



BOARD OVERSIGHT OF AI

Governing AI Through Value, Risk
and Accountability

MANOJ TAVARAJOO



EXECUTIVE SUMMARY



Board oversight of AI is increasing, but effectiveness is not keeping pace.

Boards are now expected to oversee AI as a strategic capability, yet many remain dependent on management updates rather than evidence. Oversight is present in form but often weak in practice. Boards are informed. They are not always equipped to challenge.

This creates a structural risk. AI is not a conventional technology investment. It introduces forms of uncertainty, evolving risk and uneven value realisation that traditional technology governance was not designed to handle. AI systems operate through probabilistic logic, depend on complex data and third-party models, and continue to change after deployment.



A central tension underpins this challenge. This is the boardroom AI literacy paradox. Investors and regulators are demanding visible AI oversight, and organisations are disclosing it. Yet the fluency required to make that oversight credible has not kept pace. Oversight is being declared faster than it is being built.



This paper defines a practical model for board-level AI oversight. It clarifies what boards should oversee, what management must own, and how to govern AI without becoming operational. It introduces five oversight duties: value, risk appetite, accountability, assurance and maturity. It positions the behavioural envelope as the unit of oversight, enabling boards to move from abstract principles to measurable boundaries.

Effective oversight requires a shift. Boards must move from being informed to being able to challenge, and from narrative to evidence. When boards do this well, AI governance becomes a system of value creation and control. When they do not, governance becomes documentation.

WHY AI OVERSIGHT IS DIFFERENT

AI cannot be governed in the same way as traditional technology. The differences are structural and material.

Traditional technology oversight focuses on delivery, expenditure, project status and cyber risk. These disciplines assume systems behave predictably. AI systems do not.

Behavioural risk is inherent. AI systems produce outputs that are probabilistic and context-dependent. This variability must be bounded and governed.

Risk does not end at deployment. AI systems evolve as data changes and models drift. The National Institute of Standards and Technology frames AI risk management as a continuous lifecycle of governing, mapping, measuring and managing. Oversight that concludes at implementation is incomplete.

Value is uneven. BCG's 2025 research found that only 5% of companies are capturing substantial AI value, while 60% report minimal gains despite significant investment. MIT NANDA found that 95% of enterprise generative AI initiatives are not delivering measurable revenue growth. Boards that receive activity-based updates rather than outcome-based reporting cannot distinguish investment from impact.



Risk exposure is increasing. The Stanford Institute for Human-Centered Artificial Intelligence AI Index reported 233 AI incidents in 2024, a 56.4% year-on-year increase, reflecting both increased adoption and increased exposure.

Regulatory expectations are tightening. The Financial Reporting Council Corporate Governance Code requires boards to monitor controls over emerging technologies, reinforcing that AI oversight is now a board-level governance responsibility.

This is why AI cannot be governed as a conventional technology investment.

THE GOVERNANCE PROBLEM

Despite increasing board attention to AI, oversight in practice remains largely passive.

Boards receive updates on AI initiatives, but these updates often focus on activity rather than outcomes. Use cases are presented, pilots are discussed, and progress is reported. What is often missing is a clear view of value delivered, risks taken, and whether systems are operating within acceptable limits. This creates a gap between visibility and control.



This tension reflects the boardroom AI literacy paradox. The NACD's 2025 survey found that only 17% of boards have an AI education plan, and 66% of directors report limited or no AI knowledge. PwC found that only 35% of directors say their boards have incorporated AI into oversight roles. The Diligent Institute found that while 66% of boards use AI, only 3% have integrated it into risk oversight and strategy.

Oversight is therefore present in form, but not credible in substance.

The transition required is illustrated in Figure 1.

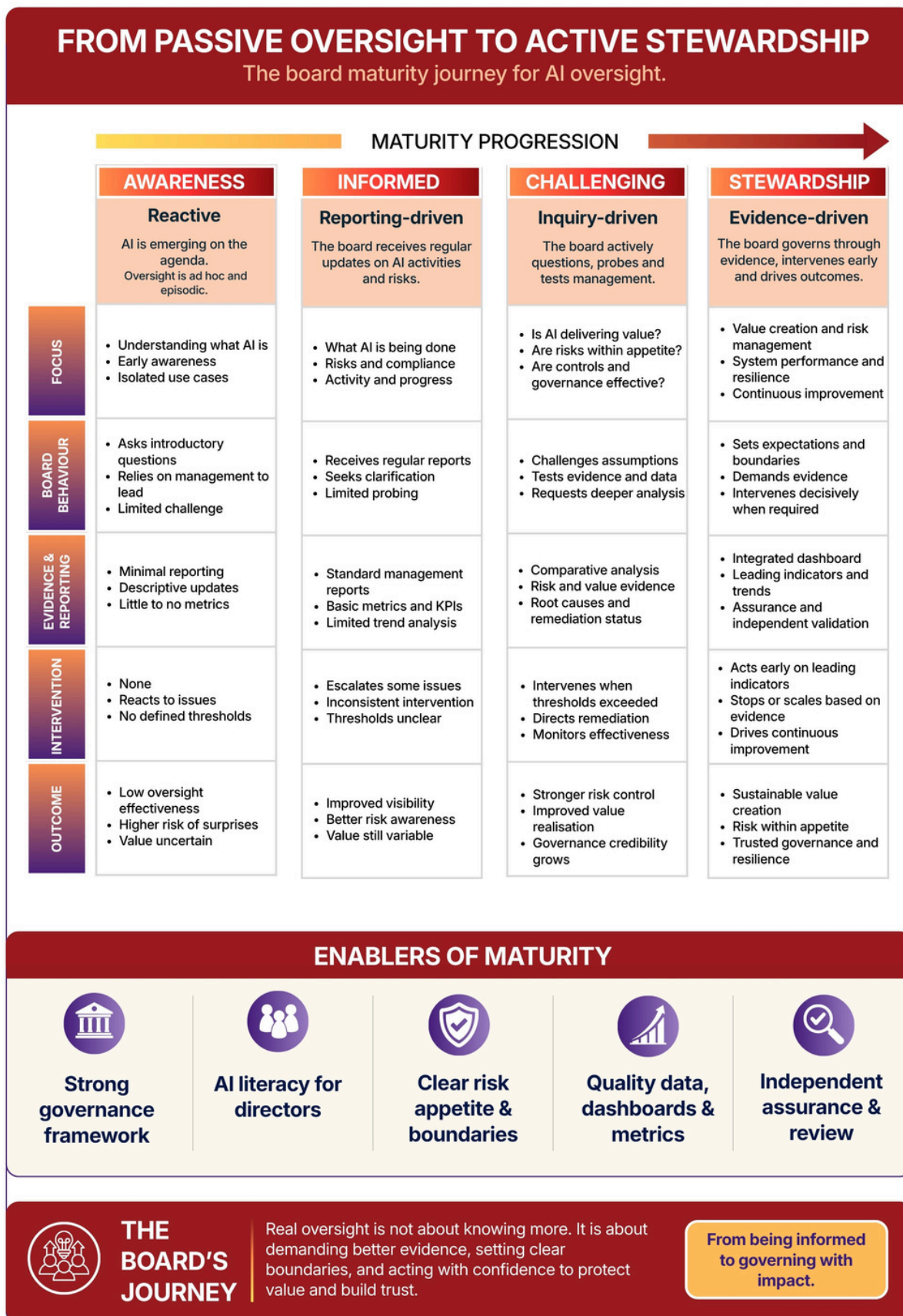







Figure 1: From Passive Oversight to Active Stewardship



The model maps four stages of board maturity. Most boards operate between being informed and being able to challenge. Effective oversight requires progression to evidence-driven stewardship, where the board governs through evidence, intervenes when thresholds are breached, and drives sustainable value creation.

The result of remaining at earlier stages is predictable:

-  The board receives regular updates, but no one asks what value has actually been realised
-  Risk appetite is discussed, but not translated into operational thresholds
-  Accountability is assigned, but not tested in practice
-  Assurance is reported, but not independently validated
-  Issues are identified, but intervention thresholds are unclear

Governance exists in structure. It does not consistently function in practice.

WHAT BOARDS SHOULD OVERSEE

Boards should not manage AI. They should oversee whether AI is being managed effectively. That distinction is consequential and must be maintained deliberately.

Effective board oversight of AI is defined through five core duties. Together, they constitute an oversight model that is active, evidence-based and appropriately bounded, as shown in Figure 2.

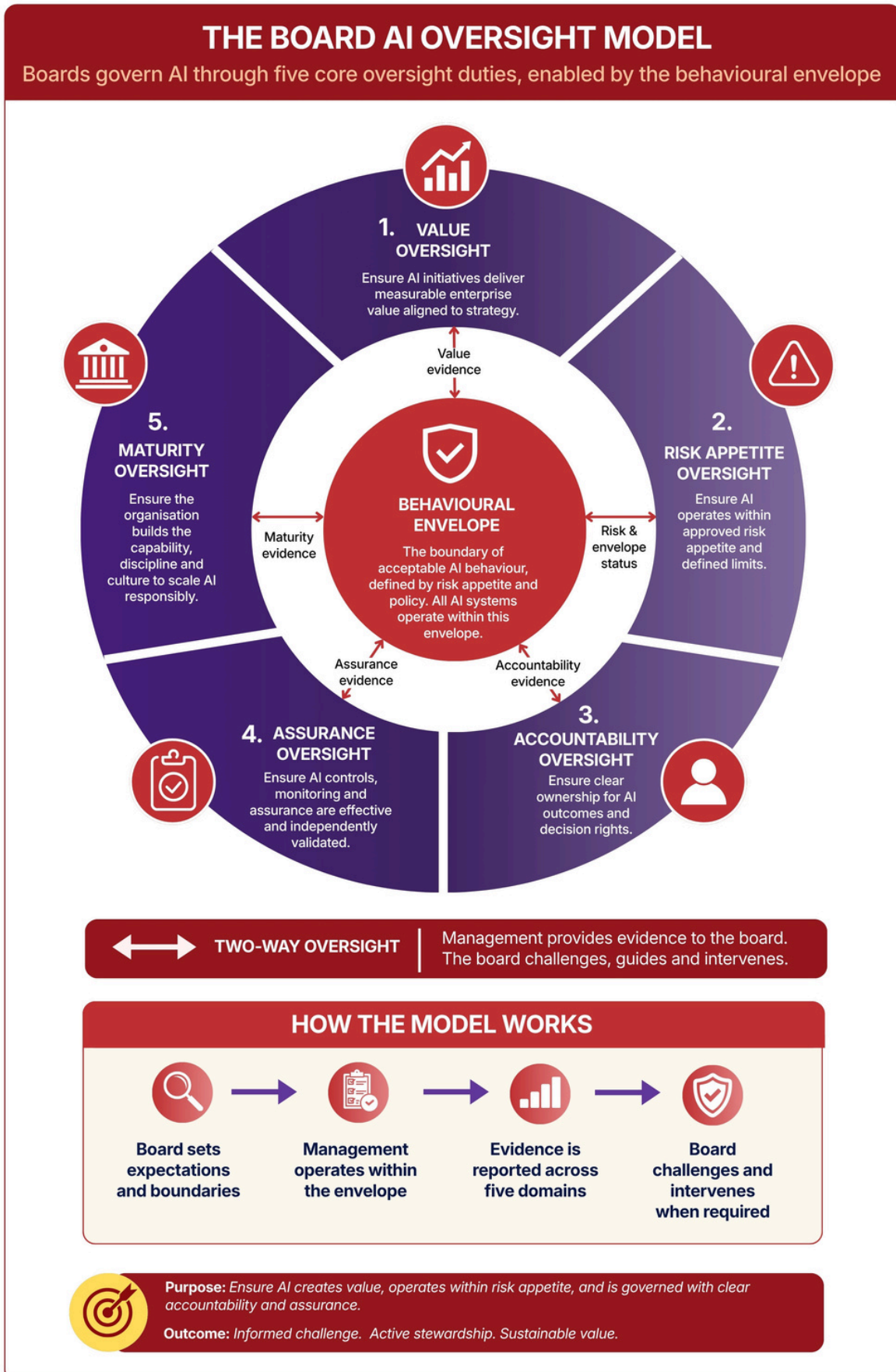


Figure 2: The Board AI Oversight Model



Value Oversight

Boards must ensure that AI is generating measurable enterprise value, not just activity. Only 5% of organisations are capturing substantial AI value despite widespread investment. Boards should expect clarity on which use cases are delivering returns against plan, which are not, and what is being done about underperforming investments.

Risk Appetite Oversight

Boards must define the boundaries within which AI operates. Without defined thresholds, executives cannot make consistent decisions and risk functions lack escalation benchmarks. A board-approved risk appetite statement should define excluded uses, tolerance levels for automated decision-making in high-impact areas, escalation thresholds, expectations for human oversight, and the organisation's position on third-party model dependency.

Accountability Oversight

Boards must ensure that ownership of AI outcomes is clear and functioning. In most organisations, AI accountability is distributed across functions without a single visible point of executive accountability. Boards should be able to confirm that a named executive is accountable for AI governance, that business owners are assigned to each material AI system, and that accountability structures are tested, not merely documented.

Assurance Oversight

Boards must require evidence that controls are working, not just documentation that they exist. Independent assurance from internal audit, and where appropriate external review, should be expected for high-risk AI systems. The question for the board is not whether controls are in place. It is whether they are working.



Maturity Oversight

Boards must ensure that the organisation is building the capability to scale AI responsibly. As AI scale increases, governance and capability requirements increase with it. Boards should have visibility of whether maturity is developing at a pace commensurate with AI ambitions.

The Behavioural Envelope

The five duties are connected through the behavioural envelope: the board-approved set of limits within which AI systems are expected to operate. The envelope defines acceptable dimensions of AI behaviour across accuracy, bias, latency, cost, drift over time and degree of permitted autonomy. Each dimension has defined limits, measurement methods and escalation triggers.

Boards do not oversee individual models. They oversee whether AI systems operate within these boundaries. The behavioural envelope converts governance from abstraction to measurement, and provides management with clear operational limits. When systems operate within the envelope, governance is functioning. When they breach it, the board has a defined basis for intervention.

WHAT MANAGEMENT MUST OWN

The board sets direction, expectations and boundaries. Management operates within them. This distinction must be explicit and consistently maintained.

Management is responsible for executing the AI strategy, identifying use cases, prioritising investments and aligning AI initiatives to business objectives. It designs, builds and deploys AI systems, and is accountable for the quality of models, data and technical infrastructure.



Management owns risk management within the board-approved appetite. It monitors system performance, detects drift, manages incidents and ensures escalation mechanisms function. It governs vendor and third-party relationships, maintains data quality and security standards, and is responsible for value realisation, not merely deployment.



Management must also provide the board with clear, timely and decision-ready evidence across all five oversight domains. Producing decision-ready information for the board is not administrative overhead. It is a governance responsibility.

A board that receives activity-based updates rather than outcome-based evidence cannot fulfil its oversight duty. The quality of board oversight depends directly on the quality of what management provides.

THE OVERSIGHT BOUNDARY

The boundary between board oversight and management execution is the most important structural element of AI governance. When it is clear, oversight functions. When it is unclear, boards either under-engage or over-reach, and both failure modes are damaging.

Figure 3 defines this boundary explicitly.

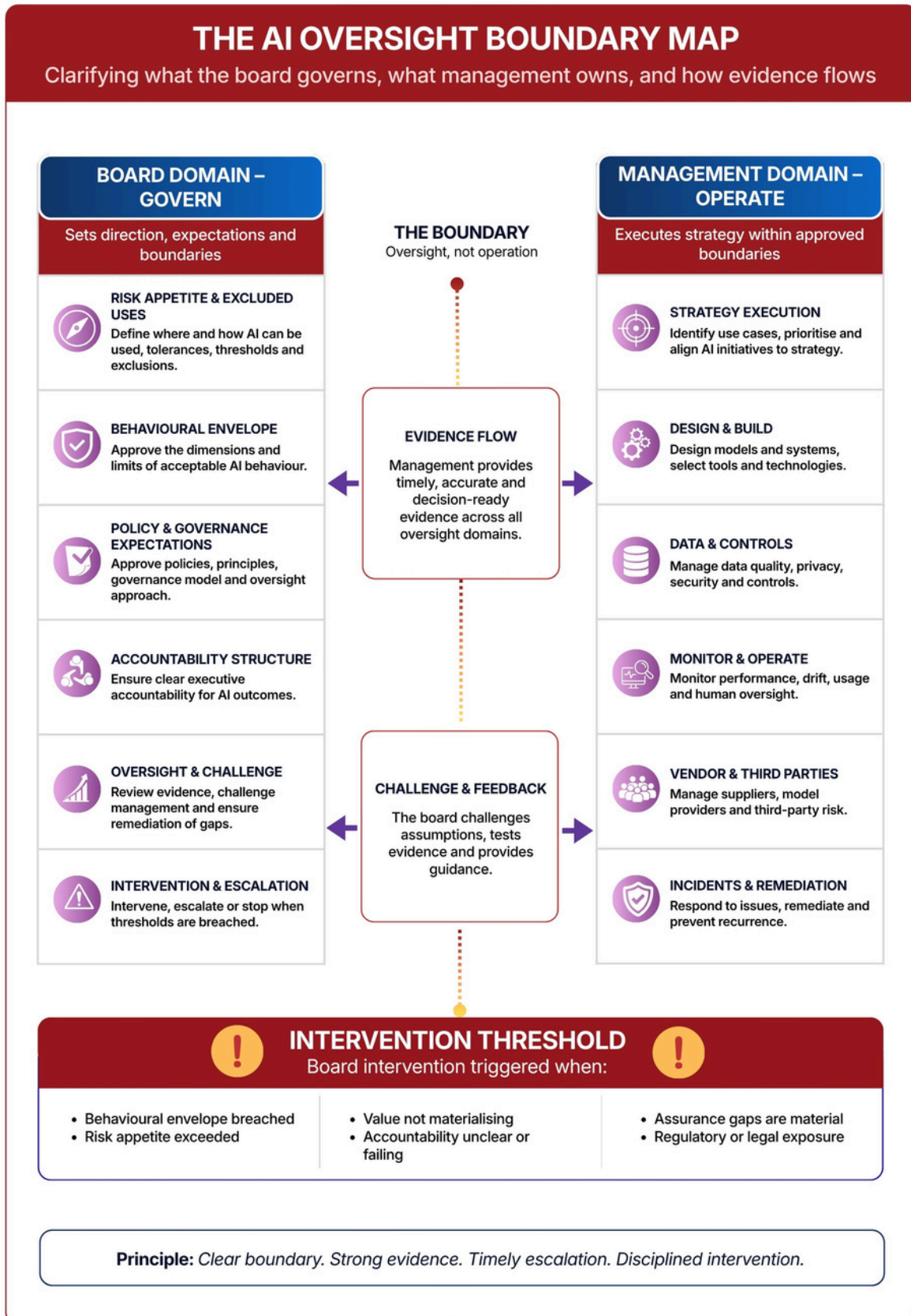


Figure 3: The AI Oversight Boundary Map



The board domain encompasses six responsibilities: defining risk appetite and excluded uses; approving the behavioural envelope; setting policy and governance expectations; confirming accountability structures; conducting oversight and challenge; and intervening when thresholds are breached.

The management domain encompasses six corresponding responsibilities: strategy execution; system design and build; data quality and controls; monitoring and operation; vendor and third-party management; and incident response and remediation.

The connection between the two domains is evidence. Management provides timely, accurate and decision-ready information. The board challenges assumptions, tests evidence and intervenes when required.

The intervention threshold is defined in advance: behavioural envelope breached, risk appetite exceeded, value not materialising, accountability unclear or failing, assurance gaps material, or regulatory or legal exposure identified. Intervention is a defined response to evidence, not a discretionary reaction to instinct.

Boards should remain alert to two failure modes on either side of the boundary. The first is under-engagement: accepting management reports without challenge, approving policies without demanding evidence, treating oversight as a compliance obligation. The second is operational interference: directing specific technical decisions or substituting board judgement for management judgement on operational matters. The boundary holds when boards are active on the duties that are theirs and restrained on the matters that belong to management.

THE AI BOARD DASHBOARD

Effective oversight requires a structured and consistent view of AI across the enterprise. Without it, board conversations default to what management chooses to present, rather than what the board needs to examine.

The dashboard in Figure 4 provides a high-level, evidence-based view across five domains, each aligned to one of the five oversight duties.

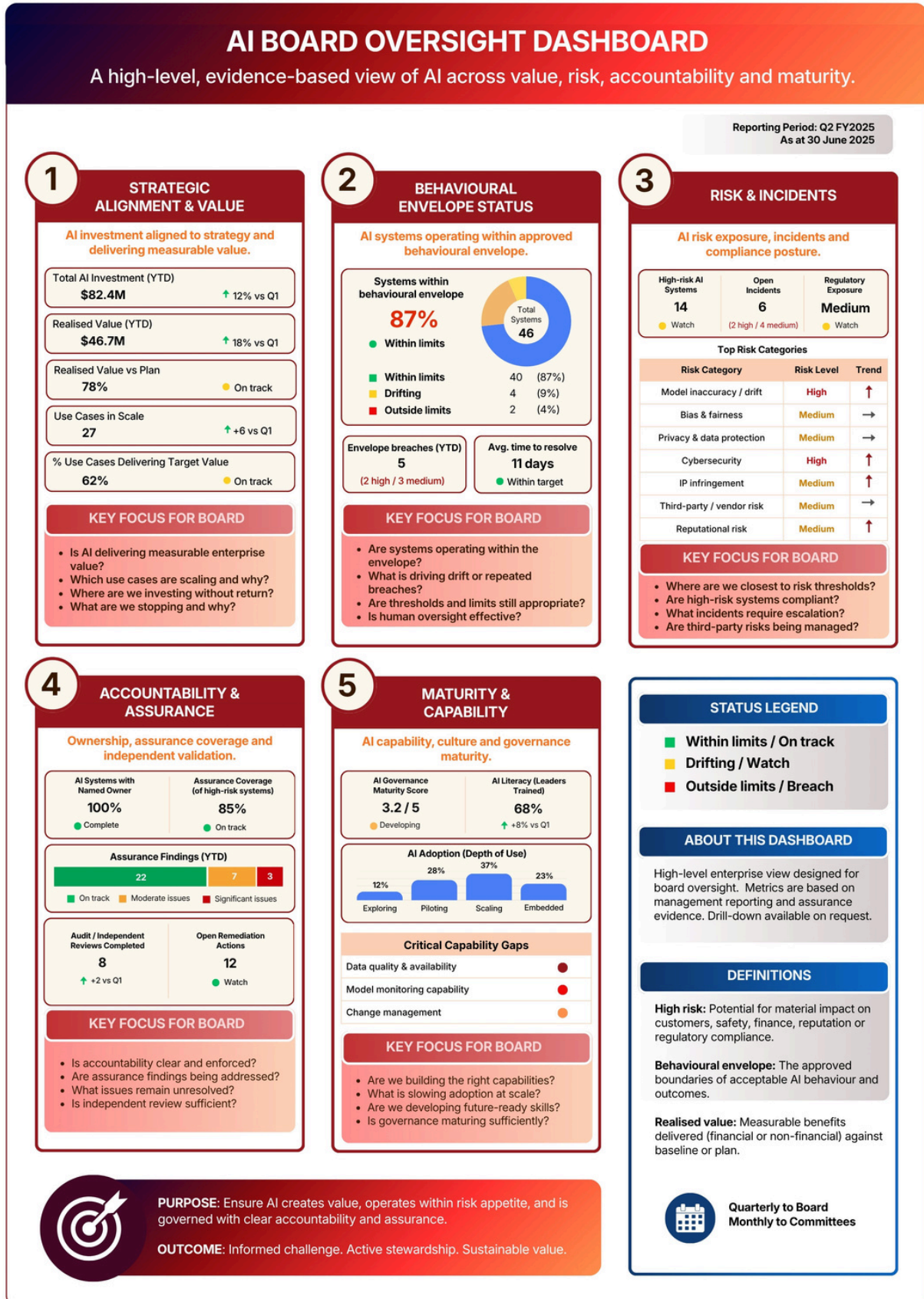


Figure 4: AI Board Oversight Dashboard



Strategic alignment and value gives the board visibility of total AI investment, realised value against plan, the number of use cases in scale, and the percentage delivering target value. The key questions are whether AI is creating measurable enterprise value, which use cases are scaling and why, and where investment is continuing without a line of sight to return.

Behavioural envelope status shows what proportion of active AI systems are operating within approved limits, how many are drifting, and how many have breached the envelope. The board's focus is whether systems are within limits, what is driving drift or repeated breaches, and whether human oversight mechanisms are functioning.

Risk and incidents provides a view of the AI risk profile: high-risk systems, open incidents, regulatory exposure and the top risk categories by level and trend. Boards should understand where the organisation is closest to risk thresholds and which systems require escalation.

Accountability and assurance shows the proportion of AI systems with a named accountable owner, assurance coverage of high-risk systems, audit and independent review results, and open remediation actions.

Maturity and capability provides visibility of AI governance maturity scores, leadership training coverage, depth of AI adoption across the organisation, and identified gaps in critical capabilities.



The dashboard is not designed to inform. It is designed to enable challenge.

The dashboard should be available quarterly to the full board and monthly to relevant committees. It should be produced by management and independently validated, so the board is not reliant on self-reported data. The shift from receiving management reports to reviewing a structured evidence dashboard is a meaningful governance upgrade.



FAILURE MODES IN BOARD OVERSIGHT

Six recurring failure modes define the boundary between oversight that functions and oversight that exists only in structure.

Passive Acceptance

The board receives AI updates but does not challenge them. Management presents progress; the board notes it. Risk appetite is accepted without interrogation. Assurance is accepted without independent validation. The appearance of oversight exists. The substance does not.

Policy Without Enforcement

The board approves an AI governance policy and considers its duty discharged. Policies define intent. They do not govern behaviour. A policy not embedded in operating practice, decision rights and accountability structures does not control outcomes.

Delegation Without Visibility

AI oversight is assigned to a subcommittee without full board awareness. Committees have an important role, but when the full board has no sight of AI governance, accountability at the top of the organisation becomes unclear.





Undefined Risk Appetite

The board has not approved a specific AI risk appetite. Without defined tolerances, executives cannot make consistent decisions and risk functions lack escalation thresholds. Risk appetite is not a technical document. It is the governing instrument through which the board makes its expectations operational.

Weak Assurance

Oversight is based on management confidence rather than independent evidence. Boards that accept management reporting without independent validation cannot confirm that controls are working.

Oversight Without Intervention




The board identifies issues but action does not follow. Effective oversight requires not only the ability to identify problems but the will and the mechanism to act on them. Observation without intervention is not governance.






BOARD QUESTIONS FOR MANAGEMENT

The quality of board oversight is significantly shaped by the quality of questions boards ask. The following questions are designed to move the board from receiving information to challenging it.




Value

-  Which AI initiatives are delivering measurable outcomes against their stated business case, and which are not?
-  What are we stopping, and on what basis have those decisions been made?
-  Where are we investing without a clear line of sight to value, and why are those investments continuing?

Risk and the Behavioural Envelope




-  What is currently operating outside our risk appetite or outside the behavioural envelope, and what is being done about it?
-  Where are we closest to our defined limits, and what is management's response plan?
-  Are the risk categories we have defined still appropriate, or has the risk profile of our AI systems changed?

Accountability




-  Who is personally accountable for the outcomes of each material AI system currently in operation?
-  How is that accountability enforced, and what happens when it is unclear or disputed?
-  Has accountability been tested in practice, including through incidents or near-misses?



Assurance

-  What independent evidence confirms that controls are working, rather than management assertion that they are?
-  What have internal audit or external reviewers identified, and how are findings being addressed?
-  Are there systems in production where assurance coverage is insufficient?

Maturity

-  Are we building the governance infrastructure, data foundations and skills required to scale AI at the pace we intend?
-  What are the critical capability gaps, and how are they being addressed?
-  Is governance maturing in step with our AI ambitions, or are we building ahead of our governance capacity?

These questions do not require technical expertise. They require governance discipline: the commitment to demand evidence, to challenge assumptions and to act when the evidence requires it.



WHAT LEADERS SHOULD DO NOW

Define AI Risk Appetite at Board Level

A board-approved AI risk appetite statement is the foundational governing instrument. Without it, executives lack clear parameters and risk functions lack escalation thresholds. The statement should define excluded uses, decision-making tolerances in high-impact areas, escalation criteria and expectations for human oversight. It should be reviewed at least annually.

Establish and Approve the Behavioural Envelope

For each material AI system or category of systems, the board should approve the boundaries of acceptable behaviour across accuracy, bias, latency, cost, drift over time and degree of permitted autonomy. This converts oversight from abstraction to measurement and provides management with clear operational limits.

Implement a Board-Level AI Dashboard

The dashboard should be structured around the five oversight duties, evidence-based rather than activity-based, delivered quarterly to the full board and monthly to relevant committees, and independently validated. Its purpose is to enable challenge, not to report progress.

Clarify and Confirm Accountability

Boards should confirm that a named executive is accountable for AI governance, that business owners are assigned to each material AI system, and that the accountability chain is tested in practice, not merely documented.



Strengthen Assurance

Internal audit should have explicit coverage of AI systems and governance. High-risk AI systems should be subject to independent review on a defined cycle. The board should see assurance findings directly, not only management summaries of them.

Develop Board-Level AI Literacy

Boards do not need technical expertise. They need sufficient literacy to ask informed questions, interpret evidence and challenge management effectively. The goal is not for directors to become AI specialists. It is for the board to become a credible governing body for AI.

CLOSING INSIGHT

Board oversight of AI is not a technical exercise. It is a governance discipline.

The organisations that will define the next decade will not be those that adopt AI fastest. They will be those that govern it most effectively: creating value at scale, managing risk within defined limits, and holding clear accountability for outcomes.

What separates effective board oversight from passive oversight is not knowledge. It is behaviour. It is the board's willingness to demand evidence, to challenge assumptions, to define the boundaries within which AI operates, and to intervene when those boundaries are breached.



Active stewardship of AI is not an intrusion into management. It is the mechanism through which governance gains its authority. When boards govern with discipline, management builds with clarity. When boards are passive, governance becomes documentation.

The question for every board is not whether AI is on the agenda. It is whether the board is governing it with the evidence, the structure and the intent that the moment requires.



REFERENCES

- ↘ Boston Consulting Group. The Widening AI Value Gap. BCG Global, 2025.
- ↘ Deloitte. AI Governance and Board Oversight. Deloitte Insights, 2025.
- ↘ Diligent Institute. What Directors Think. 2026.
- ↘ European Union. Artificial Intelligence Act. Official Journal of the European Union, 2024.
- ↘ Financial Reporting Council. UK Corporate Governance Code. FRC, 2024.
- ↘ Harvard Law School Forum on Corporate Governance. Cyber and AI Oversight Disclosures. 2025.
- ↘ Harvard Law School Forum on Corporate Governance. AI Oversight Through Three Lenses. 2026.
- ↘ Australian Institute of Company Directors and Human Technology Institute. A Director's Guide to AI Governance. 2024.
- ↘ ISO/IEC. ISO/IEC 42001 Artificial Intelligence Management System. 2023.
- ↘ KPMG. Boardroom Lens on Generative AI. KPMG International, 2025.
- ↘ MIT NANDA. State of AI in Business. 2025.
- ↘ National Association of Corporate Directors. Public Company Board Practices and Oversight Survey. NACD, 2025.
- ↘ National Association of Corporate Directors. Governance Outlook. NACD, 2025.
- ↘ National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). 2023.
- ↘ National Institute of Standards and Technology. Generative Artificial Intelligence Profile (AI 600-1). 2024.
- ↘ Organisation for Economic Co-operation and Development. OECD AI Principles. Updated 2024.
- ↘ PwC. Annual Corporate Directors Survey. PwC, 2025.
- ↘ PwC. Responsible AI Survey. PwC, 2025.
- ↘ Stanford Institute for Human-Centered Artificial Intelligence. AI Index Report. Stanford University, 2025.
- ↘ World Economic Forum. Responsible AI Playbook. WEF, 2025.



About the Author

Manoj Tavarajoo has spent over two decades working with boards, executives and senior leaders across enterprise, digital and AI transformation. His work sits at the intersection of strategy, operating model design, governance and execution, the point where good intentions either become operating reality or quietly fail.

He is the author of *Leading the AI Transformation* and *The AI Operating Model Playbook*. This paper series extends that work into the governance layer: how AI is directed, controlled, assured and held accountable at enterprise scale.

He works through MyConsultancy, an independent advisory practice based in Australia.

 [@manojtavarajoo](https://www.linkedin.com/company/manojtavarajoo)

About MyConsultancy

MyConsultancy works with boards and executives navigating the distance between AI ambition and operating reality. The firm focuses on strategy, governance and operating model design, helping organisations build the portfolio discipline and transformation assurance needed to scale AI responsibly across complex enterprise environments.



www.myconsultancy.com.au

MyConsultancy