ORBISCO GROUP

Closing the AI Value Gap

An Integrated Framework for Traditional Cost-Benefit Review, Governance Discipline and AI Measurement

White Paper | May 2026

A board-ready framework for measuring AI investment through financial return, risk control, customer impact and operational performance.

Orbisco Group | Level 17, Angel Place, 123 Pitt Street, Sydney NSW 2000 | www.orbisco.com.au

CONTENTS AT A GLANCE

A practical reader's guide to the paper and its implementation tools.

Core thesis: AI is not the engine. It is the reason to open it. Once the organisation is open, the real work is to inspect the data, systems, processes, controls and reporting that determine whether value can actually flow. The goal is not to bolt AI onto the old model. The goal is to rebuild what limits performance.

Measure the money	Govern the risk	Protect the customer	Control the model
Financial value and cost discipline	Controls, accountability and assurance	Service quality and stakeholder experience	Performance, reliability and AI operating cost

Executive Summary

Organisations are rapidly transitioning from AI experimentation to deployment; however, the discipline required to demonstrate and realise value has not advanced at the same rate. The central issue goes beyond the functionality of artificial intelligence. The more challenging question is whether organisations can demonstrate, with the rigour expected of major business investments, that AI initiatives strengthen financial performance, reduce risk, advance customer or service quality, and strengthen organisational capability.

This paper contends that AI should not be approached as an isolated technology initiative, using such terms as 'Transformation'. Instead, AI should be integrated within a comprehensive organisational value review, encompassing structured assessments of cost, process, risk, service quality, governance, workforce capability, data readiness, and technology opportunities. Traditional management disciplines such as cost-benefit analysis, benefits realisation, management accounting, KPI design, risk management, and customer service measurement stay essential. Rather than replacing these disciplines, AI increases their importance.

Orbisco identifies measurement failure as the main challenge and proposes a four-pillar framework merging financial, governance, customer, and AI-specific metrics. This paper uses that framework while improving the Orbisco perspective: leading organisations do not begin by asking where AI can be applied. Instead, they first find areas of value leakage, duplicated work, rising risk, and opportunities where technology, including AI, can deliver measurable improvements.

The evidence supports this position. McKinsey reports that high-performing AI organisations are distinguished by management practices across strategy, talent, operating model, technology, data, adoption and scaling. Gartner reports that AI adoption has moved from two out of five organisations in 2024 to four out of five in early 2026, yet only 44 percent of surveyed organisations had adopted financial guardrails or AI FinOps practices. Deloitte similarly frames AI return as more than technology deployment, measuring AI performance through direct financial return, revenue growth, operational cost savings and speed of results.

The practical outcome is the Orbisco Integrated AI Value Review, which combines traditional organisational review methods with contemporary digital and AI assessment. This approach generates a quantified benefits case, a risk-adjusted AI investment model, a governance-ready scorecard, and a sequenced implementation roadmap. Its main goal is to support boards and executive teams in transitioning from AI activity to AI accountability.

This paper recommends four executive disciplines: measuring financial outcomes, governing risk, protecting the customer, and controlling the model. Collectively, these disciplines offer a clear board-level framework to handle the complexity of pilots, tools, dashboards, policies, and vendor commitments.

The central question is not where AI can be applied, but rather where organisational value is leaking and which combination of process redesign, governance, technology, and AI will deal with these gaps.

1. The Core Argument: AI Is a Value Realisation Problem

The mistake many organisations make is to treat AI as a tool-selection problem. They compare model providers, run pilots, encourage staff to experiment and then look for productivity gains after the fact. This creates energy, but not necessarily value. It can also create risk: ungoverned data use, inconsistent customer experiences, duplicated tools, hidden cloud or inference costs, and a long list of pilots that never scale.

A more disciplined starting point is to treat AI as one component of organisational value realisation. In that model, AI is not the strategy. AI is one possible intervention inside a wider strategy for improving performance. The value case begins with the organisation, not the tool.

This matters because the major sources of value are rarely confined to the AI system itself. They often sit in the surrounding operating model: the process that should have been simplified before automation; the data that should have been cleaned before modelling; the staff capability that determines whether an AI assistant is used well or ignored; the risk appetite that determines where automation is acceptable; and the management reporting that determines whether benefits are visible or anecdotal.

From AI activity to AI accountability

AI activity is easy to create. AI accountability is harder. Activity can be measured by the number of pilots launched, licences purchased, prompts tested, chatbots deployed or workflows automated. Accountability requires a different test: whether the organisation can define the baseline, quantify the expected benefit, assign ownership, monitor leading indicators, measure realised outcomes and explain the risk position to executives and the board.

That is why this paper uses the term value gap. The gap is not simply between AI potential and AI reality. It is the gap between AI activity and auditable evidence of organisational benefit.

The Orbisco point of view

Orbisco asserts AI as a serious executive discipline. The strongest message is that AI needs to be brought into the same governance and value-realisation architecture that already governs finance, operations, customer service, risk and major transformation. It is inappropriate to have the firms selling tools, prompts, workshops or software as the drivers of AI implementation; it is imperative that an internally controlled, top-down, executive-level, capacity-enabled architecture is in place.

The Orbisco point of view can be expressed simply: AI is only valuable when it changes a measurable organisational outcome. That outcome may be financial, operational, customer-related, workforce-related, risk-related or strategic. But it must be named, measured and owned.

The four executive questions

Figure 1. The four executive questions in an AI value review

Question	What it asks	Typical evidence
1. Measure the money	What financial or economic value is expected, and how will it be proven?	Baseline cost, NPV/ROI model, benefit owner, management account, value attribution method
2. Govern the risk	What risks does the AI reduce, create or transfer?	Risk register, control map, AI system classification, assurance plan, incident protocol
3. Protect the customer	Does AI improve or damage the customer, student, citizen or staff experience?	NPS/CSAT/CES, handover metrics, service blueprint, complaints and escalation analysis
4. Control the model	Is the AI system reliable, explainable, fair, secure and cost-effective?	Model inventory, validation record, drift monitoring, bias testing, explainability evidence, FinOps reporting

Why this is not just an AI strategy

An AI strategy often begins with use cases. An AI value review begins with organisational value leakage. That distinction is important. If the organisation starts with use cases, it may automate low-value work, reproduce poor processes or purchase technology that solves a visible symptom rather than the underlying problem. If it starts with value leakage, it can choose the appropriate intervention: process redesign, policy clarification, workflow automation, data repair, staff training, customer service redesign, vendor consolidation, AI augmentation or full AI automation.

This also improves the credibility of the financial case. Savings claimed from AI are often overstated because they assume that time saved automatically translates into cost savings. In reality, time savings only become financial value when they are converted into avoided costs, increased throughput, redeployed capacity, reduced overtime, lower error costs, improved retention, faster revenue conversion, or a better customer outcome. The job of the value review is to make those conversion pathways explicit.

2. The Evidence Base: Adoption Is Rising Faster Than Value Capture

The current research landscape points in a consistent direction. AI adoption is growing rapidly, but value capture depends on management discipline, governance, process redesign and measurement capability. This is precisely where traditional consulting disciplines remain relevant.

AI adoption is no longer the differentiator

Gartner reported in March 2026 that AI deployment had grown from two out of five organisations in 2024 to four out of five organisations at the time of the survey. The differentiator is therefore not whether an organisation has AI. The differentiator is whether it has the foundations, financial guardrails and governance to derive value from it.

The same Gartner release is particularly important because it reframes ROI as more than a financial measure and identifies financial guardrails as a weakness. Gartner reported that only 44 percent of organisations surveyed had adopted financial guardrails or AI FinOps practices. This supports the need for a practical framework that links AI ambition, foundations, people, cost and governance.

High performers are management-system performers

McKinsey's 2025 State of AI work emphasises that AI high performers use management practices across strategy, talent, operating model, technology, data, adoption and scaling. The significance is clear: AI value is not simply a model-performance issue. It is a management-system issue.

This finding aligns with the practical experience of organisational transformation. Tools create capability, but operating models realise value. A workflow redesigned around AI, supported by capable staff, governed by risk controls and measured through meaningful KPIs, is very different from a workflow where staff simply have access to an AI assistant.

ROI must be measured through multiple dimensions

Deloitte's 2025 AI ROI research is useful because it does not define AI success through a single return calculation. It combines direct financial return, revenue growth from AI, operational cost savings and the speed at which results are achieved. That multi-factor view supports the Orbisco approach: boards need a balanced view of value, not a single simplistic ROI number.

Deloitte also argues that organisations should treat AI as an opportunity to rethink business models rather than merely improve efficiency. That is highly relevant to Orbisco's positioning. Efficiency may produce early savings, but strategic value often requires deeper redesign of work, service models, decision processes and capability.

Regulation and governance are now part of the value case

AI governance cannot be treated as a compliance afterthought. NIST describes its AI Risk Management Framework as a voluntary framework for better managing risks to individuals, organisations and society. ISO/IEC 42001:2023 provides an international AI management system standard, published in December 2023. The European Commission's AI Act guidance shows that non-compliance can carry significant penalties, including thresholds up to EUR35 million or 7 percent of worldwide annual turnover for certain infringements.

The implication is not that every Australian or New Zealand organisation is directly subject to every element of EU law. The point is broader: AI governance expectations are becoming more formal, more auditable and more important to procurement, assurance and stakeholder trust. Organisations that can evidence responsible AI management will be better placed than those relying on informal policies and ad hoc controls.

Traditional benefits management provides the missing bridge

The Australian Government's Benefits Management Policy for digital and ICT-enabled investments is a useful local anchor. It requires investment proposals to address purpose, anticipated outcomes, how improvements are measured, and how improvements are monitored and optimised. The UK Green Book similarly defines appraisal as the assessment of the costs, benefits, and risks of options for achieving objectives. These are not AI documents, but they provide exactly the discipline AI programmes often lack.

The UK Digital and Data Benefits Framework also shows how digital value can be made concrete through categories such as reduced failure demand, reduced paper handling, saved time, avoided duplication, reduced legacy cost and improved technology resilience. These categories translate naturally into an AI-enabled value review.

Table 1. Evidence base supporting an integrated AI value review

Source	Relevance to the Orbisco paper	Practical implication
McKinsey State of AI 2025	Value capture depends on management practices across strategy, talent, operating model, technology, data, adoption and scaling.	AI success requires organisational redesign, not tool deployment alone.
Gartner March 2026 AI Value	AI deployment is widespread, but financial guardrails and AI FinOps remain underdeveloped.	AI programmes need cost transparency and portfolio financial management.
Deloitte AI ROI 2025	AI ROI leaders measure direct return, revenue growth, operational savings and speed of results.	Use a balanced ROI approach rather than one simplistic metric.
Australian Benefits Management Policy	Digital investments should define purpose, outcomes, measurement and optimisation.	AI initiatives need benefits owners, baselines, measures and monitoring.
NIST AI RMF / ISO 42001 / EU AI Act	AI risk and governance are increasingly formalised and auditable.	Risk, control and compliance evidence must sit beside ROI evidence.

3. Why Traditional Cost-Benefit Discipline Still Matters

There is a temptation to assume that AI requires entirely new management methods. It does not and in fact, must not. AI requires some new measurements, but the foundations remain familiar: cost-benefit analysis, benefits realisation, management accounts, risk management, customer measurement, operational KPIs and governance reporting.

What changes is the way these disciplines are integrated. AI is not a conventional system implementation where cost, benefit and risk are usually visible at the point of approval. AI value can be non-linear, indirect, cumulative and dependent on adoption. Costs can be variable and hidden, particularly where usage-based compute, inference, data preparation, model maintenance and vendor consumption are involved. Risks can also be dynamic, including data leakage, hallucination, model drift, bias, overreliance, regulatory exposure and reputational harm.

The weakness of naive AI ROI

A naive AI ROI model usually looks something like this: estimate the number of hours saved, multiply by the salary cost, subtract license fees and implementation costs, then claim the difference as savings. This is rarely robust enough for an executive investment decision.

First, not all time saved becomes value. If staff save ten minutes per task but the organisation does not reduce costs, increase throughput, improve service, reduce errors, or redeploy time to higher-value activities, the value remains theoretical. Second, AI often shifts work rather than removes it. Human effort may move from drafting to review, from customer handling to exception management, or from analysis to interpretation. Third, AI can introduce new costs: data preparation, governance, monitoring, security, training, audit and change management.

A serious AI ROI model therefore needs to measure both value creation and value conversion. It should ask not only what AI can do, but how the organisation will convert that capability into a recognised benefit.

The correct sequence

Table 2. Traditional review sequence with AI-specific extensions

Step	Traditional review question	AI-specific extension
1. Baseline	What is the current cost, cycle time, volume, error rate, service level and risk exposure?	What data exists to measure AI impact before deployment?
2. Diagnose leakage	Where is waste, duplication, delay, rework, risk or poor service occurring?	Which leakage points are suited to AI, automation, analytics or better workflow?
3. Define benefit	What measurable improvement is expected?	Is the AI benefit cost reduction, capacity release, risk reduction, quality improvement, revenue uplift or service improvement?
4. Model value	What is the expected net benefit, confidence range and payback period?	How are variable compute costs, model maintenance, governance and adoption risk included?
5. Assign ownership	Who owns the benefit after the project team leaves?	Which business owner is accountable for converting AI output into realised value?
6. Govern risk	What are the key controls and assurance requirements?	What AI-specific controls are needed for data, model, bias, explainability, security and human oversight?
7. Track and optimise	How will benefits be monitored and optimised?	What dashboard, RAG thresholds and intervention triggers will be used?

Cost categories that are often missed

AI cost modelling should go beyond licences and implementation fees. The more mature the AI deployment, the more important total cost of AI ownership becomes. Cost categories include discovery and process review, data preparation, integration, platform costs, model provider costs, usage-based inference costs, cloud and compute, cybersecurity, model validation, human oversight, compliance documentation, staff training, change management, internal communications, external assurance and ongoing monitoring.

The strongest business cases make these costs visible from the beginning. That does not mean every pilot needs a heavy process. It means that material AI investments should be assessed with enough discipline to prevent apparent productivity gains from being consumed by hidden operating costs.

Benefit categories that are often undervalued

Traditional AI business cases often overstate labour savings and understate risk reduction, quality improvement and customer experience value. For example, an AI-assisted compliance review tool may not reduce headcount, but it may reduce the probability of missed obligations, improve the speed of compliance review and create a stronger audit trail. Similarly, AI-assisted customer service may not reduce staff cost immediately, but it may reduce failure demand, improve first-contact resolution and enable human staff to focus on complex matters.

A mature value framework should therefore recognise multiple benefit types: direct cost reduction, avoided cost, capacity release, increased throughput, error reduction, risk reduction, revenue uplift, working capital improvement, service quality improvement, compliance resilience and strategic option value.

4. The Orbisco Integrated AI Value Review

The Orbisco Integrated AI Value Review is a structured diagnostic and implementation framework. It begins with organisational performance rather than AI capability. Its purpose is to identify where value is leaking, determine which interventions are most suitable, and build a measurable pathway from intervention to realised outcome.

> *AI is not the first question. Value is the first question. AI is one of the possible answers.*

The eight-stage review model

Table 3. The Orbisco Integrated AI Value Review model

Stage	Purpose	Outputs
1. Executive intent	Clarify what the organisation is trying to improve and why it matters.	Executive value thesis; scope; priority areas; risk appetite boundary
2. Baseline and evidence	Establish the current position before AI intervention.	Cost baseline; process baseline; customer baseline; risk baseline; data availability map
3. Value leakage map	Identify waste, delay, duplication, rework, risk exposure and service friction.	Opportunity map; leakage categories; indicative value pools
4. Intervention options	Decide whether the solution is process redesign, technology, AI, governance, capability or a combination.	Options analysis; non-AI alternatives; AI suitability assessment
5. Benefit model	Quantify expected value, confidence level and conversion pathway.	Risk-adjusted business case; benefits register; value owner; assumptions log
6. Governance and assurance	Define the controls required to deploy responsibly.	AI system classification; control map; approval pathway; audit evidence requirements
7. Implementation roadmap	Sequence quick wins, foundational work and scaled deployment.	30-60-90 day plan; three-horizon roadmap; resourcing model
8. Measurement and optimisation	Monitor whether value is being realised and intervene where necessary.	AI Value Scorecard; RAG thresholds; review cadence; optimisation backlog

Three levels of output

The review produces outputs at three levels. At the board level, it provides a concise view of value, risk, customer impact, and governance status. At the executive level, it provides a portfolio view of cost, benefits, delivery progress, adoption, and control effectiveness. At the operational level, it creates the detailed performance, model, process, and customer metrics needed to manage the system day-to-day.

This separation matters. Boards do not need a dashboard full of model-accuracy metrics without business context. Operational teams do not need vague strategic statements that do not help them manage model drift, handover failures or cost per inference. The Orbisco approach connects the layers: operational metrics drive executive indicators, and executive indicators support board-level decisions.

AI suitability filter

Not every problem should be solved with AI. A useful value review includes an AI suitability filter before recommending a technology intervention. The filter asks whether the problem has a measurable baseline, a repeatable task or decision pattern, available data of sufficient quality, acceptable risk, a clear human oversight model, a plausible benefit conversion pathway and a business owner able to sustain the change.

Table 4. AI suitability filter

Suitability factor	Green signal	Red signal
Measurable baseline	Current cost, quality, time or risk is known.	No baseline; value would be speculative.
Process readiness	Process is stable enough to redesign or automate.	Process is chaotic, undocumented or frequently changing.
Data readiness	Relevant data is available, accessible and governed.	Data is incomplete, fragmented, sensitive or poorly understood.
Risk appetite	Use case sits within acceptable risk and control boundaries.	Use case affects vulnerable users, high-stakes decisions or legal rights without controls.
Human oversight	A clear review, escalation or exception model exists.	Automation would operate without meaningful supervision.
Value conversion	Time or quality gains can be converted into recognised benefit.	Efficiency gains cannot be converted into financial or service value.

The role of non-AI improvement

A distinctive strength of this approach is that it does not force an AI answer. Many organisations have inefficient processes, poor governance or weak data foundations that should be fixed before AI is introduced. A credible advisor should be able to tell a client when AI is not the right first step. This increases trust and protects the client from expensive experimentation that produces little organisational value.

5. The Four Pillars: Money, Risk, Customer and Model

The four pillars translate the value review into an operating framework. They are deliberately simple at the top level so that executive teams can remember them, yet sufficiently detailed underneath to support rigorous implementation.

Pillar 1 - Measure the Money

The financial pillar asks whether AI investment can be measured with the same seriousness as other capital and operating investments. It covers NPV, ROI, internal rate of return, payback, total cost of AI ownership, benefit attribution and management accounts.

A strong financial model should use scenarios rather than single-point optimism. It should identify direct savings, avoided costs, revenue uplift, risk-adjusted value and capacity redeployment. It should also include all relevant costs: technology, implementation, data preparation, integration, compute, model maintenance, monitoring, training, governance and change management.

Table 5. Financial foundations for AI value

Financial discipline	Why it matters for AI	Minimum evidence
Risk-adjusted business case	AI value can be uncertain, delayed and non-linear.	Base/optimistic/pessimistic scenarios; assumption log; sensitivity analysis
Total cost of AI ownership	Usage-based costs and governance overhead can materially alter ROI.	Cost model including licences, compute, data, integration, support and monitoring
Benefit conversion model	Time saved is not automatically cost saved.	Defined pathway from AI output to recognised benefit
AI management accounts	Executives need recurring visibility, not one-off business cases.	Monthly AI cost statement and value attribution statement
AI FinOps	Cloud, model and compute costs need active financial control.	Cost per output; utilisation; idle cost; optimisation backlog

Pillar 2 - Govern the Risk

The governance and risk pillar asks whether AI is operating within a clear system of accountability, control and assurance. This includes AI policy, risk appetite, model inventory, AI system classification, risk assessment, control design, incident management, vendor management, privacy, security, ethical review and internal audit.

The governance spine can be built around existing frameworks. ISO/IEC 42001 provides an AI management system reference point. NIST AI RMF provides a practical language of Govern, Map, Measure and Manage. The EU AI Act provides a reminder that AI risk classifications, documentation, human oversight and transparency are moving from good practice toward regulatory expectation in many contexts.

Table 6. Governance and risk evidence

Governance element	Question	Evidence
AI inventory	Do we know what AI systems are in use?	Register of systems, owners, purpose, data inputs, risk rating and status
Risk appetite	Where will we allow, restrict or prohibit AI use?	Board-approved AI risk appetite statement
Control framework	What controls prevent, detect and respond to AI risk?	Control map linked to data, model, user and vendor risks
Assurance model	Who checks that controls are operating?	Three lines model; internal audit plan; independent validation process

Governance element	Question	Evidence
Incident response	What happens when AI fails, harms or misleads?	Escalation thresholds, notification process, remediation log

Pillar 3 - Protect the Customer

The customer and service pillar asks whether AI improves the experience of the people the organisation serves. In schools, that may include students, parents, staff and administrators. In government, it may include citizens and regulated entities. In business, it may include customers, suppliers and employees.

Customer measurement should not be limited to speed. AI can make a service faster while making it less trusted, less empathetic or less effective. The minimum measurement set should include satisfaction, effort, resolution, escalation, complaints, abandonment, trust and handover quality.

Table 7. Customer and service excellence measures

Customer measure	AI-specific question	Why it matters
NPS or relationship trust	Does AI-mediated service affect loyalty or trust?	AI failures can damage trust even where speed improves.
CSAT	Was the specific AI-assisted interaction satisfactory?	Captures immediate service quality.
Customer Effort Score	Did AI make the task easier or harder?	AI should reduce friction, not transfer effort to the user.
First-contact resolution	Can the AI or AI-assisted channel resolve the matter first time?	Measures actual effectiveness, not just response speed.
Handover quality	When AI escalates to a human, is context preserved?	Poor handover is a major source of frustration and rework.

Pillar 4 - Control the Model

The AI performance pillar asks whether the model or AI system is reliable enough to support the business process it has been placed inside. This includes model accuracy, precision, recall, hallucination management, drift monitoring, bias testing, explainability, latency, uptime, throughput, data quality and security.

The key principle is that technical metrics must be translated into business consequences. A small drop in model accuracy may be immaterial in a low-risk internal drafting assistant but unacceptable in a credit, safety, employment, health, education or compliance decision system. Measurement thresholds should therefore be set according to the use case, risk appetite and impact on affected people.

Table 8. AI-specific measurement and business translation

Technical measure	Business translation
Accuracy, precision, recall and F1	Are outputs good enough for the decision or task?
Drift monitoring	Is the model degrading as circumstances, data or behaviour change?
Bias and fairness metrics	Are outcomes materially different across protected or relevant groups?
Explainability coverage	Can the organisation explain material AI-assisted decisions?
Latency and availability	Is the AI system reliable enough for operational use?
Cost per inference or output	Is the system economically sustainable at scale?

6. The Orbisco AI Value Scorecard

The Orbisco AI Value Scorecard converts the four pillars into a reporting architecture. It is designed to avoid two common failures: board reports that are too technical to support executive decisions, and technical dashboards that cannot be traced to financial or service outcomes.

The scorecard uses three reporting tiers: Board, Executive and Operational. The board tier focuses on strategic value, risk and assurance. The executive tier focuses on portfolio performance, cost, delivery, service and governance. The operational tier focuses on the detailed technical and process signals that allow teams to intervene before value or risk deteriorates.

Board level - quarterly strategic indicators

Table 9. Board-level AI Value Scorecard

No.	Metric	Pillar	Purpose	RAG example
B1	AI programme portfolio NPV / value trajectory	Money	Shows whether the AI portfolio is expected to generate positive value.	Green: on or above trajectory; Amber: within tolerance; Red: materially below or unquantified
B2	AI-attributable P&L or economic contribution	Money	Connects AI interventions to financial performance.	Green: at/above target; Red: below target without recovery plan
B3	AI governance coverage	Risk	Shows whether material AI systems are under formal governance.	Green: near complete coverage; Red: material systems outside control
B4	Regulatory and policy compliance status	Risk	Shows whether AI obligations and internal policies are being met.	Green: no critical findings; Red: critical gap or breach risk
B5	AI-mediated customer/service trust indicator	Customer	Shows whether AI is improving or harming user experience.	Green: at/above benchmark; Red: deterioration or poor experience
B6	Aggregate AI risk status	Risk	Shows whether AI risk sits within appetite.	Green: within appetite; Red: appetite breached
B7	AI value-to-cost ratio	Money	Compares realised or forecast value against total cost of AI ownership.	Green: above threshold; Red: cost growing faster than value
B8	Material AI incidents	Risk	Tracks serious failures, harms or control breaches.	Green: zero; Red: material incident or unresolved systemic issue

Executive level - monthly portfolio indicators

Table 10. Executive-level AI Value Scorecard

No.	Metric	Pillar	Purpose
E1	Total cost of AI ownership	Money	Shows recurring AI cost across licences, compute, integration, governance and support.
E2	Cost per AI-automated transaction or output	Money	Tests whether AI unit economics improve as use scales.
E3	AI FinOps savings realised	Money	Captures active optimisation of model, cloud and infrastructure cost.

No.	Metric	Pillar	Purpose
E4	AI control and governance action completion	Risk	Tracks whether governance improvements are implemented on schedule.
E5	Model validation currency	Risk	Shows whether production models remain validated and fit for purpose.
E6	Open AI audit or assurance findings	Risk	Tracks unresolved assurance gaps and critical findings.
E7	CSAT for AI-mediated interactions	Customer	Measures immediate satisfaction with AI-supported service.
E8	Customer Effort Score for AI channels	Customer	Shows whether AI makes the service easier or harder.
E9	First-contact resolution in AI channels	Customer	Measures whether AI resolves real demand.
E10	AI-to-human escalation rate	Customer	Shows whether AI is over-escalating or failing to resolve.
E11	Automation rate in target processes	Model	Tracks operational progress against the process redesign case.
E12	AI system availability and reliability	Model	Shows whether systems are operationally dependable.

Operational level - weekly or daily indicators

Table 11. Operational AI Value Scorecard

No.	Metric	Pillar	Purpose
O1	Model performance against specification	Model	Monitors accuracy, precision, recall or relevant task-specific quality.
O2	Data drift or input distribution change	Model	Detects whether the system is being used in conditions different from training or validation.
O3	Bias or fairness disparity indicator	Model	Identifies material differences in outcomes across groups.
O4	Explainability coverage for high-impact decisions	Model	Ensures material AI-assisted decisions can be explained.
O5	Response latency and throughput	Model	Measures operational performance and user experience.
O6	Compute utilisation and idle cost	Money	Identifies waste in AI infrastructure or model usage.
O7	Exception handling rate	Model	Shows how often human correction or intervention is required.
O8	AI-relevant data quality score	Risk	Shows whether the underlying data estate supports reliable AI.

RAG reporting discipline

The scorecard should not simply report numbers. It should drive decisions. Every metric needs a defined owner, reporting cadence, threshold, escalation pathway and recovery action. Where possible, thresholds

should include trend indicators. A stable amber may be less concerning than a rapidly deteriorating green. A red metric without an owner is not governance; it is theatre.

Table 12. RAG discipline for AI governance

RAG status	Meaning	Governance response
Green	Within target or risk appetite, with no material deterioration.	Continue monitoring; capture improvement opportunities.
Amber	Within tolerance but trending adversely, or minor control/performance issues present.	Assign owner; agree recovery action; monitor at increased cadence.
Red	Outside tolerance, material risk, poor performance, unquantified value or critical control gap.	Escalate; consider pause, remediation, investment decision or risk acceptance.

7. Implementation Roadmap and Governance Model

AI value measurement should be implemented in horizons. Attempting to build a complete enterprise framework in one step can create delay and resistance. The better approach is to establish minimum viable governance and measurement quickly, then deepen capability as the organisation scales.

Orbisco Four-Stage AI Transformation Framework

The four-stage transformation framework provided by Orbisco fits naturally into this paper as the practical delivery model. The Integrated AI Value Review identifies where value is leaking and what should be done. The four-stage framework explains how the organisation moves from diagnosis to scalable transformation and continuous improvement.

In simple terms, the model is: Steer, Scale, Institutionalise and Evolve. It gives executives an accessible language for the change journey, while still aligning with the deeper scorecard, governance and benefits-realisation architecture described throughout this paper.

Table 13. Orbisco Four-Stage AI Transformation Framework

Stage	Purpose	What Orbisco does	Typical outputs
Steer	Identify high-impact use cases, assess readiness and define quick wins.	Orbisco focuses first on real business impact: productivity, cost reduction, better decisions and improved service. The work includes process review, data-quality assessment, readiness analysis, KPI definition and prioritisation of practical opportunities.	AI opportunity map; readiness assessment; KPI set; 30-60-90 day roadmap; 2-3 quick wins
Scale	Build reusable models, data pipelines and strong foundations.	Orbisco turns early wins into repeatable capability. Rather than building isolated solutions, the emphasis is on reusable models, data pipelines, implementation playbooks and best-practice approaches that can be applied across teams and functions.	Reusable patterns; data and integration foundations; scaled-use-case backlog; operating model requirements; capability uplift plan

Stage	Purpose	What Orbisco does	Typical outputs
Institutionalise	Embed AI into governance, teams and everyday operations.	Orbisco helps integrate AI into normal management systems: governance, accountability, risk, privacy, ethics, compliance, business ownership and team adoption. This is where AI moves from experiment to disciplined operating practice.	Governance model; control framework; benefits ownership; adoption plan; policy and assurance artefacts; training approach
Evolve	Measure results, improve continuously and scale what works.	Orbisco establishes the feedback loops needed to track adoption, performance, business outcomes and risk. The aim is to refine solutions over time, reuse assets, reduce rework and accelerate future deployment.	AI Value Scorecard; benefits tracking; optimisation backlog; lessons learned; reusable assets; continuous improvement cadence

> *Orbisco helps organisations turn AI from isolated activity into practical, measurable transformation.*

The four stages also create a useful bridge between executive messaging and implementation detail. Steer provides direction and prioritisation. Scale converts isolated wins into organisational capability. Institutionalise embeds governance, accountability and adoption. Evolve ensures the organisation continues to measure, improve and reuse what works.

Three-horizon roadmap

Table 14. Three-horizon implementation roadmap

Horizon	Timeframe	Primary aim	Key deliverables
Horizon 1: Foundation	0-6 months	Create visibility and control.	AI asset register; baseline cost and benefit measures; priority value leakage map; initial risk register; pilot scorecard; customer/service baseline; executive reporting cadence.
Horizon 2: Integration	6-18 months	Embed AI value into normal management systems.	Full benefits register; AI management accounts; governance framework; validation process; customer measurement suite; board scorecard; FinOps optimisation; roadmap for high-value use cases.
Horizon 3: Optimisation	18-36 months	Move from reactive reporting to predictive value management.	Predictive leading indicators; mature scorecard; ISO 42001 readiness or certification pathway; continuous improvement backlog; portfolio investment model; advanced risk and assurance reporting.

90-day diagnostic sprint

For most clients, Orbisco can begin with a 90-day diagnostic sprint. This is commercially attractive because it is short enough for executives to approve, but substantive enough to produce a credible value map and implementation roadmap.

Table 15. Suggested 90-day Orbisco diagnostic sprint

Phase	Days	Focus	Outputs
Phase 1	Days 1-15	Executive alignment and evidence request	Scope; interviews; document review; data request; priority process list
Phase 2	Days 16-35	Baseline and value leakage analysis	Cost and process baseline; customer/service baseline; risk and governance scan
Phase 3	Days 36-55	Opportunity and AI suitability assessment	Opportunity map; non-AI vs AI intervention analysis; risk classification
Phase 4	Days 56-75	Benefits and governance design	Preliminary business cases; benefits register; control requirements; scorecard draft
Phase 5	Days 76-90	Executive roadmap and decision pack	Board/executive paper; three-horizon roadmap; quick wins; investment options

Governance structure

A practical AI governance structure should use existing decision forums where possible, rather than creating a parallel bureaucracy. The key is to make responsibilities clear. Executive leadership should own AI value and risk appetite. Business owners should own benefits. Risk, compliance and technology leaders should own control design and assurance. Operational teams should own performance, monitoring and incident management. Internal audit or an equivalent assurance function should provide independent review.

Table 16. AI governance responsibilities

Forum or role	Responsibility	Cadence
Board or board committee	Approve AI risk appetite; review strategic value, material risk and assurance status.	Quarterly
Executive AI Steering Committee	Prioritise use cases; approve investment; monitor portfolio value and risk.	Monthly or quarterly
AI Value Office / PMO	Maintain benefits register, scorecard, assumptions log and implementation roadmap.	Monthly
Risk, compliance and legal	Maintain control framework, regulatory mapping, policy and incident protocols.	Monthly and as required
Technology and data	Own data readiness, architecture, security, model operations and cost optimisation.	Weekly/monthly
Business process owners	Own adoption, benefit conversion and operational performance.	Monthly
Internal audit / independent assurance	Review control design and operating effectiveness.	Annual plan plus targeted reviews

Common failure modes

The most common failure is measurement after deployment. If the baseline is not established before the intervention, value becomes difficult to prove. A second failure is treating governance as documentation rather than decision-making authority. A third is allowing technical teams to report technical metrics without translating them into business consequences. A fourth is assuming that AI adoption equals productivity. A fifth is ignoring customer experience until service problems have already emerged.

The Orbisco framework is designed to prevent these failures by requiring baseline evidence, benefit ownership, risk controls, customer measurement and operational metrics from the beginning.

8. Recommendations for Boards and Executive Teams

The following recommendations are intended for board and executive teams considering material AI investment or seeking to bring existing AI activity under stronger value and governance discipline.

1. Stop asking only where AI can be used.

Begin with value leakage: cost, delay, duplication, risk, errors, poor customer experience and underused capability. Then decide whether AI is the right intervention.

2. Establish baselines before implementation.

No baseline means no credible value case. Capture current cost, volume, cycle time, quality, risk exposure and customer/service levels before AI intervention.

3. Treat time saved as a benefit hypothesis, not a realised saving.

Time saved becomes value only when converted into cost avoidance, increased throughput, redeployed capacity, improved quality, risk reduction or revenue improvement.

4. Create a total cost of AI ownership model.

Include licences, compute, data, integration, security, monitoring, governance, validation, training and change management.

5. Build financial guardrails early.

Use budget thresholds, cost-per-output metrics, AI FinOps practices and approval gates before usage-based costs scale unexpectedly.

6. Put AI into the enterprise risk framework.

AI risk should not sit in a side register. It should be mapped into enterprise risk, control and assurance structures.

7. Maintain an AI system inventory.

You cannot govern, secure, validate or optimise systems you have not catalogued.

8. Measure customer and service impact directly.

Response speed is not enough. Measure satisfaction, effort, resolution, escalation, trust and handover quality.

9. Translate technical performance into business consequence.

Accuracy, drift, bias and latency matter because of what they do to decisions, customers, risk, cost and trust.

10. Use a board-level scorecard.

Executives need a concise view of value, risk, customer impact and model control, supported by detailed operational evidence.

The final test

The final test of an AI programme is not whether it has impressive technology. The final test is whether the organisation can explain, with evidence, what value was expected, what value was realised, what risks were introduced or reduced, what customers or users experienced, and what management action was taken when performance changed.

That is the discipline required to close the AI value gap.

Appendix A: Diagnostic Questions for an AI Value Review

These questions can be used in executive interviews, workshops or document reviews. They are designed to surface whether an organisation is ready to move from AI activity to AI accountability.

Area	Diagnostic questions
Executive intent	What organisational outcomes are we trying to improve? Which outcomes matter most to the board? What risk appetite applies to AI?
Current AI activity	What AI systems, tools, pilots or informal uses already exist? Who owns them? Which are material?
Financial baseline	What is the current cost of the relevant process or service? What are the cost drivers? What data exists to prove change?
Process and operating model	Where does rework, duplication, delay, manual handling or avoidable escalation occur? Has the process been redesigned before considering AI?
Benefits	What type of benefit is expected: cost reduction, avoided cost, capacity release, revenue uplift, risk reduction, service improvement or quality improvement? Who owns it?
Data readiness	Is the required data available, accurate, accessible, secure and legally usable? What data gaps would limit AI reliability?
Risk and compliance	What could go wrong? Who is affected? What legal, regulatory, ethical or reputational exposure exists?
Customer and service	How will AI affect satisfaction, effort, trust, resolution, escalation and handover?
Model and technology	What model or system is being used? How will performance, drift, bias, explainability, latency and cost be monitored?
Governance	Who can approve, pause, change or retire the AI system? What evidence will be available for audit or assurance?

Appendix B: AI Benefits Register Template

The benefits register is the central artefact connecting AI activity to organisational outcomes. It should be owned by the business, not only by technology or project teams.

Benefit ID	Benefit statement	Baseline	Target	Measure	Owner	Timing	Confidence	Evidence source
B-001	Reduce manual processing time in selected workflow.	Current average handling time per case.	Target reduction or capacity release.	Minutes per case; cases per FTE; cost per case.	Business process owner.	Pilot, 3 months; scale, 12 months.	Medium until baseline verified.	Workflow data; timesheets; case management system.
B-002	Improve first-contact resolution for AI-assisted service enquiries.	Current FCR and escalation rate.	FCR improvement; escalation reduction.	FCR percentage; escalation percentage; CSAT.	Service owner.	6-12 months.	Medium.	CRM; survey; quality assurance review.
B-003	Reduce compliance review risk through AI-assisted document triage.	Current missed issues, review time and audit findings.	Lower review time and stronger audit evidence.	Review duration; QA pass rate; audit findings.	Compliance owner.	6 months.	Low-medium until validation complete.	Compliance records; QA sample review.

Benefit conversion rules

Every claimed benefit should specify how it becomes recognised value. Time saved must be connected to avoided cost, redeployment, increased throughput or service improvement. Risk reduction must specify the baseline risk, expected reduction and method for estimating avoided loss. Customer improvements must specify the metric, baseline and expected movement. Strategic option value should be described but kept separate from hard financial return unless it can be reasonably quantified.

Appendix C: Source and Claim Verification Register

This rebuilt version uses a conservative sourcing approach. Exact numerical claims from the supplied Markdown have been retained only where they were verified from accessible sources during this review or where they are presented as internal assumptions requiring confirmation.

Claim or source area	Verification status in this version	Use in paper
Gartner: AI deployment grew from two out of five organisations in 2024 to four out of five in 2026; 44% have financial guardrails or AI FinOps.	Verified from Gartner press release, 9 March 2026.	Used as a key evidence point for adoption rising faster than financial discipline.
McKinsey: high-performing AI organisations use management practices across strategy, talent, operating model, technology, data, adoption and scaling.	Verified from McKinsey State of AI 2025 article.	Used to support the management-system argument.
Deloitte: AI ROI Performance Index combines direct financial return, revenue growth, operational savings and speed; AI ROI leaders rethink business models.	Verified from Deloitte Netherlands article, 20 November 2025.	Used to support multi-dimensional ROI.
NIST AI RMF: voluntary framework for managing AI risk.	Verified from NIST AI RMF page.	Used to support governance and risk pillar.
ISO/IEC 42001:2023 published December 2023.	Verified from ISO standard page.	Used to support AI management system references.
EU AI Act penalties up to EUR35m or 7% of worldwide annual turnover for certain infringements.	Verified from European Commission AI Act FAQ.	Used to support regulatory risk discussion.
Australian Benefits Management Policy requires purpose, anticipated outcomes, measurement and optimisation for digital/ICT-enabled investment proposals.	Verified from digital.gov.au.	Used to connect AI ROI with benefits management.
UK Green Book defines appraisal as assessing costs, benefits and risks of options for achieving objectives.	Verified from GOV.UK Green Book page.	Used to connect AI investment to traditional appraisal.
Fewer than 5% of enterprises capture measurable AI value.	Not verified from a primary accessible source during this pass.	Removed from core claims; may be reinstated if source is supplied.
Only 29% of executives can measure AI ROI confidently.	Not verified from a primary accessible source during this pass.	Removed from core claims; may be reinstated if source is supplied.
67% of AI impact attributable to organisational factors versus 32% technology.	Not verified from a primary accessible source during this pass.	Not used as a firm statistic; argument retained qualitatively.

References

- McKinsey & Company. The State of AI: Global Survey 2025, 5 November 2025.
- Deloitte Netherlands. Turning AI into ROI: what successful organisations do differently, 20 November 2025.
- Gartner. Gartner Identifies Three Pillars for Deriving Value from AI, 9 March 2026.
- National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0), released 26 January 2023.
- International Organization for Standardization. ISO/IEC 42001:2023, Artificial intelligence - Management system.
- European Commission. Navigating the AI Act, FAQ updated 28 January 2026.
- Australian Government Digital Transformation Agency. Benefits Management Policy, last updated 8 October 2025.
- HM Treasury and Government Finance Function. The Green Book, updated 5 February 2026.
- UK Government. Digital and Data Benefits Framework, 7 April 2026.

- Association for Project Management. Benefits management and project success, Body of Knowledge 8th edition summary.

Appendix D: Glossary

Term	Meaning
AI FinOps	Financial operations discipline applied to AI and cloud costs, including cost visibility, accountability and optimisation.
AI management accounts	Recurring management reports that show AI costs, benefits, risks and value attribution.
AI system inventory	A register of AI tools, models and systems, including owners, purpose, data inputs, risk rating and governance status.
Benefit conversion	The pathway by which an AI-enabled improvement becomes recognised organisational value.
Data drift	A change in the input data distribution that may reduce model reliability.
Explainability	The ability to explain why an AI system produced a particular output or recommendation.
Model validation	Independent testing and review of whether an AI model is fit for its intended use.
RAG reporting	Red/Amber/Green status reporting used to translate performance and risk data into governance signals.
Total cost of AI ownership	The full cost of deploying and operating AI, including technology, compute, data, integration, governance, monitoring and change management.
Value leakage	The loss of organisational value through avoidable cost, delay, rework, risk, poor service or underused capability.

End of paper.