AI Investment · Executive Working Playbook

For boards, CEOs, CFOs, CIOs, CTOs, COOs, and Chief Data / AI Officers

AI investments deserve ROI cases that that survive the boardroom.

A working playbook for building, defending, and staging-gate staging-gate AI business cases that finance, risk, and the the board will sign.

Audience

CEO · CFO · CIO · CTO · COO · CDO/CAIO · Strategy · Board

Outcomes

Board-ready business cases · Risk-adjusted ROI · 90-day action plan

The question on the table

How does the executive team construct an AI ROI case that captures captures the **total value of change**, prices the **total cost to serve**, **serve**, and still wins board scrutiny?

1 · VALUE	2 · COST	3 · MEASURE	4 · GOVERN
Hard, soft, and strategic stack	Eight categories, not a SaaS bill	Six-layer KPI architecture	Trustworthy AI, board-grade

Synthesis prepared for executive briefing · All cited sources documented in the appendix

Five disciplines separate board-ready AI cases from pilot theater. theater.

A credible AI ROI case requires more than a vendor deck and a labor-savings spreadsheet. The cases boards approve look like the cases boards can defend.

01	**Treat AI as transformation, not procurement**	Stage against an operating-model redesign — workflows, data, governance, talent — not against a SaaS replacement cycle.
02	**Capture the full value stack**	Hard financial value, soft operating value, and strategic option value — labor savings alone understates the prize and overstates the conviction.
03	**Price the full cost to serve**	Inference, integration, verification, governance, change, and lock-in — the cloud bill is roughly the smallest line on the page.
04	**Measure what the business owns**	Six metric layers — adoption, productivity, quality, risk, financial, strategic — with the counterfactual explicit and the cohort tracked.
05	**Govern for board-grade risk**	Trustworthy-AI controls, stage gates from pilot to production to scale, and quarterly board review tied to risk-adjusted ROI.

Synthesis · see slides 03 through 11 for the supporting argument and slides 14 through 15 for what comes next.

The credibility problem

Most AI cases are rejected long before they reach scale — and boards have good reasons.

25% of AI initiatives delivered their expected ROI in IBM's 2025 C-suite study — only **16% scaled enterprise-wide**.

Board engagement matters

26% of boards discuss AI at every meeting.

In 63% of high-ROI organisations AI is on every agenda vs 13% in low-ROI counterparts.

Four reasons boards reject pilot-stage cases

Pilots overstate wins and understate cost

Cherry-picked cohorts, no counterfactual, hand-tuned prompts, and rollout costs billed to other budgets make pilots look better than the production system will.

Value follows a J-curve, not a step

Learning curve, verification tax, and pipeline adaptation leave early ROI negative; Google Cloud calls out realistic financial modeling as the lever.

Adoption, not model performance, is the bottleneck

Workflow redesign, verification, and data readiness slow realization; Microsoft reports 79% see AI as crucial yet 59% struggle to quantify productivity gains.

Outcomes are not tied to business metrics

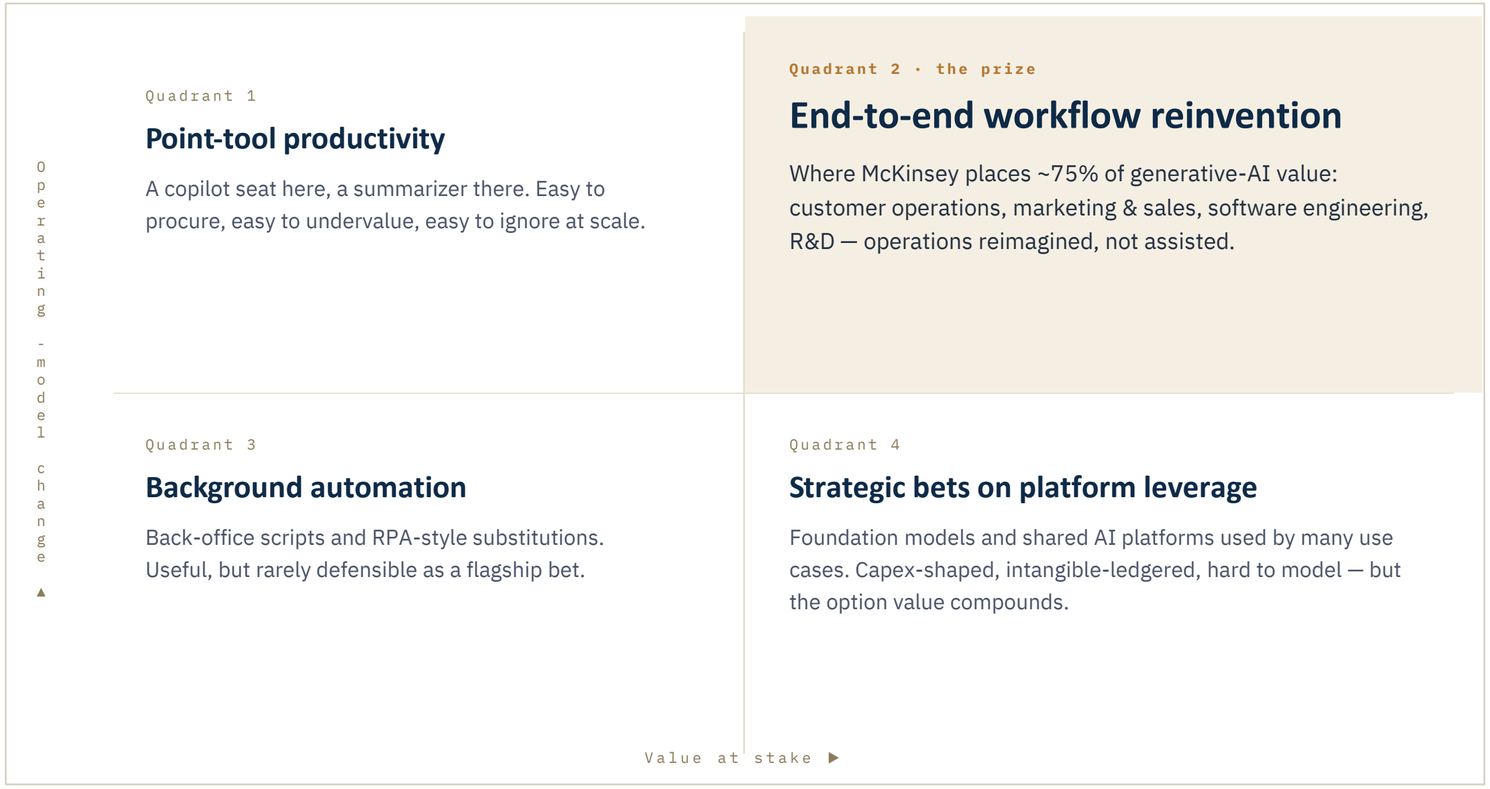
Activity metrics (prompts, calls, hours saved) replace value metrics (margin, NPS, risk reduction); boards discount what cannot be linked to P&L or risk.

Sources: IBM 2025 C-suite study · Google Cloud / DORA · Microsoft Worklab · Protiviti global board survey.

Reframing the investment question

AI is an operating-model investment — price it that way and the ROI discussion changes.

Boards fund transformations; they approve tool purchases. The same dollars read differently when scored against different lens — McKinsey warns against paying a "gen-AI premium" for projects rankable on business case alone.



So what

Stop sizing AI cases against software software contracts. Size them against against the operating model.

If a project sits in Quadrant 1 or 3 it can be funded inside existing OpEx; the discussion is payback and adoption. If it sits in Quadrant 2 or 4 the discussion is transformation, portfolio mix, and capital allocation — which is exactly where boards belong.

McKinsey (2024) places the generative-AI prize at **$2.6T–$4.4T annually**, with about 75% concentrated in four functions.

Synthesis · value-at-stake axis calibrated to McKinsey, Deloitte, and IBM C-suite framing.

Total value of change · the upside half of the case

A credible case captures hard financial value, soft operating value, and strategic option value — and refuses to collapse them into “labor savings”

Tier 1 ▸ Hard financial value

What shows on the P&L this this year.

Direct cost reduction

Headcount substitution, processing cost, channel cost, vendor consolidation.

Revenue uplift

Conversion lift, win-rate, cross-sell, expanded wallet share.

Margin expansion

Mix shift, premium capture, productivity gains net of reinvestment.

Working capital

DSO/DPO shifts, inventory turns, capex deferral.

Tier 2 ▸ Soft / intermediate value

What compounds before it converts.

Throughput

Cycle time, queue depth, time-to-decision, jobs-per-FTE.

Quality & rework

Error rates, first-pass yield, compliance exceptions.

Experience

NPS, CSAT, employee eNPS, ease-of-use.

Speed

Time to insight, response time, prototype-to-production cadence.

Tier 3 ▸ Strategic / option value

What buys the next decade.

Innovation cadence

Shorter idea-to-launch loop, more experiments per dollar.

Platform leverage

One foundation model, many use cases — IBM advises open orchestration.

Resilience & risk

56% of gen-AI users report improved security posture (Google).

Data & regulatory option

Curated data assets and audit trails convertible to future products.

Synthesis · Microsoft and Google frame value across operational efficiency, revenue growth, experience, and strategic advancement / risk mitigation. Microsoft reports AI value lives in six dimensions; this slide condenses those into three tiers for board presentation.

Total cost of AI enablement

Price the full stack — eight categories boards expect, two of which most which most decks forget.

Common blind spots: Governance, change management, and verification effort are routinely left out of board materials — and are routinely the largest variance between projected and actual costs.

#	Cost category	What it covers	Typical share
01	**Model, API & subscription**	Licensed model seats, foundation-model APIs, vendor subscriptions.	15–25%
02	**Token, inference & agentic**	Per-token spend & orchestration across multi-step / multi-agent workflows. **Scales non-linearly** with autonomy.	**10–30%**
03	**Infrastructure & data**	GPU/compute, vector stores, data pipelines, observability, custodianship.	10–15%
04	**Integration & workflow redesign**	Systems integration, UX redesign, exception paths, telemetry back into source systems.	10–15%
05	**Risk, compliance & auditability**	Responsible-AI controls, bias testing, security review, regulatory readiness. **Often omitted.**	**5–15%**
06	**Human review, QA & exceptions**	Spot-checks, edge cases, "human-in-the-loop" overrides, model reviews. **Scales with autonomy.**	**5–20%**
07	**Change, training & adoption**	Enablement, communications, incentive redesign, productivity dip coverage.	5–15%
08	**Lock-in & exit cost**	Architecture flexibility, model portability, re-platforming insurance.	2–10%

What changes the math

Visible cost ≠ total cost. Keep the visible visible line honest and the invisible line line billable.

Google Cloud illustrates the pattern: a single chatbot estimated at roughly **$24,910 monthly value** against **~$2,700 monthly cost** — payback in under two weeks — but the model explicitly priced governance and verification, not only the API.

Sourced breakdown; per AWS prescriptive guidance, baseline cost assessment and autonomy-level selection set the error-tolerance envelope.

Shares are working ranges for a typical enterprise pilot-to-scale programme; ties to AWS measurement guidance, Google Cloud value framework, and IBM open-orchestration advice. Token / inference is the row most under-modelled.

A seven-step framework that aligns business case, baseline, governance, and stage governance, and stage gates before any scale money moves.

Synthesis Built from McKinsey, Deloitte, IBM, Google Cloud, AWS, Microsoft, NIST, and FinOps frameworks — not a sourced single-source framework. Sources cited per step.

1 Select the use case

Rank the top 20–30 value-accretive projects regardless of AI tag; avoid a "gen-AI premium"; pick a small number of high-impact use cases.

McKinsey "A guide for CFOs"

2 Define baseline & counterfactual

Use a control group or matched cohort. Without a counterfactual the case is a memo, not a measurement.

IBM · AWS prescriptive guidance

3 Map value drivers

Tie each driver to one of the three tiers on slide 5; refuse to leave any line as "productivity" without an operational definition.

Microsoft six dimensions · Google value framework

4 Price total cost to serve

Build the eight-category ledger from slide 6; contingency for governance, verification, and exit; no SaaS-bill blinders.

AWS · IBM · Google Cloud

5 Risk-adjusted return

Apply scenario weights (P25 / P50 / P75). Pair payback, NPV / IRR with risk-adjusted ROI; let the board weigh tail risk explicitly.

Synthesis · Deloitte balanced portfolio

6 Operating metrics

Adoption, productivity, quality, risk, financial, strategic — all six layers, with leading / lagging split and cohort tracking.

See slide 9 · Microsoft, Google, IBM

7 Stage-gate cadence

Pilot → production → scale. Each gate has a decision rule, an owner, and a stop condition. Set the next-review cadence.

AWS decision points · Google J-curve

Reading order Use this framework as the spine of the business case. Slides 8 through 12 unpack steps 1, 6, and 7 — and the safeguards orgs forget.

Step 1 of the framework · selecting where to play

Pick use cases by value-and-feasibility, with the counterfactual written down — written down — not by executive enthusiasm.

Feasibility & readiness ▲ data · workflow · risk · capability

Quadrant A

Quick proof points

Low risk, fast to deploy. Fund inside existing OpEx. Build conviction and reusable foundations. Reinvest savings to fund bigger bets.

Counterfactual: it is tempting to skip one — write it down anyway.

Quadrant B · place flagship bets here

End-to-end reimagination

Where McKinsey estimates ~75% of generative-AI value lives: customer operations, marketing & sales, software engineering, R&D.

Total prize

$2.6T–$4.4T / yr

Concentration

≈ 75% in 4 fns

Quadrant C

Defer or descope

Low value, low readiness. Either park or use as cheap learning vehicles — do not let them consume programme leadership.

Quadrant D

Capability-build first

High value, low readiness. Treat as multi-quarter investments — invest in data, governance, and talent first; sequence the business cases behind the capability.

Value at stake ▶ cost reduction · revenue · risk · option

Counterfactual discipline

Without a counterfactual the case is a memo.

For every shortlisted use case, write down **three things** before the proposal leaves the working group:

1 · The do-nothing baseline

What the metric looks like 12 months in if no one does anything. Include the natural improvement the team would have made anyway.

2 · The matched cohort

Region, segment, role, or process twin that does not receive AI in the same window. Track it quarterly.

3 · The autonomy envelope

AWS prescriptive guidance: pick the autonomy level (assist / recommend / act) and the error-budget that goes with it. Cost follows.

Synthesis · McKinsey, IBM, AWS, and Deloitte informed the axes. The choice of axes is synthesis; the underlying benchmarking data is sourced.

Step 6 of the framework · measurement architecture

Track six metric layers, split leading from lagging, and never measure outcomes measure outcomes without the cohort that did not see the system.

Layer 1

Adoption

Active users, weekly depth, abandonment, workflow penetration.

Leading

Layer 2

Productivity

Cycle time, throughput per FTE, tasks closed, hours-to-task-output.

Leading

Layer 3

Quality

First-pass yield, error & rework, compliance exceptions, NPS / CSAT shifts.

Mixed

Layer 4

Risk

Bias incidents, security findings, audit exceptions, near-miss rate.

Lagging

Layer 5

Financial

Cost-to-serve, ARPU / margin, payback period, NPV / IRR against counterfactual.

Lagging

Layer 6

Strategic

New products shipped, data reuse, optionality created, resilience score.

Lagging

Timing — leading & lagging

Adoption and productivity move first; financial lags by one to two quarters.

Boards that ask only about cost and revenue at the 90-day review end up killing projects that are tracking. Pair lagging outcomes with leading signals so the scorecard reads in both directions.

Counterfactual

Always pair the AI cohort with a matched non-AI non-AI cohort.

Without it, every improvement looks attributable to AI — including the improvements the team would have shipped anyway. The cohort is the cheapest line item on the measurement budget and the most defensible.

Reference benchmark

45% of gen-AI users reporting reporting productivity gains saw saw productivity at least double.

74% of surveyed enterprises with gen AI report current ROI. 56% report improved security posture. Set your scorecard against comparable benchmarks, not against an abstract target.

Microsoft Worklab cautions that time saved alone is insufficient — pair with quality, effort, and creativity.

Sources: Google Cloud ROI study · Microsoft Digital AI Value Framework · IBM C-suite study · AWS prescriptive guidance. The six-layer split is a synthesis.

Board oversight · governance as risk surface

AI is a board-grade risk domain — boards that treat it that way post higher ROI; boards that do not, post less.

Board engagement correlates with ROI

The board's cadence is itself a itself a financial decision.

63% of high-ROI organisations have AI on every board agenda.

13% of low-ROI equivalents do.

26% of all boards discuss AI at every meeting — most do not.

Source: Protiviti global board survey.

Quadrant I · Accountability

Name an executive sponsor for every dollar of AI spend.

AI crosses legal, security, HR, finance, and product. Without a named sponsor the work routes to committee and never to a number.

Quadrant II · Risk controls

NIST AI RMF: govern, map, measure, manage — embedded across the lifecycle. lifecycle.

Trustworthy-AI controls: bias testing, transparency, accountability, security, privacy, validation. Risk-adjust every ROI to make tail events readable.

Quadrant III · Cadence

Establish a quarterly AI portfolio review — not — not annual.

Score the portfolio the way a private-equity firm scores its deals: kill, double, or hold. Document the decision and the data behind it each cycle.

Quadrant IV · Disclosure

Board materials align to audit-committee committee readability.

Plain-English risk language with named consequences. Shadow-AI inventory and stop conditions are reported in the same cover as the wins.

Synthesised from NIST AI RMF, Protiviti global board survey, and Deloitte balanced-portfolio framing. Cadence and disclosure quadrants are synthesis.

Five safeguards · adapted from prior field research

Five Doctor of DT safeguards plus an adoption lever — the cards boards typically boards typically forget, adapted for the ROI setting.

Adapted & synthesis Source field: **doctorofdt.com/ai-enabled-framework-research**. Five core safeguards plus an adoption lever that determines whether the value is realised — each card is an adaptation of the underlying ideas to the ROI-case setting.

Safeguard 1

Value-stream mapping beyond beyond cost savings

Map each use case against end-to-end value streams, not tasks. Capacity creation and margin expansion live in the seams between functions; per-FTE savings miss them.

Safeguard 2

Reversibility vs consequence consequence matrix

Sort every initiative by reversibility (can we walk it back?) and consequence (what if it fails?). Reversible, low-consequence moves earn speed; irreversible ones earn stricter governance.

Safeguard 3

Pre-mortems and red-teaming teaming

Before each gate, run a structured pre-mortem and an adversarial review. Boards fund evidence that the team has stress-tested the case, not only defended it.

Safeguard 4

Automated circuit breakers and stop stop conditions

Encode thresholds — accuracy floor, cost ceiling, drift breach, OOD rate — that auto-pause or auto-fallback a workflow before a human catches the failure. Pre-define who has the kill-switch authority.

Safeguard 5

Shadow-AI containment protocol

Untracked AI use is a parallel ledger most CIOs do not see. Inventory it quarterly, route shadow use through the same risk gate as sanctioned use, and decide what to standardise, sandbox, or sunset — before the next audit cycle finds it.

Safeguard 6 · adoption lever

Cognitive ROI & human value

Microsoft finds adoption slows when quality, effort, and creativity are ignored. Pair time-saved with cognitive-load and skill-development metrics — without them, realised ROI reliably under-delivers projected ROI.

Adapted from the AI-Enabled Transformation framework (doctorofdt.com/ai-enabled-framework-research). Pairings of safeguards with board governance draw on NIST AI RMF and Microsoft Worklab value dimensions.

The atomic unit of AI value

Cost per token is necessary; cost per successful outcome is sufficient. sufficient.

Token economics is one row of the cost ledger from slide 6 — but it is a row whose framing quietly determines whether the ROI case ages well or ages badly.

The narrow lens	The board-ready lens
Question	Question
What is our cost per token?	**What is our cost per successful outcome?**
Optimisation	Optimisation
Cheapest model, shortest prompts, aggressive caching.	Right model for the task — accuracy, retry-rate, and human review are part of the unit.
Failure mode	Failure mode
A cheaper model that ships wrong answers and forces downstream rework.	A more expensive model that costs less because it solves the problem the first time.
Multi-agent risk	Multi-agent risk
Hidden — costs multiply across agent calls and reasoning chains.	Made visible — showback, chargeback, and per-task cost ceilings are part of the case.

The metrics that matter

Three numbers FinOps & field research research put on the case.

Cost per inference

Combine model, hosting, and orchestration costs. Show per task and per cohort.

Token yield rate

Successful outputs per million tokens consumed. Each retry and re-prompt erodes it.

Inference efficiency index

Value delivered per dollar of inference — the unit the board actually cares about.

Field references: FinOps "token economics — the atomic unit of AI value" · Nandagopal IJAIS (2025) "Tokens as Currency" · arXiv 2605.09104 · arXiv 2605.17410.

Sourced anchors for the framework

Four sourced reference cases keep the framework close to observable outcomes — not ambitions.

Reference 1 · Google Cloud · value framing

A single chatbot, priced honestly, paid back in under two weeks.

Est. value / mo	Est. cost / mo	Payback
≈ $24,910	≈ $2,700	< 2 weeks

Cloud.google.com /blog/topics/cost-management/measure-the-value-and-impact-of-your-ai

Reference 2 · IBM 2025 C-suite study · scaling reality

A quarter of initiatives deliver on expected ROI — fewer still fewer still scale.

25%
delivered their expected ROI

16%
scaled enterprise-wide

IBM recommends starting with cost-saving use cases, baselining time/cost/quality, and using open orchestration to avoid lock-in.

Reference 3 · Google · survey of 2,500+ execs

Most enterprises now report ROI; the conviction gap is in the in the data behind the case.

74%
report current ROI with gen AI

45%
saw productivity at least double

56%
report improved security posture

Reference 4 · Microsoft · M365 Copilot value framework

Time saved alone is insufficient — six dimensions span value, experience, and risk.

- Revenue impact
- Security & risk
- Quality improvement
- Productivity & efficiency
- Employee / customer experience
- Cost savings

Microsoft Worklab: 79% see AI as crucial; 59% struggle to quantify productivity gains.

Sources per case are linked in the appendix. Use these as anchors — substitute the user's own enterprise benchmarks before the board meeting.

The pre-funding checklist

Ten items the executive team must have on the table before asking for scale funding. funding.

Three columns — what we said, what we proved, what we'll do next. Sponsor signs each row; absent signature, the case does not move to scaling.

What we said we'd do	What we proved	What we'll do next
01	**Rank use case among top value-accretive, AI-tag-indifferent** No "gen-AI premium." Documented ranking against 20–30 candidates; AI tag absent from decision criteria.	
02	**Map value to all three tiers — not labor alone** Hard, soft, strategic; each line owned by a named executive; quantified where data permits, bounded where not.	
03	**Price the full eight-category cost stack** Governance, change, and verification are line items. Stack documented per use case; risk and exit cost separately.	
04	**Define baseline and matched cohort explicitly** Cohort identified; tracking cadence agreed; baseline frozen before pilot; reported as paired deltas each quarter.	
05	**Risk-adjusted return and stop conditions** P25 / P50 / P75 explicit; auto-circuit breakers; named kill authority; thresholds re-tested at every gate.	
		06 **Six KPI layers tracked end-to-end** Adoption → productivity → quality → risk → financial → strategic. Cadence + dashboards live; one owner per layer.
		07 **Governance: NIST controls + named owners** Trustworthy-AI controls embedded; shadow-AI inventory quarterly; tighten controls after each incident.
		08 **Architecture flexibility & lock-in analysis** Optionality priced in; multi-provider or abstraction strategy documented; annual re-platform drill booked.
		09 **Cognitive ROI, change, and adoption evidence** Quality, effort, creativity, eNPS — adoption's leading edges. Adoption depth, drop-out cause, pulse survey.
		10 **Board-grade cadence and disclosure standard** AI on the agenda every quarter; plain-English risk with named consequences; decisions and data archived.

Decision hygiene, not paperwork · sign each row or hold the case

Synthesis · cross-references slide 7 (framework), slide 9 (KPIs), slide 10 (governance), and slide 11 (safeguards).

Activating the playbook

What the next 90 days should look like when the playbook goes from cover slide to working program.

Phase 1 Days 1 – 30 · align

Stakeholder map & sponsor

CFO + CIO + line business sign the executive sponsor. Map legal, security, HR, and data owners.

Rank the top 20–30 candidates

Score on value, feasibility, autonomy envelope. Force-rank without AI-tag bias.

Inventory shadow AI

Survey + telemetry sweep. Decide what to standardise, sandbox, or sunset.

Governance baseline

Map current controls to NIST AI RMF; name owners; agree a stop-condition policy.

Phase 2 Days 31 – 60 · build the case

Top 3 use cases: full ROI pack

Apply the seven-step framework. Build the 8-category cost stack and 3-tier value.

Cohort & baseline freeze

Match each pilot to a non-AI cohort. Freeze baseline; document the do-nothing path.

Measurement architecture live

Wire the six layer KPIs. Pair leading and lagging; auto-collect where possible.

Risk-adjusted scenarios

P25 / P50 / P75 explicit. Name the kill-switch authority and the circuit breakers.

Phase 3 Days 61 – 90 · present & stage-gate

Board pre-read & decision pack

Names, numbers, owners. Decisions required on the front page; risks on the back.

Pilot-gate decision

Continue, expand, pause, or kill. Document the decision and the data behind it.

Scale-funding request

Risk-adjusted ROI + payback + NPV pair. Capex shape clearly named; cash-flow line.

Cadence & escalation

Quarterly board review booked; audit-committee material staged; escalation path live.

Cadence
Quarterly board review · monthly exec stand-up · weekly use-case use-case cadence

Next checkpoint
End of Day 90 · board

Stop condition
Cohort underperforms control → control → revert

Owner of the playbook
Named executive sponsor + cross-functional team

Cadence references — board engagement benchmark (Protiviti), governance baseline (NIST), and stage-gate practice (AWS).

Appendix

Sources cited cited across the deck.

Primary public references that grounded the framework, the cost stack, the value tier model, the governance framing, and the measurement architecture.

Reading order

McKinsey & Microsoft for value framing.
Google Cloud & IBM for measurement and scaling reality.
NIST & Protiviti for governance and cadence.
Doctor of DT and FinOps for safeguards and cost discipline.

URLs live at time of writing.

Sources · 13 references referenced across slides

McKinsey

Economic potential of generative AI

$2.6T–$4.4T annually · ~75% in 4 functions.

mckinsey.com / capabilities / tech-and-ai / our-insights / the-economic-potential-of-generative-ai-the-next-productivity-frontier

McKinsey · CFOs

Gen AI: a guide for CFOs

Rank top 20–30 projects · avoid gen-AI premium.

mckinsey.com / capabilities / strategy-and-corporate-finance / our-insights / gen-ai-a-guide-for-cfos

Google Cloud

How to measure the business value of generative AI

J-curve; learning, verification tax, pipeline adaptation.

cloud.google.com / blog / products / ai-machine-learning / how-to-measure-the-business-value-of-generative-ai

Google Cloud

Measure the value and impact of your AI

≈ $24,910/mo value vs ≈ $2,700/mo cost · payback in < 2 weeks.

cloud.google.com / blog / topics / cost-management / measure-the-value-and-impact-of-your-ai

Google · DORA

Generating value from gen AI — ROI study

2,500+ execs · 74% ROI · 45% 2× productivity · 56% improved security.

cloud.google.com / transform / survey-generating-value-from-generative-ai-roi-study

Microsoft

Measuring the impact of M365 Copilot and AI

Microsoft Digital AI Value Framework · six dimensions.

microsoft.com / insidetrack / blog / measuring-the-impact-of-microsoft-365-copilot-and-ai-at-microsoft

Microsoft Worklab

How we measure the value of AI at work

Quality, effort, creativity · 79% crucial · 59% struggle to quantify.

microsoft.com / en-us / worklab / how-we-measure-the-value-of-ai-at-work

IBM

Realize ROI of AI agents

25% deliver expected ROI · 16% scale enterprise-wide · open orchestration.

ibm.com / think / insights / realize-roi-ai-agents

AWS

Agentic AI economics — measuring success

Baseline, autonomy level, break-even, decision points.

docs.aws.amazon.com / prescriptive-guidance / latest / agentic-ai-economics / measuring-success.html

Deloitte

AI tech investment ROI

Balanced digital portfolio · shared value scorecard.

deloitte.com / us / en / insights / topics / digital-transformation / ai-tech-investment-roi.html

Protiviti

AI in board-meeting discussions — global survey

26% boards every meeting · 63% high-ROI orgs on every agenda.

protiviti.com / us-en / press-release-ai-board-meeting-discussions-global-survey

NIST

AI Risk Management Framework

Trustworthy AI built into design, development, use, and evaluation.

nist.gov / itl / ai-risk-management-framework

FinOps

Token economics — the atomic unit of AI value

Cost per inference · token yield · showback / chargeback.

finops.org / insights / token-economics-the-atomic-unit-of-ai-value

Doctor of DT

AI-Enabled Transformation — research

Adapted for safeguards on slide 11.

doctorofdt.com / ai-enabled-framework-research

Field research · token economics

Nandagopal IJAIS 2025 · Akpan 2024 · arXiv 2605.09104 · arXiv 2605.17410. **Supplements slide 12.**

Substitute your own enterprise benchmarks before the board meeting · preserve audit trail · prefer primary sources.