisations make against whether any evidence exists behind them. Human-centred AI asks the companion question: when a person is kept in the loop, is the loop designed, or is it a box ticked? Read against the eight reports, both questions point to five things a board should govern and four it should leave to management. This section is my position and is not attributed to any of the sources.

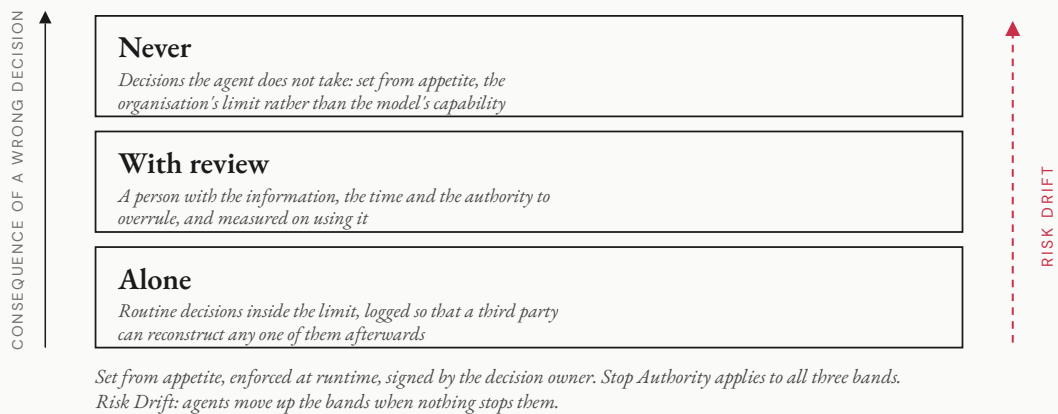
FIVE THINGS A BOARD SHOULD GOVERN

- I **Govern the decision boundary, not the model.** A model is probabilistic and will stay that way; a 95% accuracy figure says nothing about what the other 5% cost once the agent executes rather than advises. The board's lever is the boundary: which classes of decision an agent may take alone, which with review and which never, set from appetite and enforced at runtime rather than written in a policy. Agents move from low-risk to high-risk decisions when nothing stops them. I call this **Risk Drift**. It is a mechanism I propose, not one the surveys measure; the nearest signal is that 58% of leaders expect decision rights to change [2] and few have written down how.
 - II **Govern the decision owner, not the system owner.** Every AI paper names a system owner. Almost none names a decision owner: the person accountable for the decisions the system makes without a human. The first is a technology role; the second is a governance role, and the question a board will be asked afterwards is whether it took reasonable steps to make sure someone was answerable. IBM's three tiers [8] are the first consultancy model I have seen that puts a name on the top tier.
 - III **Govern the evidence, not the policy.** A policy is **Process Evidence**: a record of intent. A log of what the agent did, what it was allowed to do and who reviewed it is **System Evidence**: a record of operation. A log is evidence that something happened; a control tested against the log, with a date and a finding, is evidence that it worked. Boards are shown the first and asked to rely on it as if it were the second. For an agent acting at machine speed a quarterly review is a historical document; the record has to be continuous, and a third party has to be able to reconstruct any decision after the fact. Process is not proof.
-

IV **Govern the stop.** Someone must hold the authority and the physical means to halt an agent, and the board should know who, and whether it has ever been exercised. **Stop Authority** is the simplest control to specify, and in the AI papers I have read it is the one least often written down.

V **Design the human's job; do not assume it.** A person in the loop is not a control unless that person has the information, the time and the authority to overrule, and is measured on doing so. 86% of workers treat AI output as a starting point [7] and 84% of organisations have not redesigned the jobs those people hold [3]. A reviewer who cannot say no is a signature, not a control.

FIGURE 3 · THE DECISION BOUNDARY: THE BOARD'S LEVER



LEAVE TO MANAGEMENT

The model's internals and the vendor selection, since weights, prompts and which vendor to use are management's to choose and operate; what the board oversees is the consequence: concentration, dependency, jurisdiction, and what happens if the vendor does not survive the contract. The sandbox, since **Pilot Governance**, governance built for a pilot, cannot govern production and the board's concern is the gate between them. The feature list, since platform comparisons belong to the CIO. And the urge to ban, since 94% will keep investing regardless [2], agents are arriving inside tools already licensed [7], and a blanket ban governs nothing. Prohibiting particular uses is governance; that is what the top band in Figure 3 is for.

In the three reports that look, governance surviving production is the condition the scaled organisations share; the associations do not show that it is the cause. Accenture's aligned 18% show 2.2x the revenue growth of their peers [1]; IBM finds organisations 5.4 times more likely to adopt autonomous workflows when six capabilities, governance among them, mature together rather than one at a time [8]; Microsoft finds leadership and culture explain the result [7]. Governance set at the decision boundary is what makes the next approval defensible, and the one after it faster.

The next stage, and the limits of this one

A working method, not a validated instrument

The questions in Appendix A are free to ask, and management will answer every one of them. What a list of questions does not give a board is the level of answer each one requires, or a way to grade what comes back.

A board that has not decided in advance what an acceptable answer looks like accepts the answer it is given, and the record shows that it did.

The doctoral work takes the grading used in Section 04 and applies it to the AI funding papers boards actually receive. It treats this as a question of assurance rather than of risk assessment: a fixed set of questions every AI funding paper must answer, the level each answer must reach set before the paper arrives, and each answer graded against that level, so that directors read the grading the way they read the paper and the committee finds the problem itself. The grading standard has been applied in two practitioner sessions with directors and executives in Sydney in August and September 2026, first to governance questions the room itself wrote and then to a mock AI programme board paper. The general specification of what an acceptable answer looks like, across system types, is the subject of the doctoral work and is not published here. It is a working method, not a validated instrument.

LIMITATIONS

The paper reads eight reports, chosen because they are prominent and travel together, not because they are the best evidence available; seven were published by organisations with a commercial interest in their findings, and the eighth aggregates data that includes some of the other seven. It contains no primary data of the author's own. The grading in Section 04 was applied by one reader to each report's headline claim and has not been checked by a second reader or applied to every figure. Australian coverage in the eight is thin: four include Australian respondents and none reports on Australian boards. The reports describe conditions in late 2025 and early 2026 and will date. Section 06 is a position, grounded in the author's doctoral work and earlier published writing, and is offered as such.

For three years boards were asked what AI could tell them. The eight reports read here say the question has changed, and none of them says what the change means for the people who answer for the organisation.

What changes for a board is not the technology but where decisions happen. The questions that follow are the ones a board has always asked of anything it funds: how much of this risk will we carry, who answers for each decision, and what record exists when one goes wrong. The reports stop before they reach them.

PROCESS IS NOT PROOF. THE RECORD IS WHAT A BOARD WILL BE ASKED FOR.

References

The eight reports were opened and read in full on 17 October 2026 (AEDT). Each entry notes what a reader should know before repeating its figures.

1. Accenture Strategy and Accenture Research (2025). *The new rules of platform strategy in the age of agentic AI*. 18 December 2025. 1,031 C-suite and senior executives, 12 countries, 10 industries, fieldwork 2025; 20 interviews; ANCOVA on 2024 outcomes for 602 public companies. accenture.com/en/insights/strategy/new-rules-platform-strategy-agentic-ai. The 37% figure is labelled an EBITDA lift on the landing page and profitability in the executive summary; the growth and profit advantage is correlational.
2. Boston Consulting Group (2026). *AI Radar 2026: As AI investments surge, CEOs take the lead*. 15 January 2026. 2,360 executives including 640 CEOs, 16 markets, nine industries; fielding dates not stated. bcg.com/publications/2026/as-ai-investments-surge-ceos-take-the-lead. A web article and 29-slide deck; every return figure is an expectation.
3. Deloitte AI Institute (2026). *State of AI in the Enterprise: The untapped edge*. 21 January 2026. 3,235 senior leaders, 24 countries including Australia, six industries, fieldwork August to September 2025; 15 executive interviews. [deloitte.com\(PDF\)](https://deloitte.com(PDF)). Branded 2026 on 2025 fieldwork; chart bases vary between 3,199 and 3,325; the 25% production figure is self-estimated.
4. Anthropic (2026). *The 2026 State of AI Agents Report*. Publication date not stated; first third-party coverage 5 March 2026. More than 500 US technical leaders, fielded by Material in late 2025, plus Anthropic Economic Index usage data and 12 customer case studies. [resources.anthropic.com\(PDF\)](https://resources.anthropic.com(PDF)). Respondents are adopters of one vendor's products; returns are self-reported and undefined; the report contains no material on safeguards, permissions or monitoring.
5. McKinsey & Company (2026). *Building the foundations for agentic AI at scale*. 2 April 2026. Ten-page article adapted from the data-architecture chapter of *Rewired*, second edition (Wiley, 14 April 2026). mckinsey.com/capabilities/tech-and-ai/our-insights/building-the-foundations-for-agentic-ai-at-scale. The headline adoption figures are footnoted to McKinsey's June 2025 agentic AI report, which does not contain them; they match McKinsey's State of AI 2025 survey (1,993 respondents, June to July 2025).
6. Stanford Institute for Human-Centered AI (2026). *The 2026 AI Index Report*. 13 April 2026. 423 pages, nine chapters aggregating third-party data. hai.stanford.edu/ai-index/2026-ai-index-report. Figures in this paper come from the R&D, Technical Performance, Responsible AI, Economy, Policy and Governance, and Public Opinion chapters. Agent failure rates are benchmark results, not field data.
7. Microsoft WorkLab (2026). *2026 Work Trend Index Annual Report: Agents, human agency, and the opportunity for every organization*. 5 May 2026, PDF revised September 2026. 20,000 knowledge workers who already use AI, about 2,000 in each of 10 countries including Australia, fielded 18 February to 7 April 2026; Microsoft 365 telemetry and about 105,000 Copilot conversations. microsoft.com/en-us/worklab/work-trend-index. The sample selects on existing AI use; the 67% and 32% figures are associations, not causes; the PDF was revised without a changelog.
8. IBM Institute for Business Value (2026). *The blueprint for agentic operations: How to build an interconnected enterprise*. 12 May 2026. Two surveys of 1,000 each, C-suite and senior operations executives, 17 industries, 16 countries (Australia 5%); fielding dates not disclosed. ibm.com/thought-leadership/institute-business-value/en-us/report/agentic-ai-enterprise-operations. Every value figure is an expectation to 2027. An earlier IBV and Oracle report, *Agentic AI's strategic ascent* (October 2025), sits at a web address ending in *agentic-ai-operating-model*, the likely origin of the mis-titling in circulation.
9. Australian National Audit Office (2026). *Artificial Intelligence and Medicare Benefits Integrity*. Auditor-General Report No. 5 2026–27, Performance Audit, Department of Health, Disability and Ageing. Presented 21 September 2026. Cited in Section 04 and Appendix A, question 20; the register count and the fraud-detection system's status are at Table 4.1 of the report.
10. National Institute of Standards and Technology (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1, January 2023. nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf. Cited in Section 04 for risk tolerance and accountability; not one of the eight reports read.

The question paper

The 38 questions the review arms a board to ask, drawn from the eight reports and, in one case, from the Auditor-General's report, grouped the way an audit and risk committee chair thinks.

The red number is the source each rests on; the grey number is the situation in Section 05 it serves: 01 running pilots, 02 agents in production, 03 writing a strategy. There is no appetite group, because none of the eight arms one. That question is the board's own, and Section 05 sets it before spend.

RELIANCE: WHAT THE NUMBER RESTS ON

- 1 Is this investment tied to the business and platform strategy, or a separate paper with its own sponsor? [1] 03
- 2 What is our AI spend as a share of revenue, this year and next, and what does the board expect for it? [2] 03
- 3 Is continued funding conditional on anything the board can observe? [2] 01 03
- 4 Which executives beside the CEO have signed the numbers? [2] 03
- 5 This paper rests on a pilot. Which pilot conditions does the production case assume will hold? [3] 01
- 6 The productivity or revenue claim: measured here, or borrowed? [3] 01 03
- 7 Whose customers was the returns figure measured on? [4] 03
- 8 What has been realised to date, as against expected by 2027? [8] 02

AUTHORITY: WHO DECIDES, WHO ANSWERS

- 9 When agents act without an approval chain, which of this board's decision rights change, and is it written down? [2] 02
- 10 Which decisions will these agents take without a person, and does the governance for that exist now, inside the existing risk structure? [3] 01 02
- 11 Who owns day-to-day governance, who owns the guardrails, who owns the agents' identities: three names or one? [5] 02
- 12 Who holds AI governance as their job, and is it in their KPIs? [6] 02 03
- 13 Which decisions are routine and go to the agent alone, which go to a team, which to a named person? [8] 02
- 14 Who is accountable, by name, for the top tier? [8] 02
- 15 What share of tasks is fully delegated rather than assisted, and who approved the move? [4] 02

EVIDENCE: THAT THE CONTROLS OPERATE

- 16 Are coding agents writing production code today, and who reviews it before it runs? [4] 02
- 17 Is every action an agent takes logged where the board could inspect it, and who reads the log? [5] 02
- 18 What is our own AI incident record for the year, and who sees it? [6] 02
- 19 What error rate does this paper assume for the agent, and what catches the error? [6] 01

- 20 Is this system on the organisation's AI register, and is the register complete? The Auditor-General's 2026 audit of a federal department found 55 AI uses on the register and a fraud-detection system that had run an AI model on personal data for eighteen months not on it. [9] 02
- 21 Who checks AI output before it leaves the organisation: a job, or an assumption? [7] 02
- 22 Are agents managed as identities with permissions and a lifecycle, like staff accounts? [7] 02
- 23 Is this agent in our cyber risk assessment as a threat as well as a tool? [2] 02

DATA AND SYSTEMS

- 24 Where in our application estate will the agents act, and were those systems in scope for the risk assessment? [1] 01
- 25 What data will the agents act on, who owns its quality, and has it been verified for agents rather than for reports? [5] 01
- 26 Which end-to-end workflows is this limited to, and who decided the boundary? [5] 01
- 27 Of IBM's six capabilities, which are in place and which are assumed? [8] 01 03
- 28 What share of our data can the agents reach, and what are they acting on in the meantime? [8] 01
- 29 Is the integration and data-access cost in the budget, or below the line? [4] 03

PEOPLE AND LEADERSHIP

- 30 What has been redesigned so the tool is used in the work rather than beside it? [3] 01 02
- 31 What has been spent on training against what has been spent on licences? [1] 03
- 32 Can the executive team state the AI priority in the same words? [7] 03
- 33 Do the managers use the tools this paper asks the board to fund, and is that measured? [7] 02
- 34 How many agents are running inside our licensed tools today, and who created them? [7] 02

VENDORS AND JURISDICTION

- 35 Which platform vendor does this depend on, for how long, and what happens if it does not survive the contract? [1] 03
- 36 Where will this system and its data sit, and did country of origin enter the vendor decision? [3] 03
- 37 What do we know about the models we depend on beyond what the vendor discloses? [6] 03
- 38 Which jurisdictions' AI laws reach this system, including where our customers and vendors are? [6] 03