The Board-Grade AI Business Case: Funding Gates before Scale

Requiring Baseline, Counterfactual, Adoption, Control, Cash and Replication Evidence before Capital Release

Authored by Chennakeshav Adya, Independent Researcher

September 2026

Abstract

Artificial-intelligence proposals can move from demonstration to enterprise funding before their economic and control evidence is mature. A successful prototype may show that a model can produce an output. It may not show that the output improves an authorised operating decision, that users adopt the changed workflow, that quality remains acceptable across representative conditions, or that the improvement converts into cash. Capital released on the strength of a demonstration can therefore fund integration, data, licences, compute and change before the board has evidence of scalable value.

This paper develops a board-grade AI business-case framework. It begins with a controlled baseline and counterfactual, then connects use-case performance to workflow adoption, operating outcomes, control evidence, full lifecycle cost and cash conversion. Six evidence gates govern movement from problem definition to validation, bounded deployment, replication, scale and realised value. Machine assistance can organise evidence and run scenarios, while accountable executives and directors retain approval and risk decisions.

Five original figures and five decision tables present the business-case anatomy, counterfactual design, funding gates, benefit waterfall and board decision dashboard. A worked example uses a hypothetical company and illustrative assumptions. All amounts, percentages, probabilities, scores, weights and scenarios are management assumptions for analytical design. The paper does not provide accounting, audit, legal, regulatory, tax, investment, employment, valuation, security, privacy or technology advice and does not recommend any AI system, vendor or capital commitment.

JEL Classification: G31, M15, O32, D81, C52

Keywords: artificial intelligence, business case, stage gates, benefits realisation, counterfactual, adoption, model governance, capital allocation, transformation, board oversight

1. Begin with the board decision

An AI proposal becomes board-grade when it states the decision requested, the capital at risk, the evidence supporting value, the controls required and the conditions for further funding. A prototype, vendor presentation or technical benchmark can inform the proposal. None of them alone establishes that the company should commit enterprise capital.

The board decision may be to fund problem validation, approve a bounded deployment, release integration capital, authorise scale or stop. Each decision needs a defined evidence threshold. The requested approval should include scope, legal entity, user population, data, vendor, operating owner, control owner, funding, duration and next review date.

The paper should distinguish observed evidence from forecasts and proposed actions. A measured task time is evidence. A forecast reduction in cost is an assumption until the operating mechanism

and cash conversion occur. A planned security control is not an implemented control. Clear classification allows the board to approve conditional learning without treating future benefits as realised.

2. Define the operating problem

The business case should identify a decision, task or process whose current performance creates material cost, delay, risk, revenue loss or working-capital exposure. It should state the affected population, transaction volume, cycle time, quality, exceptions, controls and owner. A broad ambition to use AI is not an operating problem.

Problem materiality needs evidence. Finance can reconcile labour, external spend, rework, claims, inventory, customer loss or delayed cash. Operations can document bottlenecks, variability and failure pathways. Risk and control functions can identify exposure and obligations. The proposal should show which part of the problem is addressable by the proposed system.

Alternative responses should remain visible. Process simplification, policy change, conventional software, analytics, outsourcing, recruitment or no action may be credible routes. The AI case should win through evidence against these alternatives rather than assume technology preference.

The problem statement should identify the decision latency created by the current process. A slow task may be immaterial when it does not delay a customer, payment, control or management action. A relatively short task can be valuable when it sits on a critical path. This distinction prevents the programme from selecting a use case because its activity volume is easy to count. It directs capital toward an observable constraint and gives the committee a measurable reason for intervention.

3. Construct the business-case anatomy

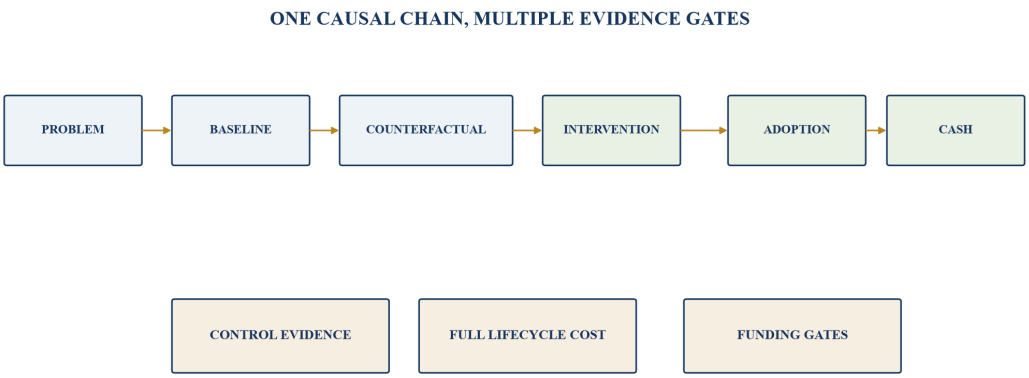
The anatomy connects eight elements: problem, baseline, counterfactual, intervention, adoption, controls, economics and evidence gates. A weakness in any one element can break the value chain. Strong model accuracy cannot compensate for absent data rights. High adoption cannot compensate for a workflow that produces no cash effect. A positive forecast cannot compensate for unacceptable legal or safety risk.

The proposal should show the causal sequence from model output to human or automated action, operating outcome and financial effect. Every link needs an owner and evidence. The model may classify an invoice, a reviewer may accept the classification, a corrected workflow may reduce rejection and the cash date may improve. The result depends on the complete sequence.

The anatomy should also distinguish a platform case from a use-case case. A platform can provide reusable data, security, evaluation and integration services. Its value depends on approved use cases that consume those services. The board can fund a limited platform foundation when it has identifiable demand and an allocation policy. It should avoid attaching every speculative future use case to the initial return. Platform capital and use-case capital can have separate gates and benefit ledgers.

The anatomy also identifies lifecycle obligations. Data preparation, integration, monitoring, incident response, model change, vendor management and retirement continue after launch. Their cost and control requirements belong in the initial business case.

Figure 1. Anatomy of a board-grade AI business case



Capital follows a connected chain from a controlled problem to evidence, authorised operation and realised value.

Table 1. Minimum board-grade AI proposal record

Field	Required content	Evidence	Board use
decision	approval requested, amount, scope and date	delegated authority and funding envelope	authority
problem	process, population, materiality and owner	operating and financial records	relevance
baseline	current outcome, cost, quality and controls	reconciled period and defined population	comparison
counterfactual	expected outcome without the programme	committed plan, trend and external drivers	attribution
intervention	model, workflow, people and system change	design, test and integration evidence	feasibility
risk and control	data, model, human, vendor and legal controls	implemented tests and acceptance	residual exposure
economics	cash pathway, lifecycle cost, range and funding	model with source-linked assumptions	return and liquidity
gates	acceptance, stop, owner and capital release	signed decision record	conditional commitment

Every claim should connect to a source, owner, decision and next evidence requirement.

4. Establish the programme perimeter

The perimeter should identify the legal entities, jurisdictions, users, customers, employees, vendors, models, data sources and systems included. It should also identify excluded populations and interfaces. A narrow prototype can create a misleading impression when scale introduces new data, languages, products, regulations or user behaviour.

The perimeter should match the proposed capital. Funding for one process should not be justified by benefits from an unapproved enterprise rollout. Shared platform costs and reusable capabilities

should be allocated through a transparent policy. The board should see which future use cases depend on the platform and which can operate independently.

Third-party dependencies need named contracts and rights. The proposal should cover data use, model provider, intellectual property, service levels, change, security, audit, portability, termination and transition. Qualified review is required for the actual arrangement.

5. Build a controlled baseline

The baseline is the current operating and financial performance against which change will be evaluated. It needs a defined period, population, source systems, exclusions, adjustments and owner. Measures can include task time, throughput, error, rework, cycle time, conversion, loss, customer outcome and cash.

Finance should reconcile material cost and benefit baselines to management accounts or controlled operational records. Time saved requires a valid measurement method. Error reduction requires stable definitions. Revenue loss requires a customer and contribution-margin pathway. Risk exposure requires an approved method rather than a generic percentage.

The baseline should capture variability and segments. Average performance can conceal high-risk tails or populations where the intervention fails. A representative baseline can support test design and later replication. Weak baselines should lead to measurement work before a large capital release.

Baseline governance needs a cut-off and refresh policy. A rapidly changing operation can make an old period unrepresentative. A very short period can exaggerate seasonality or a temporary backlog. Management should explain why the chosen period reflects the decision population. It should preserve raw extracts, transformation rules and reconciliation so that a later reviewer can reproduce the measure. This evidence is part of the investment asset because it supports both validation and realised-value assessment.

6. Define the counterfactual

The counterfactual estimates what would occur without the proposed programme. It should include committed process improvements, replacement systems, hiring plans, demand, inflation, regulation and normal productivity. Freezing the current process for several years can overstate AI value when other changes would occur.

The company can use matched groups, phased rollout, historical controls or structured forecasts depending on feasibility and ethics. The design should document selection, contamination and external changes. A counterfactual does not need experimental perfection to improve decision quality, but its limitations should be visible.

The board should see the incremental result relative to this path. If the company would have reduced processing cost through ordinary automation, only the additional effect attributable to the AI programme belongs in its case. Counterfactual discipline protects against paying twice for the same improvement.

Counterfactual uncertainty should appear in the value range. When management cannot identify a credible comparison, it can calculate several without-programme pathways and observe whether the decision remains attractive. A project whose return disappears under a modest alternative path

should receive a stronger learning gate. A project whose return remains positive across reasonable counterfactuals can have a more robust economic case, subject to its control and execution evidence.

7. Map the causal value chain

The value chain begins with an input, produces a model output, supports an action, changes an operating outcome and converts to financial value. Each step needs a definition and evidence. A forecast that does not change a decision has limited operating value. A changed decision that does not improve the outcome has no demonstrated benefit.

Human roles should be explicit. The system may recommend, draft, prioritise or automate within approved limits. Reviewers need training, authority, time and escalation. If the control design requires human review, the business case should include its cost and performance. Hidden manual work can erase apparent productivity.

The chain should include failure pathways. Incorrect output, missing data, delayed response, override, gaming, drift or system outage can change the result. The case should show prevention, detection, response and fallback according to consequence.

8. Design the counterfactual test

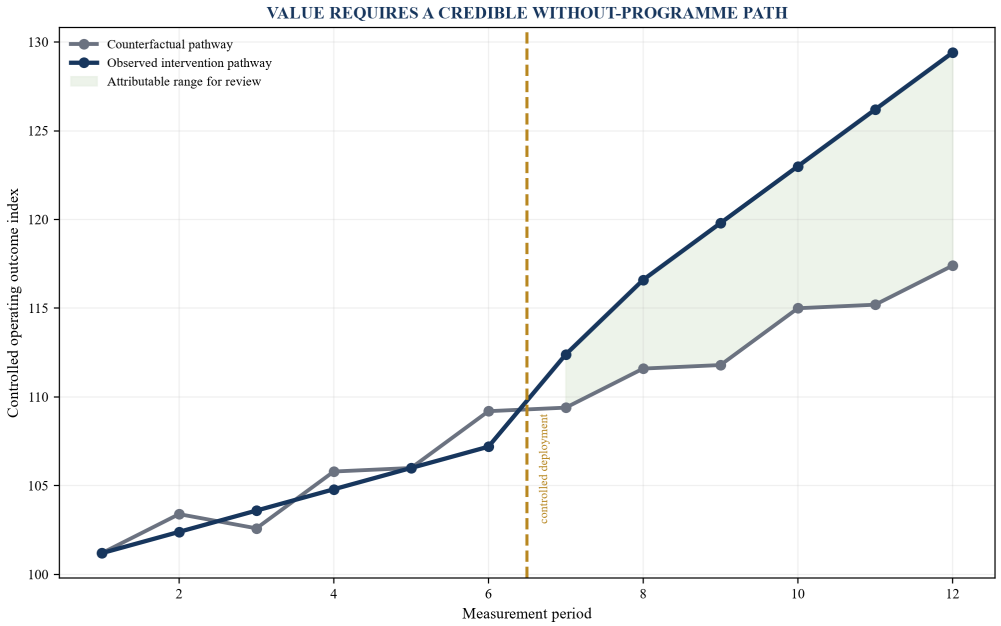
The test should select representative cases, users, periods and operating conditions. It should predefine the intervention, outcome measures, control measures and analysis. Changing the target after results are known reduces the reliability of the evidence. Material deviations should remain in the record.

Where randomisation is feasible and appropriate, it can improve causal evidence. Where it is not, phased rollout, matched comparison, interrupted time series or expert adjudication can be used with limitations. Legal, ethical, customer and workforce considerations determine the suitable design.

The board does not need a single universal method. It needs evidence strong enough for the consequence and capital requested. A low-cost internal aid can use a proportionate test. A system affecting credit, employment, safety, regulated advice or material customer outcomes generally requires stronger validation and oversight.

Test contamination should be monitored. Users in a comparison group can adopt techniques from the intervention group, while intervention users can continue old tools outside the measured workflow. Vendor staff can provide unusually intensive support during a pilot. These effects can make the observed result difficult to reproduce at scale. The test record should identify operational support, deviations and concurrent changes so the committee can judge whether the result represents normal operation.

Figure 2. Counterfactual design from baseline to attributable outcome



The comparison design should isolate programme contribution while documenting selection, contamination and external change.

Table 2. Counterfactual evidence choices

Method	Useful condition	Main limitation	Board question
randomised test	representative units can be assigned appropriately	operational and ethical feasibility	does assignment support causality?
phased rollout	implementation can occur in waves	time and learning can contaminate groups	are waves comparable?
matched comparison	similar non-treated cases exist	unobserved differences remain	what drives selection?
historical baseline	process is stable and records are controlled	market and process change	is history still representative?
expert adjudication	outcomes require professional quality judgment	reviewer consistency and bias	are criteria and blinding controlled?
structured forecast	no suitable control group exists	assumption dependence	which evidence would change the forecast?

The method should be proportionate to decision consequence, feasibility and the capital requested.

9. Set value hypotheses before testing

Each value hypothesis should specify the population, intervention, expected outcome, mechanism, measurement period and minimum effect that matters economically. It should also state the evidence that would disconfirm the claim. A hypothesis such as improving productivity is too broad for capital approval.

The minimum effect should reflect lifecycle cost and opportunity cost. A statistically observable improvement can remain economically immaterial. Conversely, a rare risk reduction can matter even with limited frequency when severity is high and the method is approved.

Multiple hypotheses should be prioritised. The company should avoid adding measures after testing to find a positive result. Exploratory findings can inform a future test and should be identified as exploratory. This preserves a disciplined route from learning to capital.

10. Measure model performance in context

Model performance should match the operating decision. Accuracy, precision, recall, calibration, error distribution, robustness and latency can matter differently by use case. Aggregate performance can hide material failures by segment, language, period or exception. The proposal should define acceptance before deployment.

Performance should be compared with relevant baselines, including current human or rules-based methods. The model should also be tested under representative and adverse conditions. Changes in data, workflow, users or model version can alter results after validation.

Technical performance is one evidence layer. The business case should connect it to decision and outcome performance. A small model improvement can have high value in a constrained process. A strong benchmark can have little value if integration, adoption or control costs dominate.

Threshold selection should reflect the cost of different errors. False approval and false rejection can have different customer, financial and legal consequences. The company can use separate thresholds, human review or abstention for uncertain cases. The proposal should show the volume and cost created by these choices. A high headline accuracy can still produce an unacceptable number of material errors in a large population.

11. Define adoption as changed work

Adoption is demonstrated when the intended users follow the approved workflow and the model output influences the defined action. Login, licence or prompt counts do not establish operating adoption. The business case should measure eligible cases, actual use, override, exception, completion and outcome.

The workflow needs role, authority, training, time and escalation. A reviewer may ignore a tool because it arrives late, lacks source evidence or creates personal risk. A user may accept every recommendation without adequate challenge. Both patterns can damage value. Adoption measures should therefore include appropriate use and control quality.

Change costs include process redesign, communication, training, support, performance management and temporary productivity loss. Complementary investment can determine whether AI produces value. OECD research on firm adoption highlights the importance of digital capability, skills and organisational change, with causal limitations in observed productivity relationships.

Adoption evidence should be segmented by role, business unit, case type and experience. A small group of expert enthusiasts can make an early average appear strong. New users can need different support, and managers can create inconsistent expectations. The programme should track why users accept, override, avoid or escalate. Qualitative evidence can identify design problems while controlled workflow data determine whether the behaviour is widespread and economically relevant.

12. Design human decision rights

The proposal should state which decisions remain human, which can be automated under approved limits and which require escalation. Authority should follow legal, regulatory, policy and risk requirements. A general human-in-the-loop statement does not define accountability.

Reviewers need the information required to challenge an output. Source evidence, confidence, reason, limitations and alternative may be necessary. The interface should avoid creating false certainty. Time allowed for review should match consequence. Overrides should be possible where authorised and recorded with proportionate rationale.

The business case should include reviewer cost, capacity and performance. Automation that moves work into an unbudgeted control team can transfer cost rather than reduce it. Material decisions need qualified governance design.

Decision rights should include the authority to suspend the model. Operations may need immediate fallback during an outage or unexplained performance change. Risk or compliance may need to restrict a population. Technology may need to roll back a version. The governance record should show who can act, how users are informed and how accumulated cases are handled. These operating rights make the control design executable.

13. Map data readiness

Data readiness includes rights, provenance, quality, representativeness, timeliness, security, access and retention. A model can perform well on a curated prototype dataset and fail when source systems, definitions or missingness change. The scale case should use representative production pathways.

The data map should identify source owner, permitted purpose, transformation, quality control, lineage, interface and fallback. Sensitive or regulated data require qualified review. Vendor access and model-provider use should be documented. The company should know whether data are used for training, service improvement or cross-customer purposes.

Data remediation cost belongs in the lifecycle case. If the programme depends on a new master-data, integration or governance capability, the board should see that investment and any reusable value. Platform benefits should not be claimed by multiple use cases without an allocation rule.

Representativeness should be tested against the production population. Historical data can exclude new products, customers or adverse events. Labels can reflect prior policy, bias or inconsistent judgment. The programme should document who created labels, how disagreements were resolved and which cases remain uncertain. A decision to exclude an unsupported population can improve reliability and provide a clear boundary for later evidence work.

14. Establish control evidence

Control evidence should cover governance, legal requirements, privacy, security, model performance, human oversight, third parties, change, monitoring, incident response, fallback and retirement. Policies are inputs. The funding gate needs evidence that relevant controls are designed, implemented and accepted for the proposed scope.

NIST AI RMF provides voluntary govern, map, measure and manage functions. Its core includes defining business value, risk tolerance, roles, monitoring and decommissioning. ISO/IEC 42001 provides an AI management-system standard. These references require adaptation to the organisation and do not replace law or sector-specific obligations.

The residual risk statement should identify exposure after controls, accountable acceptance and conditions. An unresolved critical issue can stop scale regardless of forecast return. The board should avoid a weighted total that hides a mandatory gate.

Control cost and control effectiveness should be evaluated together. A manual review can reduce risk and remove the expected productivity gain. A monitoring rule can produce too many alerts and become ineffective. The business case should test the complete controlled workflow rather than an unconstrained model. Where a control is necessary and expensive, the company can redesign the use case, narrow the scope or recognise a lower return.

15. Govern third-party AI

Third-party systems can create dependencies across data, model, infrastructure, intellectual property, security, service continuity and pricing. The proposal should identify subcontractors, hosting, model versions, data locations, update rights and termination consequences. Procurement price is only one part of the economic case.

Due diligence should be proportionate to use and consequence. Evidence can include architecture, testing, security, incident process, performance, data practice, financial capacity and customer references. Contract terms should reflect actual risk and require qualified legal review.

The company should test portability and fallback. A contractual exit right can have limited value when prompts, workflows, embeddings, integrations or staff capability cannot move. Transition time and cost belong in downside scenarios.

16. Calculate full lifecycle cost

Lifecycle cost includes discovery, data, model, software, integration, infrastructure, compute, licences, security, assurance, legal review, training, change, monitoring, incident response, maintenance, vendor management and retirement. Internal staff time and parallel operation should be included when material.

Costs should reflect scale drivers such as users, transactions, tokens, storage, latency, support and control intensity. Variable cost can change the unit economics of high-volume use. Pricing changes and model migration can affect the forecast. A range is often more informative than one amount.

The case should distinguish committed, avoidable and sunk costs by gate. This supports stop decisions and preserves option value. It also prevents the board from continuing because prior expenditure is large.

Lifecycle estimates should include the cost of evidence itself. Evaluation datasets, independent assurance, legal opinions, control testing and user studies are investments that reduce uncertainty or risk. Their value appears through better decisions and lower exposure rather than direct output. The board can assess whether the evidence cost is proportionate to the capital and consequence at stake. This creates a rational basis for stronger validation in higher-risk cases.

17. Build the cash conversion bridge

The proposal should convert operating outcomes into incremental cash. Time saved becomes value through avoided hiring, reduced external spend, authorised workforce change, additional contribution or redeployed capacity. Error reduction becomes value through less rework, claim, loss or delay. Revenue improvement requires customer, price, volume and margin evidence.

Working-capital effects need timing and sustainability. Risk reduction needs an approved expected-loss or capital method and should not be presented as certain cash. Strategic learning and option value can remain separate from realised cash.

Finance should control definitions and reconcile results where feasible. The bridge should include implementation, tax, working capital, financing and recurring cost. Gross vendor benefits should not substitute for net company cash.

18. Create the benefit waterfall

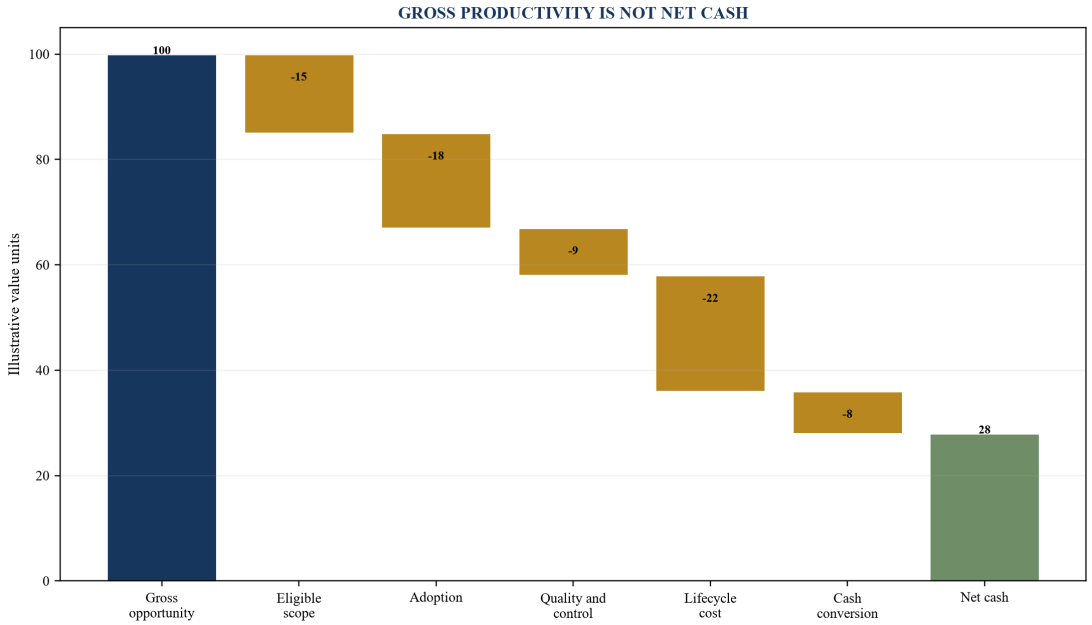
The waterfall begins with gross operational opportunity, then removes ineligible population, adoption loss, quality and control constraints, displacement, implementation, recurring cost and cash-conversion leakage. The result is an attributable net cash range. Each adjustment requires evidence or a labelled management assumption.

The order matters because several deductions can interact. Low adoption can reduce volume and also increase support cost. Quality controls can reduce throughput while protecting loss. The model should avoid applying independent percentages when the drivers are correlated.

The board should see the uncertainty around each layer. A wide range can lead to a smaller bounded deployment or stronger evidence requirement. Precision should not be manufactured through a long spreadsheet.

The waterfall should reconcile to the proposal's cash-flow model. If adoption changes, related licence, support and benefit volumes should change consistently. If a quality control adds reviewers, both cost and risk reduction should move. A linked model reduces the chance that one assumption improves benefit without carrying its economic consequence. Sensitivity analysis should focus on the few drivers that can change the decision.

Figure 3. Benefit waterfall from gross opportunity to attributable cash



Values are hypothetical management assumptions for framework illustration.

Table 3. Benefit recognition controls

Benefit	Operating measure	Cash evidence	Exclusion risk
labour productivity	time, throughput, quality and workload	avoided spend, hiring or monetised capacity	released time without conversion
revenue	conversion, retention, price or volume	realised incremental contribution	gross revenue or market growth
rework	error, exception and resolution cost	reduced internal or external spend	change in case mix
working capital	event timing and balance	attributable sustainable cash release	seasonality or collection action
risk	exposure, probability and severity	approved reduction method	avoided loss presented as certain cash
strategic option	right, evidence and exercise route	changed future decision set	hypothetical opportunity without feasibility

Each benefit requires an attributable mechanism, evidence and a financial treatment.

19. Model uncertainty and downside

The business case should present ranges for adoption, performance, timing, cost, cash conversion and residual risk. A base case alone can conceal the exposure that determines funding design. Scenarios should use coherent combinations rather than changing one variable at a time when drivers are linked.

Downside analysis should show maximum funding need, liquidity effect, operational disruption, control failure, vendor dependency and exit cost. The company should test whether it can stop, recover or operate through failure. Tail exposure can make a small expected gain unsuitable.

The board should identify which uncertainty can be reduced through a test and which risk requires control or transfer. Funding a learning gate differs from accepting an uncontrolled exposure. Every scenario remains a management assumption rather than a forecast.

20. Establish the six funding gates

Gate One confirms a material problem and decision owner. Gate Two confirms baseline, alternatives and counterfactual. Gate Three confirms technical and control validation. Gate Four confirms bounded operating adoption and outcome. Gate Five confirms replication under representative variation. Gate Six authorises scale and realised-value governance.

Each gate should state evidence, acceptance, capital release, owner, authority, stop condition and next date. The sequence can be adapted to risk and use. A low-consequence internal tool may combine gates. A high-consequence system can require independent assurance and regulator engagement.

The ability to stop should be practical. Contracts, architecture and operating plans should preserve it. If most capital is committed before the validation gate, the process provides limited protection.

Gate evidence should have an expiry or review rule. Model performance, vendor terms, regulation, data and operating processes can change between approval and release. A programme that delays execution should confirm that its accepted evidence remains current. The committee can delegate routine confirmation while retaining material exceptions. This keeps conditional capital connected to the facts that supported it.

21. Link capital to gate evidence

The first commitment should buy the cheapest decision-relevant evidence. Problem measurement, data assessment or a controlled test can cost less than enterprise integration. The company should compare the cost of evidence with the expected avoided loss and improved choice.

Later commitments can fund production integration, change, controls and scale after the relevant evidence is accepted. Contingency should be released through defined authority rather than absorbed automatically. Unused capital returns to the portfolio.

The business case should show sunk, committed and conditional amounts at each gate. This gives the board an exposure calendar and makes reversibility visible. A positive full-programme NPV can coexist with a staged route that has better risk-adjusted value.

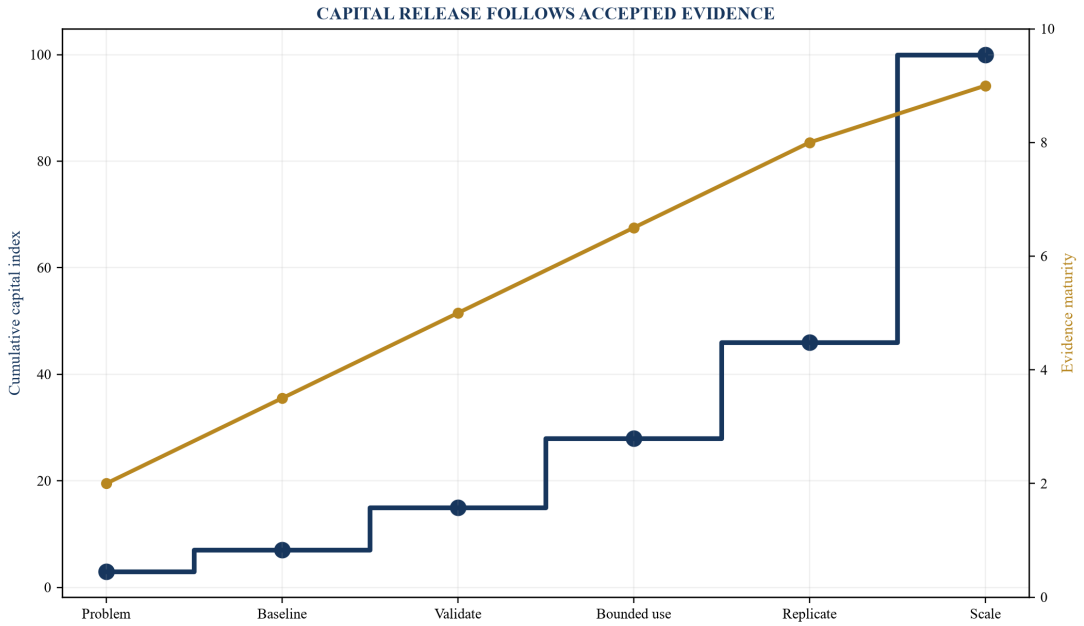
22. Draw the funding-gate map

The gate map connects evidence maturity to cumulative capital. It should show the decision, not only programme progress. A gate can release, hold, narrow, redesign or stop. The map also identifies evidence deadlines and authority.

The capital curve should remain low while the programme resolves the largest uncertainties. Major integration and rollout expenditure should follow representative adoption and control evidence. Some platform investments may need earlier commitment and should disclose the resulting exposure.

The board should compare planned and actual capital by gate. Early overspend can reduce the option to stop. Repeated extensions can indicate that acceptance criteria were weak or that the thesis is deteriorating.

Figure 4. Funding gates before enterprise scale



Capital amounts and evidence scores are hypothetical assumptions; each gate requires approved transaction-specific criteria.

23. Define validation evidence

Validation should cover model performance, data, workflow, human review, controls and operating consequence. It should be independent from model development to the degree required by risk and policy. The report should state purpose, scope, method, data, limitations, findings, residual issues and acceptance.

Federal Reserve SR 11-7 provides supervisory guidance on model risk management in banking, including validation and governance concepts. Its applicability depends on institution and context. A company outside that perimeter can still study the principles without presenting them as mandatory.

Validation should be repeated after material model, data, workflow or use changes. A vendor's general benchmark is not a substitute for company-specific validation. The funding gate should state which unresolved issues block operation or scale.

24. Test bounded deployment

Bounded deployment places the system in a controlled operating scope with representative users and real workflow while limiting exposure. Boundaries can include value, customer, geography, product, decision type, user, volume and duration. The fallback process should remain available.

The programme should measure appropriate use, quality, exceptions, incidents, cycle time, operating outcome and cash pathway. Users should report friction and workarounds. Control teams should test monitoring and escalation. Management should avoid expanding scope during the test without recording the change.

The result can support scale, redesign or stop. A technically successful test can still fail the adoption or cash gate. A commercially promising test can still fail a critical control gate.

Bounded deployment should include incident and near-miss capture. A small population may not produce a statistically frequent event, but a credible failure pathway can still be material. Users need a simple reporting route and protection from incentives to hide exceptions. The committee should review the severity, detectability and control response, not only incident count. This provides richer evidence for scale design.

25. Require replication before scale

Replication asks whether the result survives new users, periods, cases, locations, data and operating conditions. Enterprise value depends on generalisation beyond the original team and curated sample. The replication plan should identify the variation that scale introduces.

The company should predefine acceptable degradation, new controls and support requirements. Replication can reveal that performance depends on one expert, clean data or unusual process discipline. These dependencies can be addressed or used to narrow the scope.

Replication should also test economics. A new business unit can have lower transaction volume, different labour cost or higher integration expense. The same performance may therefore produce a weaker return. The business case should update unit economics by population and avoid assuming that technical replication guarantees financial replication. Scale can be selective where only some populations meet both value and control thresholds.

Scale capital should be released only for the population supported by evidence. Additional geographies, languages or regulated contexts can receive their own gate. This protects value and avoids representing one successful use case as an enterprise capability.

26. Preserve architecture and vendor options

The business case should compare build, buy, configure, partner and hybrid routes. Each route changes speed, cost, control, capability, intellectual property and dependency. The company can use stage gates to delay irreversible architecture choices until evidence improves.

Modularity can preserve the ability to change models, providers or workflows. It can also add integration cost. The board should understand which components are portable and which are proprietary. Data, evaluation sets, prompts, interfaces and operating knowledge can be strategic assets when governed properly.

Vendor selection should follow the approved use and evidence requirement. A general enterprise contract should not automatically approve every use case. Material changes in model, terms, data handling or price should trigger review according to consequence.

27. Integrate legal and regulatory review

The legal and regulatory perimeter depends on jurisdiction, sector, users, data, decision and consequence. The proposal should identify applicable requirements and qualified owners. It should avoid generic claims that a tool is compliant without a fact-specific analysis.

CBUAE has published guidance concerning consumer protection and responsible adoption and use of AI and machine learning within its perimeter. DFSA has reported growing generative-AI adoption and developing governance within the DIFC. Organisations should consult the current official materials and obtain qualified advice for their facts.

Regulatory evidence can change product scope, human review, records, explanation, validation and monitoring. These requirements are economic inputs. A delayed approval or added control can change NPV and timing.

28. Account for workforce consequences

AI can change task design, skills, workload, supervision and accountability. The business case should identify affected roles, current work, future work, training, transition and employee-relations requirements. Workforce action needs qualified review and authorised management decisions.

Released time is not automatically a saving. Management should specify whether it supports capacity, service, quality, avoided hiring, reduced external spend or an approved role change. The cash case should reflect the selected mechanism and timing.

Workforce evidence also affects adoption. Employees can identify process exceptions and control needs that a technical team misses. Participation should be structured, documented and consistent with applicable law and policy.

29. Govern security and resilience

Security design should cover data, identities, access, model and prompt attacks, software supply chain, logging, incidents and recovery. The proposal should identify the business consequence of compromise and the fallback process. Controls should be tested before the relevant operating gate.

The UK National Cyber Security Centre and international partners have published secure AI system development guidance. NIST has also published cybersecurity and AI risk resources. These materials can inform the programme when adapted to architecture and obligations.

Resilience includes vendor outage, model withdrawal, degraded performance and loss of connectivity. The cost of redundancy, manual fallback and recovery belongs in the business case. A system that creates critical dependency can require a different risk appetite and return threshold.

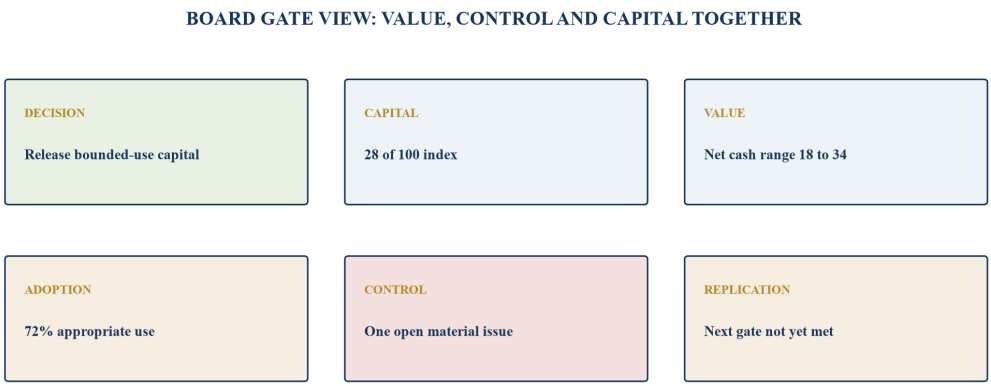
30. Build the board decision dashboard

The dashboard should show decision requested, cumulative and conditional capital, baseline, counterfactual, performance, adoption, control status, net cash range, residual risk, gate result and next deadline. It should identify source dates and accountable owners. A single green status can conceal a failed value or control pathway.

The board should see both level and trend. Model performance can remain above threshold while adoption falls. Adoption can rise while cash conversion remains absent. Control incidents can change the acceptable scope. The dashboard connects these signals.

Exceptions should link to decisions. Each red or amber item should state exposure, action, owner, deadline and authority. Narrative should explain material changes since the previous gate rather than repeat activity.

Figure 5. Board decision dashboard for an AI funding gate



Scores and values are hypothetical management assumptions; critical gates remain visible outside the aggregate view.

Table 4. Board funding-gate dashboard

Dimension	Board measure	Gate evidence	Decision consequence
capital	spent, committed, conditional and reserve	reconciled funding and contracts	release or preserve
value	attributable net cash range and option value	baseline, counterfactual and bridge	continue, redesign or stop
performance	task and outcome measures by segment	validated representative results	approve scope
adoption	eligible, appropriate use, override and exception	production workflow records	scale change programme
control	critical findings, incidents and residual risk	accepted control evidence	block or condition
replication	result across users, time, data and contexts	predefined replication test	authorise population

Every status should connect to source evidence, an owner and a decision consequence.

31. Define stop and redesign conditions

Stop conditions can include failure to establish materiality, unacceptable performance, absent rights, critical control failure, adoption below threshold, cash conversion outside range, cost above limit or replication failure. The conditions should be observable and linked to authority.

Redesign can be appropriate when the problem remains material and the evidence identifies a fixable pathway. Scope can narrow, workflow can change, data can improve or a vendor can be replaced. The committee should approve revised cost, timing and evidence rather than roll the programme forward informally.

Restart conditions should also be defined. A stopped programme may resume after new evidence or remediation. The record should preserve prior findings so that the same weak thesis is not repeatedly funded under a new label.

32. Prevent benefit inflation

Benefit inflation can arise from broad population, frozen baseline, absent counterfactual, gross revenue, released time, double counting, omitted cost or unsupported replication. The business case should contain explicit tests for each risk. Finance should own the benefit policy and reconciliation.

Sponsors can be rewarded for learning and disciplined stops, not only scale. Otherwise the governance system can encourage optimistic assumptions and continuation. Decision quality should include capital preserved through early disconfirmation.

Independent challenge can test evidence, method and incentives. Material overrides should be documented with reason and owner. Historical bias by benefit type and sponsor can improve future contingencies without assuming misconduct.

33. Use machine assistance within the business case

Machine assistance can index source documents, extract assumptions, detect inconsistent values, generate sensitivity cases and monitor evidence. It should maintain links to source and expose confidence. Generated claims require verification before a material decision.

The tool used to prepare a business case is itself an AI use case when relevant. Its data, security, model, human-review and record controls should match consequence. Confidential transaction, customer and workforce information requires approved handling.

The board should receive the accountable management recommendation, not an opaque model score. A model can support calculation and challenge. It cannot assume fiduciary, statutory or delegated responsibility.

34. Measure realised value after approval

Realised-value reporting should begin at bounded deployment and continue after scale. It should preserve the approved case, current forecast, counterfactual and actual result. Variances should be attributed to scope, timing, performance, adoption, cost, market and management action.

Finance can reconcile material cash effects to management accounts and bank or ledger evidence where feasible. Operating owners confirm causation and continuing performance. Control owners report incidents and residual risk. The combined view protects against reporting financial benefit while quality or risk deteriorates.

Benefits should have sunset and review rules. A one-off saving should not become recurring run-rate. A baseline can change after reorganisation. The governance method should remain consistent enough to support learning.

35. Create the benefit ledger

The benefit ledger records each mechanism, baseline, counterfactual, formula, owner, evidence, cost, timing, status and approval history. It separates gross opportunity, validated operating outcome, recognised run-rate and realised cash. This prevents different teams from using the same measure with different meanings.

Negative benefits and disbenefits belong in the same ledger. Customer loss, quality degradation, staff burden, incident cost and vendor lock-in can offset value. The board should see net results.

Changes require version control. When the sponsor revises scope or assumption, the ledger should show the bridge from the approved case. This makes realised-value review constructive and evidence-led.

Table 5. Benefit ledger fields

Field	Definition	Owner	Evidence
mechanism	operating change that creates or removes value	business owner	process and decision map
baseline and counterfactual	controlled current and without-programme paths	finance and operations	reconciled records and method
gross outcome	measured operating effect before conversion	operations	representative outcome data
deductions	adoption, quality, displacement and cost	finance and control owners	accepted calculations
net cash	attributable realised or forecast cash	finance	accounts, contracts and cash evidence
residual risk	exposure after implemented controls	risk owner	control testing and acceptance
approval history	gate, conditions, changes and authority	programme office	signed decision record

The ledger provides a traceable bridge from proposal to operating evidence and realised cash.

36. Run a hypothetical worked example

Consider a hypothetical company evaluating an AI-assisted contract-review workflow. It processes 60,000 documents a year. The controlled baseline is an illustrative 42 minutes per eligible document, 6 per cent material rework and USD 5.8 million annual internal and external processing cost. These figures are assumptions and do not describe a company.

The proposal forecasts an illustrative 35 per cent reduction in handling time and 30 per cent reduction in rework. Full three-year lifecycle cost is assumed at USD 4.6 million, including data, integration, licences, assurance, training, monitoring and retirement. The initial gross benefit estimate is USD 8.2 million over three years before adoption, quality, displacement and cash conversion.

The committee does not release full scale capital. It approves USD 350,000 for baseline confirmation, data rights, workflow design and validation. The gate requires representative documents, predefined quality, reviewer evidence and security acceptance.

37. Apply bounded deployment to the example

The hypothetical validation shows acceptable performance on routine agreements and material weakness on non-standard liability clauses. The scope is narrowed. A bounded deployment covers defined low-risk documents with mandatory review and escalation. The committee releases a further illustrative USD 900,000 for integration and controlled operation.

After three months, appropriate adoption is assumed at 74 per cent. Handling time falls by an illustrative 24 per cent within scope, while reviewer exception work is higher than forecast. No customer or legal outcome is assumed. Finance reduces the net benefit range and the programme funds interface and training improvements.

This evidence supports continued bounded use. It does not support enterprise scale. The non-standard document population remains outside the approved scope. The decision record preserves the limits.

38. Test replication and cash conversion

The hypothetical replication adds two business units, a new document period and different reviewers. Performance remains within the predefined range for the bounded document class. Adoption rises to an illustrative 83 per cent after workflow changes. External-review spend declines, while internal capacity is partly redeployed.

Finance recognises only the reduced external spend and an approved avoided-hiring amount in the realised cash bridge. Released time that has not converted to cost or revenue remains a capacity benefit. This produces an illustrative three-year net cash range of USD 2.1 million to USD 3.4 million after lifecycle cost.

The committee can now decide whether this return, residual risk and strategic capability justify further scale. All values remain analytical assumptions. The example demonstrates method rather than a forecast or recommendation.

39. Implement the framework in ninety days

Days One to Thirty can define governance, proposal template, benefit policy, evidence taxonomy, gate authority and AI inventory. Finance and operations can select two live cases and establish controlled baselines. Risk, legal, privacy, security and technology owners can define critical gates within their mandates.

Days Thirty-One to Sixty can rebuild the cases using counterfactual, lifecycle cost, adoption, control and cash pathways. The company can test the benefit ledger and dashboard. Any machine assistance used for evidence management should receive approved data and review controls.

Days Sixty-One to Ninety can run the first gate committee, record decisions, release conditional capital and establish realised-value reporting. The board should review exceptions and method limitations. The next cycle should use observed accuracy, adoption and conversion evidence to calibrate assumptions.

40. Make evidence the unit of scale

An AI demonstration answers whether a system can produce an output under selected conditions. A board-grade business case answers whether a defined operating change can create attributable value within acceptable risk, cost and control. The distinction determines capital discipline.

Baseline and counterfactual establish the comparison. Adoption and workflow evidence establish operational use. Control evidence establishes the permitted scope. Lifecycle economics and cash conversion establish financial value. Replication establishes the population that can support scale.

Funding gates allow the company to buy evidence before it buys enterprise exposure. They preserve the ability to narrow, redesign or stop. When realised outcomes update future assumptions, the organisation develops a repeatable AI investment capability governed by accountable directors and executives.

References

[1] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework 1.0, 2023. https://doi.org/10.6028/NIST.AI.100-1

[2] National Institute of Standards and Technology, AI RMF Core. https://airc.nist.gov/airmf-resources/airmf/5-sec-core/

[3] National Institute of Standards and Technology, AI RMF Playbook. https://airc.nist.gov/airmf-resources/playbook/

[4] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, 2024. https://doi.org/10.6028/NIST.AI.600-1

[5] Central Bank of the UAE, Guidance Note on Consumer Protection and Responsible Adoption and Use of Artificial Intelligence and Machine Learning, 2026. https://rulebook.centralbank.ae/en/rulebook/guidance-note-consumer-protection-and-responsible-adoption-and-use-artificial-intelligence

[6] Dubai Financial Services Authority, AI Survey 2025: Generative AI Adoption and Governance in the DIFC, 2025. https://www.dfsa.ae/news/new-dfsa-ai-survey-generative-ai-adoption-has-nearly-tripled-within-difc-last-12-months-governance-continues-develop

[7] OECD, AI Adoption by Small and Medium-Sized Enterprises, 2025. https://www.oecd.org/en/publications/ai-adoption-by-small-and-medium-sized-enterprises_426399c1-en.html

[8] OECD, The Adoption of Artificial Intelligence in Firms, 2025. https://www.oecd.org/en/publications/the-adoption-of-artificial-intelligence-in-firms_f9ef33c3-en.html

[9] OECD, A Portrait of AI Adopters across Countries, 2023. https://doi.org/10.1787/0fb79bb9-en

[10] UK Government, Digital and Data Benefits Framework, 2026. https://www.gov.uk/government/publications/digital-and-data-benefits-framework/digital-and-data-benefits-framework

[11] HM Treasury, The Green Book: Central Government Guidance on Appraisal and Evaluation. https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government

[12] UK Government, Managing Your Artificial Intelligence Project. https://www.gov.uk/guidance/managing-your-artificial-intelligence-project

[13] Financial Stability Board, The Financial Stability Implications of Artificial Intelligence, 2024. https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/

[14] International Organization for Standardization, ISO/IEC 42001 Artificial Intelligence Management System. https://www.iso.org/standard/81230.html

[15] International Organization for Standardization, ISO 31000 Risk Management. https://www.iso.org/iso-31000-risk-management.html

[16] US Federal Reserve and Office of the Comptroller of the Currency, Supervisory Guidance on Model Risk Management SR 11-7 and OCC 2011-12. https://www.federalreserve.gov/supervisionreg/srletters/sr1107.htm

[17] UK National Cyber Security Centre, Guidelines for Secure AI System Development, 2023. https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development

[18] National Institute of Standards and Technology, Cybersecurity Framework 2.0, 2024. https://doi.org/10.6028/NIST.CSWP.29

[19] IFRS Foundation, IAS 38 Intangible Assets. https://www.ifrs.org/issued-standards/list-of-standards/ias-38-intangible-assets/

[20] IFRS Foundation, IAS 36 Impairment of Assets. https://www.ifrs.org/issued-standards/list-of-standards/ias-36-impairment-of-assets/

[21] IFRS Foundation, IAS 7 Statement of Cash Flows. https://www.ifrs.org/issued-standards/list-of-standards/ias-7-statement-of-cash-flows/

[22] Committee of Sponsoring Organizations of the Treadway Commission, Enterprise Risk Management Framework. https://www.coso.org/enterprise-risk-management

[23] OECD, OECD AI Principles. https://oecd.ai/en/ai-principles

[24] European Commission, Ethics Guidelines for Trustworthy AI, 2019. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

[25] UK Government, AI Playbook for the UK Government, 2025. https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government

About the Author

Authored by Chennakeshav Adya, Independent Researcher

Chennakeshav (CK) is a corporate finance and investment banking executive with 25+ years of global experience in deal origination, structuring and execution across M&A, growth capital and corporate strategy. He has led value-creation mandates for founders, corporates and funds — bridging the boardroom view to hands-on execution and close.

His career spans Morgan Stanley, HSBC, Lloyds Banking Group, EWEC, ADQ portfolio companies and Emirates Growth Fund, across TMT, real estate, fintech, deeptech, cleantech, infrastructure and energy. He has partnered with C-suite leaders, private equity and venture funds, sovereign wealth funds and family offices to finance complex fund raises and scale-up ventures, and has led M&A due diligence, post-merger integration and business-transformation initiatives to create value.

At Matchpoint Partners he is Managing Partner, leading the firm's corporate finance, M&A and capital-raising practice. He holds an MBA from London Business School, an engineering degree from VTU and a Master of Laws (LLM, in progress) from UCL London.

An active start-up mentor, CK mentors at Techstars, DIFC FinTech Hive, Startup Grind, Founder Institute and IN5, serves as Entrepreneur Mentor in Residence (EMiR) at London Business School, and judges the Entrepreneurship World Cup.