

The AI Value-Capture Contract: Aligning Vendors, Employees and Customers to Measured Outcomes

A commercial framework for baselines, counterfactuals, data rights, service levels, gainshare, liability, audit and dispute resolution

Authored by Chennakeshav Adya

Independent Researcher

14 September 2026

Abstract

Artificial-intelligence contracts often define access, implementation and subscription charges more clearly than the economic result that justifies the investment. A buyer may expect lower cost, faster cycle times or higher revenue. A vendor may claim influence over those outcomes. Employees and customers may supply the actions and data that make the result possible. Unless the parties agree the value event, baseline, counterfactual, evidence and allocation rules before deployment, a successful technical implementation can still produce a commercial dispute.

This paper presents an AI Value-Capture Contract framework. It converts an AI business case into a governed measurement system and a set of contractual schedules. The framework defines the outcome perimeter, baseline tree, comparison method, data rights, service levels, human responsibilities, gainshare waterfall, liability allocation, change control, audit trail, dispute ladder and exit conditions. It also separates operational benefit from accounting recognition and from marketing claims.

A wholly hypothetical illustration uses a USD 20.0 million annual addressable cost base. Reported gross benefit of USD 4.20 million is reduced for external effects, unsupported attribution, quality failures and incremental operating costs. The resulting verified net benefit is USD 2.38 million. Under the illustrative waterfall, the vendor receives USD 0.43 million of outcome consideration and the customer retains USD 1.95 million. Sensitivities show how the allocation changes when the counterfactual, evidence quality or service performance changes.

Every amount, rate, score and outcome in the illustration is hypothetical. The framework is a decision aid, not a contract form or an estimate for any organisation. A live arrangement requires verified process data, customer and employee impacts, system performance, intellectual-property and data rights, applicable law, tax and accounting treatment, and advice from qualified legal, privacy, technical, employment, accounting and regulatory specialists.

JEL Classification: D23, D86, L14, M15, O32

Keywords: artificial intelligence, outcome-based contracting, gainshare, value measurement, counterfactual, data rights, service levels, audit, vendor governance

1. Define the commercial decision

The first decision is whether an AI initiative should be purchased as capacity, as a deliverable, as a measured outcome or through a hybrid of those structures. Capacity pricing pays for access to software, models, infrastructure or people. Deliverable pricing pays for agreed implementation outputs. Outcome

pricing links part of consideration to a measured change in the customer's business. A hybrid can fund unavoidable delivery cost while reserving variable consideration for independently supported value.

The commercial choice should follow the controllability of the result. An outcome fee is most defensible when the value event is observable, the baseline is stable enough to measure, the vendor can materially influence performance, and the customer can perform the complementary actions required. A fixed fee is usually easier where outcomes depend heavily on macroeconomic conditions, customer discretion, multiple vendors or long and uncertain causal chains.

The contracting team should write the decision in one sentence: the organisation is considering a defined AI-enabled change, within a defined perimeter, to improve a defined outcome over a defined measurement period, subject to named safeguards and a maximum authorised economic exposure. That sentence becomes the controlling scope for the business case, statement of work and measurement schedule.

NIST organises AI risk management around govern, map, measure and manage. It states that measurement outcomes should inform continuing risk monitoring and response. [1] This supports a contract architecture in which commercial payment depends on evidence produced by an operating control system rather than a retrospective negotiation.

The investment committee should approve the target outcome, fixed and variable consideration, measurement method, risk tolerance, data use and stop conditions. Legal terms should follow that approved economic design. If the measurement design remains unresolved, the parties should avoid presenting the variable fee as objectively determined.

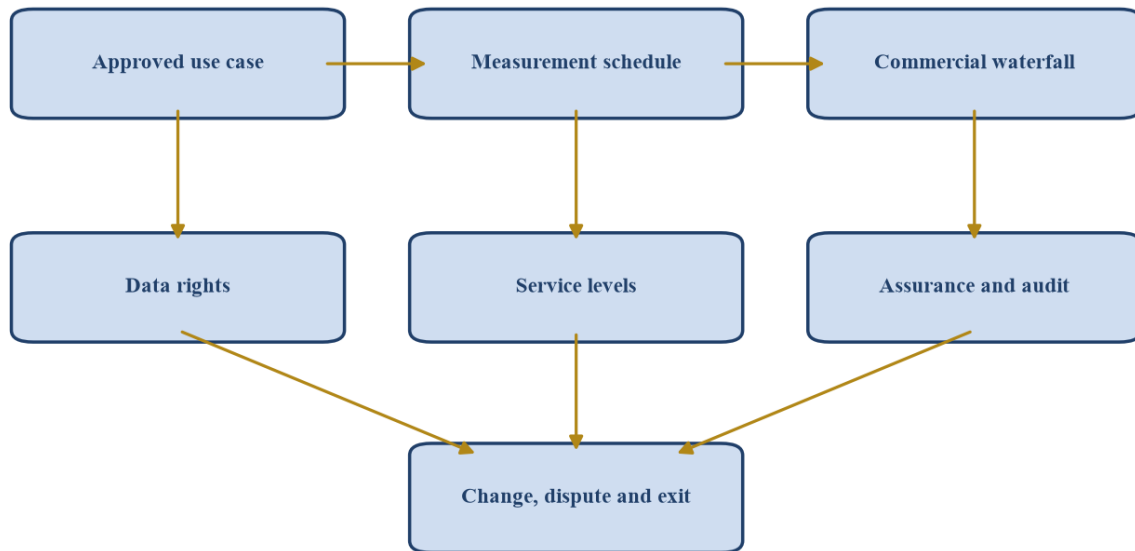
2. Build the value-contract architecture

The contract should be assembled as a set of linked schedules. The master agreement addresses general legal and commercial terms. The use-case schedule identifies the workflow and approved purpose. The measurement schedule defines the value event, baseline, counterfactual and calculation. The data schedule governs access, processing, retention and deletion. The service schedule sets availability, performance and support. The responsibility schedule records vendor, customer and employee actions. The assurance schedule defines testing, evidence, audit and incident reporting. The commercial schedule applies the gainshare waterfall and caps.

Each schedule should use the same defined terms and version identifiers. A model name alone is insufficient because the delivered service can also depend on retrieval data, prompts, policies, software orchestration, human review, infrastructure and downstream decisions. The contracted system should be identified as a versioned configuration with an owner and change history.

The architecture should distinguish obligations from dependencies. A vendor may be obliged to maintain response latency and model availability. The vendor may depend on the customer to provide timely, lawful and representative data. The customer may be obliged to train authorised users and operate a human-review control. The measured outcome may depend on customer adoption, process redesign and management decisions. Recording these dependencies makes attribution and remedies more credible.

Figure 1. Proposed architecture of an AI value-capture contract
From approved outcome to governed payment



Original framework. The legal structure, terminology and required schedules should be adapted to the transaction and applicable law.

Table 1. Proposed contractual schedule map

Schedule	Core question	Minimum evidence	Primary approval
Use case	What system may do what, for whom and where	Process map, intended users, prohibited uses	Business and risk owner
Measurement	What change counts as value	Metric definition, baseline, comparison method	Finance and operations
Data	What information may be used	Data map, lawful basis, permissions, retention	Data owner and privacy lead
Service	What performance must be delivered	Availability, latency, accuracy, support records	Service owner
Assurance	How claims are tested and challenged	Test results, logs, model and control versions	Independent assurance owner
Commercial	How verified value is allocated	Waterfall, caps, exclusions, invoices	Finance and authorised signatories

Original framework. Qualified counsel should determine the enforceable document structure.

3. Define the value event

A value event is the smallest observable occurrence that connects system use to an economic outcome. Examples include a correctly resolved customer request, a prevented duplicate payment, an accepted engineering design, an eligible lead converted into a sale, a shipment cleared without avoidable delay or a claim processed within the approved standard. Broad descriptions such as productivity improvement or better decisions are not sufficiently precise for payment.

The event definition should identify the unit, starting point, completion point, value date, data source, acceptance rule and reversal rule. If a customer case is reopened within thirty days, the original resolution may need to be reversed. If a predicted maintenance event is followed by an inspection showing no defect,

the avoided-cost claim requires a pre-agreed treatment. The contract should state how duplicate, incomplete, cancelled and exceptional events are handled.

Economic value may sit several steps beyond the operational event. A faster proposal does not automatically create revenue. A lower forecast error does not automatically reduce inventory. A successful recommendation does not automatically change a customer's behaviour. The measurement schedule should draw a causal chain from system output to employee action, operating event, financial effect and cash consequence. Payment can be attached to the point where evidence and controllability are strong enough.

The event should also have a quality boundary. A reduction in handling time may destroy value if errors, complaints or rework increase. The primary metric should therefore be paired with guardrail metrics such as accuracy, safety, customer outcome, employee workload, privacy incidents and downstream loss. A value event fails when a material guardrail breach occurs, even if the headline metric improves.

4. Construct the baseline tree

The baseline is the expected result without the contracted intervention. It should be defined before implementation using a hierarchy of evidence. The highest-quality available method may be a contemporaneous control group or randomised rollout. Other methods include phased introduction, matched units, difference-in-differences, a forecast built from pre-period drivers, or a simple historical average adjusted for agreed changes.

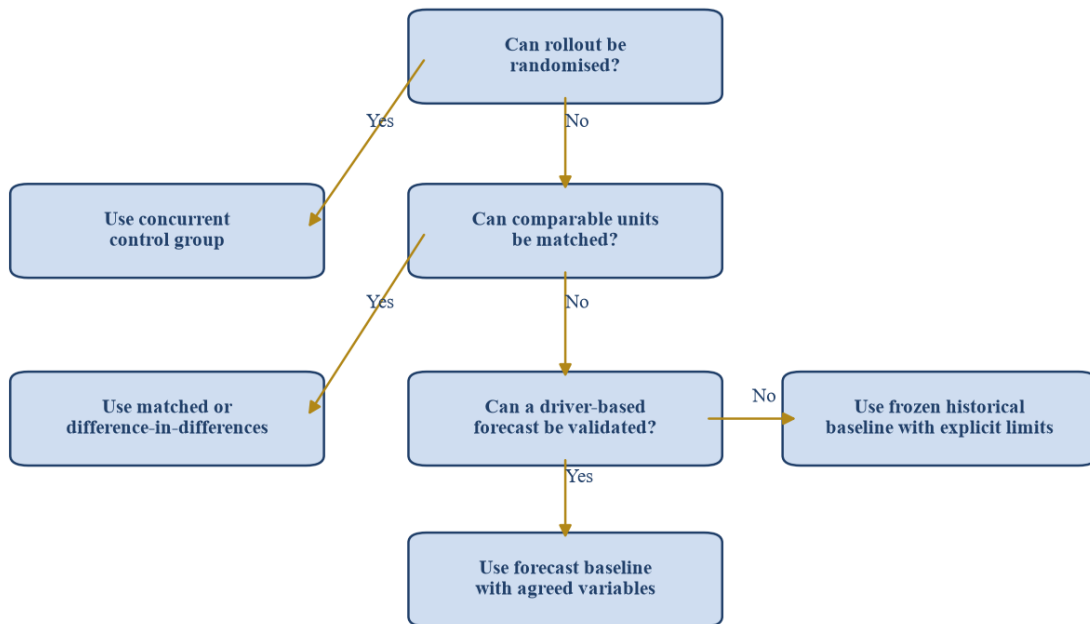
World Bank guidance explains the counterfactual problem as the need to estimate what would have happened to the same unit without the intervention. It describes experimental and quasi-experimental methods for causal inference. [5] Commercial contracts rarely reproduce a full research design, but the same discipline helps prevent all post-deployment improvement from being credited to the AI system.

The baseline tree should specify the preferred method and fallbacks. It should identify the treatment population, comparison population, pre-period, measurement period, seasonality, material external variables, exclusions and minimum sample. Each branch should state who can activate it and how the effect on payment is calculated.

Historical baselines need change rules. Volume, product mix, wage rates, input prices, regulation, customer behaviour, acquisitions, closures and other programmes can alter the expected result. The parties should agree which variables are normalised and which remain part of ordinary business risk. Excessive normalisation creates a model that cannot be audited. Insufficient normalisation can reward the vendor for unrelated improvement or penalise it for an external shock.

Figure 2. Proposed baseline and counterfactual decision tree

Choose the strongest feasible comparison before deployment



Original framework. Method selection depends on data, operational feasibility, ethics and the use case.

5. Freeze metric definitions and source lineage

Every payment metric should have a data dictionary. It should state the business definition, formula, system of record, extraction method, owner, refresh frequency, permitted adjustments, rounding and treatment of missing or late data. The definition should also record whether the metric represents an operational event, accounting amount or cash receipt.

Lineage matters because an AI service can alter the data used to measure itself. A system that classifies cases as resolved may also write the resolution status on which its performance fee depends. The measurement design should retain independent evidence, such as customer confirmation, downstream completion, accounting records or a separate control system. Self-generated evidence can support the calculation, but it should not be the sole determinant where manipulation or drift is plausible.

The parties should freeze data and calculation versions for each measurement period. Corrections should remain possible through a controlled restatement process. Logs should identify the model, prompt or policy version, input data set, human reviewer, output, final action and subsequent economic event. Access should follow privacy, confidentiality, security and privilege requirements.

The GAO AI accountability framework groups practices around governance, data, performance and monitoring, and supplies questions for management and assessors. [8] A contractual evidence pack can use the same four-part logic: authorised purpose, reliable data, measured performance and continuing monitoring.

6. Allocate data rights and responsibilities

The data schedule should distinguish customer data, vendor data, licensed third-party data, employee data, customer-generated feedback, system logs, prompts, outputs, derived features and model improvements. For each class, define ownership or permitted use, access, security, location, retention, deletion, portability, audit and use after termination.

Controller and processor labels should reflect actual decisions about purposes and means. The UK Information Commissioner's Office states that organisations determining the purposes and means of personal-data processing are controllers regardless of the label used in a contract. It also recommends documenting roles across the AI supply chain and setting written instructions, security measures, sub-processor controls, end-of-contract treatment and audit rights. [3] Applicable requirements vary by jurisdiction and activity.

Training and improvement rights should be explicit. Permission to process data for delivering the service does not automatically establish permission to use the same data for a vendor's separate model development. The schedule should identify whether information can be aggregated, de-identified or used to improve a shared model, and should define the evidence required for the chosen treatment.

Measurement rights must survive long enough to settle the final fee and audit disputes. That period should be limited and proportionate. A data-retention clause designed around operating service termination can inadvertently delete evidence needed for payment or regulatory response. The parties should reconcile deletion duties with legal holds, audit windows and final settlement.

7. Connect service levels to value

Conventional service levels measure uptime, response time and support. AI value contracts also need performance and control levels linked to the use case. Depending on the system, these may cover precision, recall, error cost, abstention rate, escalation quality, latency, throughput, robustness, security, explainability, data freshness and human-review compliance.

A model-level metric should not substitute for an end-to-end service result. High benchmark accuracy can coexist with poor business performance if inputs are stale, users ignore outputs, integration fails or customer conditions change. The service schedule should identify which metrics are vendor controlled, customer controlled or jointly controlled.

NIST's Generative AI Profile focuses on governance, content provenance, pre-deployment testing and incident disclosure. [2] These disciplines can be translated into acceptance tests, provenance logs, release gates and incident obligations. The contract should identify material model or system changes that require retesting before the revised version can contribute to an outcome fee.

Service failure should affect payment through an agreed mechanism. Possible treatments include exclusion of affected events, a quality deduction, service credit, gainshare reduction, cure period or termination right. The same failure should not create unintended duplicate recovery unless the contract expressly provides it. Liability, indemnity and service-credit provisions should be reconciled with the commercial waterfall.

Table 2. Proposed outcome metric and guardrail record

Metric class	Example definition	Evidence	Contract effect
Value event	Eligible case completed and accepted	Workflow and downstream system	Enters gross-benefit calculation
Quality	Reversal or substantiated rework within defined window	Case audit and customer record	Reverses event or applies deduction
Safety	Material harm or prohibited action	Incident and review record	Excludes period and triggers escalation
Service	Availability, latency and capacity	Independent monitoring	Service credit or fee reduction

Metric class	Example definition	Evidence	Contract effect
Adoption	Authorised users following approved process	Access and workflow logs	Tests customer dependency
Control	Required human review completed	Review record and sample testing	Conditions eligibility for payment

Original framework. Thresholds should be supported by use-case evidence and approved risk tolerance.

8. Define human and organisational dependencies

AI outcomes usually require changes in work. Employees may need to review outputs, correct data, redesign queues, change customer conversations or accept new performance measures. Managers may need to remove duplicate controls, revise decision rights and fund integration. These actions should be identified as customer dependencies rather than assumed.

The responsibility schedule should name the accountable role, due date, evidence and consequence for each dependency. Consequences should be proportionate. A missed training deadline might extend the measurement period. A failure to supply agreed data might suspend the affected metric. A deliberate bypass of a required safety control may make events ineligible and trigger incident governance.

Employee incentives should align with the contracted outcome and guardrails. Paying only for throughput can encourage superficial completion. Penalising overrides can suppress necessary human judgment. The operating design should reward correct escalation, evidence quality and sustainable adoption. Employment, consultation, monitoring and discrimination requirements need jurisdiction-specific review.

Customers also shape value. A system that accelerates credit decisions may increase applications but create confusion if explanations and appeal routes are weak. A customer-service agent may resolve cases faster while customers reopen them later. The measurement schedule should include delayed or downstream indicators where they are material.

9. Build the attribution bridge

Attribution begins with the observed change and removes amounts unsupported by the contract. The bridge should identify volume, price, mix, inflation, seasonality, policy, staffing, other technology, process redesign, vendor contribution, customer contribution, quality failure and incremental cost. Each adjustment needs an owner, data source and approval rule.

The parties should avoid converting every uncertainty into a discretionary committee decision. Pre-agreed formulas are useful where inputs are observable. A bounded expert determination can address specialised disputes. A joint governance body can approve documented exceptional adjustments within a monetary limit. Material changes beyond that limit should require authorised amendment.

Attribution should recognise interaction. The vendor's model may create no value without customer process redesign, while the redesign may have limited value without the model. A negotiated allocation factor can reflect shared contribution where causal separation is impractical. The factor should be fixed or derived from a transparent rule before results are known.

Marketing statements should remain consistent with the evidence. The US Securities and Exchange Commission has taken enforcement action concerning false or misleading statements about the use of AI by investment advisers. [7] A value contract should preserve a clear difference between verified internal benefit, externally communicable claims and regulated disclosure.

10. Design the gainshare waterfall

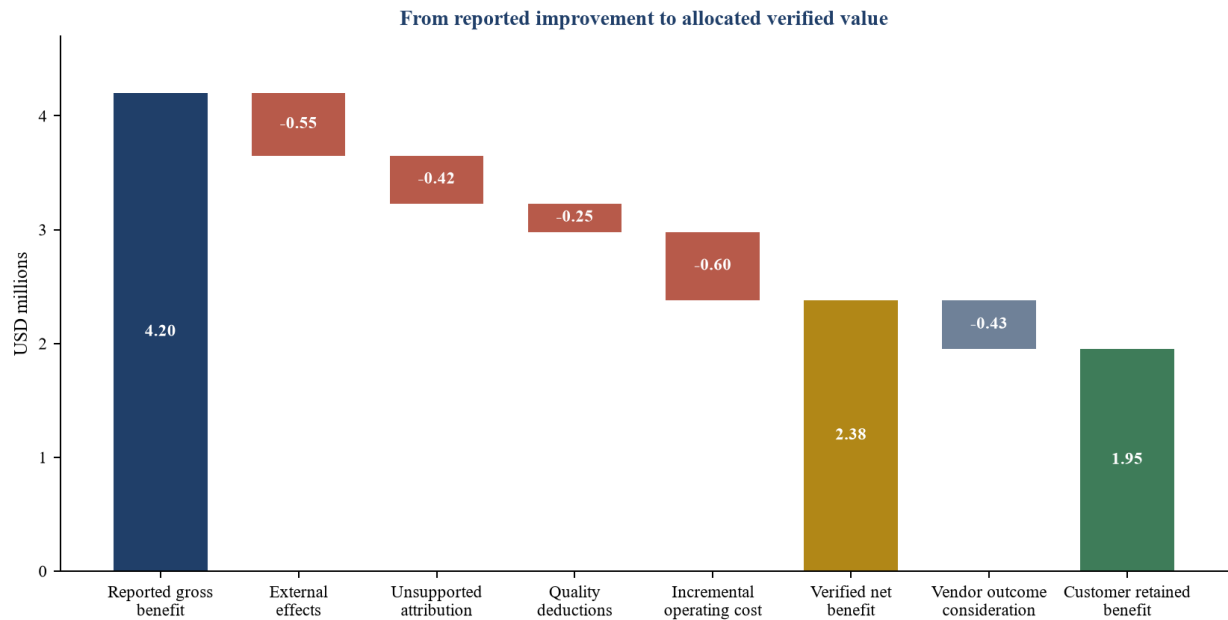
The waterfall should start from an agreed gross-benefit measure and deduct items in a fixed order. A typical sequence is external effects, unsupported attribution, reversals and quality failures, incremental customer operating costs, required third-party costs, vendor fixed fees already paid, threshold benefit, gainshare and caps. Taxes and accounting entries should follow applicable advice.

The contract should prevent double counting across metrics. A reduction in handling time and a reduction in employee cost may describe the same benefit. Revenue increase and contribution margin should not both be counted without a reconciliation. Working-capital release differs from recurring profit and should have its own treatment.

Gainshare tiers can reward exceptional performance while preserving customer economics. A minimum threshold can ensure that the customer first recovers implementation and disruption cost. A declining vendor share at higher benefit levels can recognise that broader organisational actions drive scale. A cap can limit variable consideration when measurement uncertainty remains material.

Payment timing should follow evidence maturity. Operational events may be measured monthly, while reversals, customer claims or cash collection emerge later. The parties can use provisional statements, holdbacks and annual true-up. The audit period should be long enough to observe agreed reversals without creating indefinite uncertainty.

Figure 3. Hypothetical AI value-capture waterfall



Original illustrative model. All amounts are USD millions and wholly hypothetical.

Table 3. Hypothetical central value waterfall

Step	USD million	Calculation or treatment
Reported gross benefit	4.20	Observed change before attribution review
Less external effects	0.55	Volume, price and market changes
Less unsupported attribution	0.42	Benefit without sufficient causal evidence
Less quality deductions	0.25	Reversals, rework and guardrail failures

Step	USD million	Calculation or treatment
Less incremental operating cost	0.60	Integration, review, infrastructure and support
Verified net benefit	2.38	Amount entering the commercial waterfall
Vendor outcome consideration	0.43	Hypothetical tiered gainshare after threshold
Customer retained benefit	1.95	Verified net benefit less vendor outcome consideration

Original illustrative model. Values describe no organisation and are unsuitable as benchmarks.

11. Test the hypothetical economics

The illustration assumes an annual addressable cost base of USD 20.0 million. The baseline is frozen before deployment and adjusted for agreed volume and wage-rate movements. The measurement period reports USD 4.20 million of gross improvement. This equals 21.0 per cent of the addressable base, but that percentage is not treated as verified value until the attribution bridge is complete.

External effects account for USD 0.55 million. Unsupported attribution accounts for USD 0.42 million. Reversals and quality failures remove USD 0.25 million. Incremental infrastructure, integration, assurance and human-review cost is USD 0.60 million. Verified net benefit is therefore USD 2.38 million.

The hypothetical gainshare gives the customer the first USD 0.50 million of verified net benefit. The vendor receives 25 per cent of the next USD 1.50 million and 15 per cent of benefit above USD 2.00 million. The resulting vendor outcome consideration is USD 0.43 million after rounding. Customer retained benefit is USD 1.95 million. Fixed implementation and subscription charges would be treated separately and should be included when the investment committee assesses total economics.

The illustration should be read as a calculation design. A live contract could use different tiers, fixed fees, caps, currencies and measurement periods. Its inputs would need to be supported by verified operational and financial data. The value schedule should identify whether figures are before or after tax and whether cash, accounting profit or avoided expenditure is being measured.

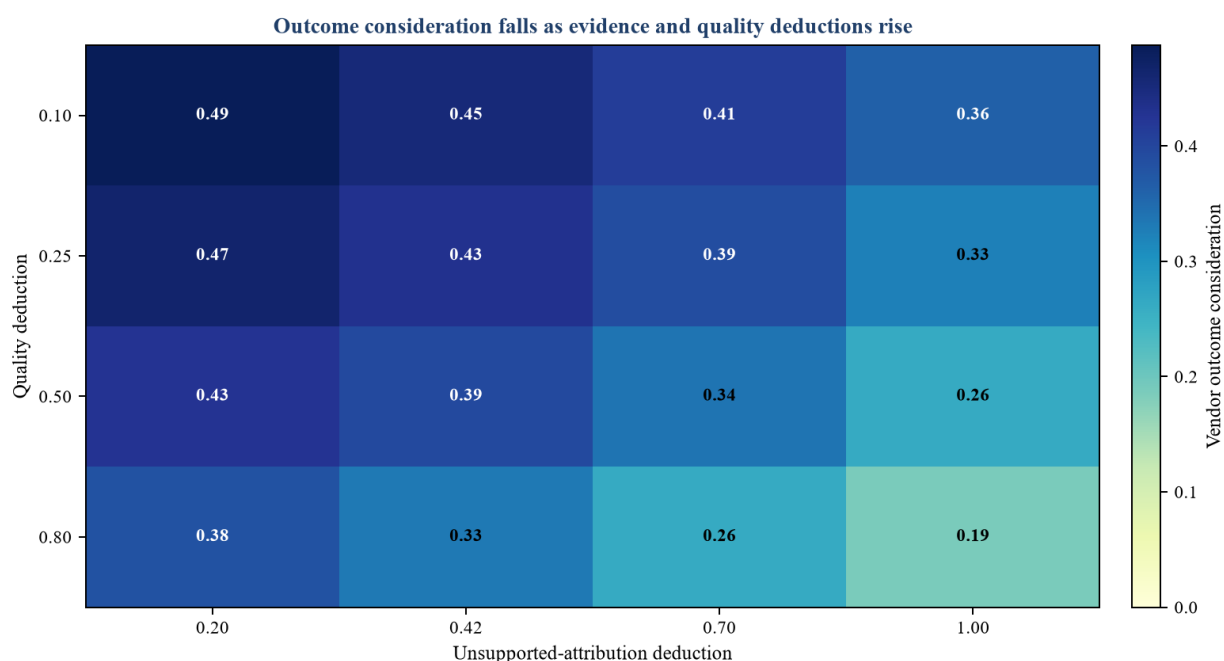
12. Run sensitivity and break-point analysis

The economic committee should test the variables that can reverse the commercial conclusion. These commonly include the baseline trend, comparison-group validity, adoption, quality failures, vendor attribution, incremental operating cost, benefit persistence and fee cap. A single central estimate hides the dispute risk carried by those assumptions.

In the hypothetical model, lower evidence confidence increases the unsupported-attribution deduction. A weak comparison method can therefore reduce verified benefit even when the observed improvement is unchanged. Higher quality failure also reduces eligible value. The interaction matters because poor evidence and poor quality may affect the same events; the contract should define the order of deductions and avoid duplicate adjustment.

The payment cap should be tested against the customer's retained benefit. If the vendor can earn the full cap while the customer has not recovered implementation and operating cost, the structure may reward technical output without economic value. The contract can address this through a customer-first threshold, minimum retained-benefit ratio or cumulative true-up.

Figure 4. Hypothetical vendor outcome consideration by evidence and quality deductions



13. Allocate liability and remedies

Liability should follow control, information and the ability to prevent harm. The vendor may control model development, service security and disclosed limitations. The customer may control purpose, deployment context, user access and final decisions. A third-party model or data provider may control an important upstream component. The responsibility matrix should inform warranties, indemnities, exclusions, caps and insurance requirements.

The EU Artificial Intelligence Act establishes risk-based obligations for providers, deployers and other actors within its scope. Its requirements include governance, technical documentation, logging, transparency, human oversight, accuracy, robustness and cybersecurity for relevant systems. [4] The applicable role and obligation require system-specific legal analysis. Contract labels cannot remove statutory responsibility.

Remedies should address the nature of failure. Incorrect invoices need correction and interest rules. Service failure may need credit and cure. Data misuse may require deletion, notification and indemnity. Repeated model-performance failure may require retraining, rollback or termination. A harmful decision may require customer remediation and regulatory response. The contract should preserve urgent protective action while later allocating cost.

Outcome fees can create conduct risk. A vendor paid for cost reduction may have an incentive to recommend fewer reviews. A customer paid team may suppress incidents that reduce the fee pool. The governance design should protect independent risk, compliance, safety and audit functions from commercial pressure.

14. Govern model change and drift

AI systems change through new model versions, prompts, retrieval corpora, policies, integrations, data distributions and user behaviour. A contract that fixes a performance fee while allowing uncontrolled system change creates an unstable measurement object. The change schedule should classify modifications by materiality.

Minor changes may be logged and included after routine tests. Material changes should require impact assessment, regression testing, baseline review and authorised release. A change that alters the workflow or eligible population may require a new measurement period. Emergency security changes may proceed under an expedited route with subsequent review.

The OECD AI Principles call for traceability across datasets, processes and decisions and for continuing risk management through the AI lifecycle. [6] The contract can operationalise this through version records, monitoring, incident thresholds and rights to override, repair or decommission the system.

Drift monitoring should connect technical indicators to economic and customer outcomes. A fall in precision may increase rework before it appears in headline cost. A change in customer mix may reduce model performance and invalidate the baseline. The governance body should define warning levels, stop levels, owners and evidence for restart.

15. Establish assurance and audit

Assurance should test the calculation, data, system and controls. Finance can reconcile the value bridge. Internal audit or an agreed independent reviewer can test lineage and control operation. Technical specialists can evaluate model performance and security. Privacy, legal, employment and sector specialists can assess obligations within their competence.

Audit rights should be proportionate and protect intellectual property, confidentiality, security and personal data. The customer may need sufficient evidence to validate fees without receiving model weights or unrelated vendor information. Possible mechanisms include agreed reports, secure inspection, third-party assurance, sampling, clean-room review and expert determination.

CISA and the UK National Cyber Security Centre's joint secure-AI guidance addresses secure design, development, deployment and operation and emphasises ownership of customer security outcomes. [9] A value contract should treat security as a continuing service obligation and a condition of eligible value where a breach affects the measurement period.

The evidence pack should be generated during operation. Retrospective reconstruction is expensive and invites dispute. A monthly pack can include version logs, population reconciliation, baseline inputs, event files, guardrail results, dependency performance, adjustments, incidents, approvals and provisional fee calculations.

Table 4. Proposed monthly evidence pack

Evidence item	Owner	Control test	Retention trigger
Eligible event population	Operations	Reconcile source to calculation	Final settlement plus audit period
Baseline and adjustments	Finance	Reperform formula and approval	Contract and accounting policy
Model and system versions	Technology	Match production logs to approved release	System and incident requirements
Guardrail outcomes	Risk and service owner	Sample failures and reversals	Complaint and regulatory requirements
Customer dependencies	Programme owner	Confirm dated performance or exception	Measurement and dispute period
Provisional fee statement	Finance	Independent review and sign-off	Invoice and statutory record rules

Original framework. Evidence access and retention require privacy, security and legal review.

16. Create the dispute ladder

The dispute mechanism should begin with evidence and escalate by issue type. Operational data differences can be reconciled by named analysts. Accounting classifications can be referred to finance specialists. Model-performance questions may need a technical expert. Contract interpretation belongs to the agreed legal process. A single general committee may lack the competence or independence for all four.

The schedule should state notice periods, undisputed-payment obligations, evidence access, expert appointment, standard of review, allocation of cost and finality. The expert's question should be narrow. For example, determine whether the calculation applies the agreed baseline formula, rather than decide what is commercially fair.

Materiality thresholds can keep immaterial differences out of formal escalation. Repeated small differences may still indicate a control problem and should be aggregated. Urgent safety, privacy or security matters should bypass the commercial ladder and follow incident procedures.

The parties should preserve a complete decision record. A settlement may amend future calculations, reverse prior amounts or close a one-off exception. The effect should be stated explicitly so that the same issue does not reappear under a later measurement period.

17. Separate economics, accounting and disclosure

Verified operational benefit, contractual consideration, accounting recognition and cash payment are different concepts. The contract should define commercial calculations without asserting the accounting result. Finance should assess the applicable standards, entity facts and contract terms.

IFRS 15 addresses variable consideration and the constraint applied when consideration may reverse. [10] IAS 38 distinguishes research expenditure from development expenditure that meets specified recognition criteria. [11] These requirements can affect the timing and presentation of AI-related revenue and cost, but they do not determine whether the operating initiative created value.

The business case should show gross benefit, incremental operating cost, fixed vendor charges, variable vendor charges, implementation investment, working capital, tax and cash timing. This prevents a large operational benefit from obscuring a weak cash return. It also supports reconciliation between the investment approval and realised outcome.

External disclosure should use evidence appropriate to the claim. A company may disclose an AI initiative, an operating metric or a financial effect. Each statement needs a defined basis, period and control. Contract confidentiality and customer-data restrictions also apply. The governance body should approve who can use outcome data in marketing, fundraising or investor communications.

18. Execute through 30, 90 and 180-day gates

The first thirty days should settle the commercial hypothesis. The parties identify the use case, value event, guardrails, baseline options, data population, responsibilities, fixed and variable fee boundaries and legal perimeter. The committee should stop or redesign a proposed outcome fee when attribution cannot be made sufficiently objective.

By day ninety, the parties should have a signed measurement design, tested data lineage, approved system version, acceptance results, operating responsibilities, incident route and provisional reporting pack. A limited pilot can test whether events and reversals are observable. Pilot success should be defined before results are known.

By day 180, the organisation should have completed at least one controlled measurement cycle, reconciled the baseline, tested guardrails, issued a provisional value statement and resolved exceptions. The

committee can then approve scale, revise the contract or discontinue the outcome mechanism. The decision should consider customer retained benefit, control performance and evidence quality together.

The final operating state is a repeatable value-control system. Each fee statement can be traced to approved definitions, versioned data, observable events, agreed adjustments and authorised approvals. Each material change enters through change control. Each incident has an owner and remedy. Each external value claim has evidence.

This structure can improve alignment because it makes contribution, dependency and uncertainty visible before money is allocated. It cannot remove causal uncertainty or replace professional judgment. Its purpose is to place uncertainty inside a governed commercial process that boards, vendors, employees, customers, auditors and advisers can examine.

19. Adapt the contract to the use case

The common architecture should be adapted to the economics and risk of each use case. A customer-service system may produce large event populations, short feedback cycles and observable reversals. A research assistant may influence decisions through a longer chain and require qualitative review. A predictive-maintenance system may generate avoided-cost claims whose counterfactual cannot be observed directly. A pricing system may affect revenue, customer fairness and regulatory obligations at the same time.

For customer service, the value event can be an accepted resolution that remains closed through a defined window. The baseline can compare matched case types, channels and customer segments. Guardrails can include complaints, escalation, vulnerable-customer outcomes and agent overrides. The calculation should distinguish deflection from abandonment and should recognise the cost of human review, quality assurance and remediation.

For software development, lines of code or generated suggestions are weak value measures. Better events include accepted change, escaped defect, lead time, service incident and rework. The baseline should control for project difficulty and team composition. Security, licensing, provenance and review compliance are essential guardrails. Productivity claims should identify whether saved time becomes lower cost, faster release, additional scope or unused capacity.

For sales and marketing, an AI-generated lead or message is remote from cash. The event chain should progress through eligibility, accepted opportunity, signed contract, delivered service and collected cash. Attribution should account for campaign, sales effort, pricing, product and market conditions. Customer consent, marketing law, discrimination and brand conduct may constrain system use.

For finance operations, duplicate-payment prevention, reconciled invoice handling and collection prioritisation can create observable events. The control design should protect segregation of duties, approval authority and evidence retention. Avoided loss requires proof that the payment or error would probably have occurred under the agreed comparison method. Working-capital improvement should be separated from recurring earnings.

For industrial and infrastructure applications, value can depend on physical reliability, safety and asset condition. A predicted failure that prompts maintenance may avoid an outage, but the counterfactual is uncertain. The parties can use engineering evidence, historical failure rates and bounded expert determination. Safety and regulatory controls should remain outside commercial override.

The use-case schedule should record why its selected measurement method is appropriate. This avoids copying a standard gainshare formula into a context where events, causality or safeguards differ materially. The same vendor may therefore operate under distinct measurement schedules across functions while using common data, assurance and dispute principles.

20. Integrate procurement, finance and operations

An effective value contract begins before a request for proposal is issued. Procurement should ask bidders to describe the value event, dependencies, evidence, limitations, data use, performance tests, change process and commercial model. Comparable pricing requires a common scenario and workload. A low unit price can be misleading when integration, human review, infrastructure, assurance and exit costs sit elsewhere.

Finance should own the benefit definition and reconciliation. Operations should own the process event and adoption evidence. Technology should own system configuration, integration and performance. Risk and legal functions should approve safeguards and obligations. Procurement should preserve competitive tension and contractual consistency. The executive sponsor should decide trade-offs and remain accountable for the business result.

The business case should use an auditable cost taxonomy. Initial expenditure can include discovery, data preparation, integration, security, testing, process design, training and change. Continuing expenditure can include licences, model usage, hosting, monitoring, human review, assurance, support and remediation. Exit expenditure can include data return, migration, replacement, parallel operation and contract settlement.

Commercial comparison should test at least three structures: fixed subscription, fixed plus milestone, and fixed plus outcome consideration. The analysis should show expected and downside customer cost, vendor incentives, measurement burden, termination flexibility and value retained. The chosen structure should reflect the organisation's ability to operate the measurement controls, not only the apparent alignment in the headline fee.

Procurement should verify subcontractors and upstream dependencies. A service may rely on a foundation-model provider, cloud platform, specialist data set and human-review supplier. Changes in those dependencies can affect price, rights, security and performance. The agreement should require relevant notice, flow-down obligations and an approved route for material substitutions.

Renewal should depend on cumulative evidence. A vendor can meet a first-year target through unusual conditions or deferred cost. The renewal review should examine benefit persistence, customer retained value, control performance, incidents, model changes, unresolved dependencies and total cost. An extension option should not automatically carry forward a baseline that has become obsolete.

21. Use a board-level value ledger

The board or investment committee needs a compact record connecting approved AI investments to measured outcomes and residual risk. A value ledger can list the use case, owner, approved cost, target event, baseline method, measurement period, verified benefit, cash effect, vendor consideration, customer retained value, guardrail status, material incidents and next decision.

The ledger should preserve the calculation behind every summary number. Management estimates should be identified as estimates and separated from verified results. Forecast benefit should not be combined with realised benefit. Operational improvement should be reconciled to financial statements or cash records when such a connection is claimed.

Portfolio reporting can expose concentration. Several use cases may depend on the same model provider, data platform, integration team or customer process. Their benefits may overlap. A shared reduction in support headcount cannot be allocated fully to multiple systems. The portfolio ledger should reconcile shared costs and benefits and identify common failure points.

The ledger can also support capital allocation. Initiatives with strong evidence and positive retained value may qualify for scale. Initiatives with technical performance but weak adoption may need process

redesign. Initiatives with uncertain attribution may move to fixed pricing. Initiatives with repeated guardrail failure may require suspension or closure.

Board reporting should include negative results. A transparent record of failed hypotheses, stopped pilots and reversed benefits improves future selection and baseline design. It also reduces the incentive to keep uneconomic systems operating to protect a public success narrative. Confidentiality and privilege should be maintained according to the organisation's governance and legal advice.

The value ledger does not replace statutory accounts, operational dashboards or risk registers. It connects them for a specific decision: whether the organisation should continue allocating capital and management attention to the contracted AI system. Its strength comes from a stable chain between approved purpose, measured event, evidence, economics, safeguards and action.

Management should reconcile the ledger at a fixed cadence. The operating owner confirms the event population and exceptions. Finance verifies calculations, cost and cash treatment. Technology confirms the production configuration and performance record. Risk functions report guardrail breaches and open remediation. Procurement confirms charges and contractual milestones. The accountable executive signs the resulting decision and records any dissent.

The ledger should also show remaining uncertainty. A benefit can be verified under the contractual calculation while still depending on a short observation period or a narrow customer population. The committee can therefore distinguish measurement completion from evidence maturity. It can retain a holdback, extend monitoring or limit scale until persistence is demonstrated.

At termination, the final ledger becomes part of the closeout pack. It reconciles provisional and final fees, unresolved claims, data return or deletion, system access, transition support and continuing obligations. This gives the parties a common record of the economic outcome and reduces dependence on separate commercial, technical and legal narratives.

Frequently asked questions

What is an AI value-capture contract

It is a contract structure that links part of commercial consideration to a defined and measured business outcome. It includes a baseline, comparison method, evidence rules, guardrails, responsibilities, allocation waterfall and dispute process.

Should every AI vendor be paid through gainshare

No. Gainshare is most workable when the value event is observable, attribution is sufficiently credible and the parties can govern dependencies. Fixed or milestone pricing may be more suitable where outcomes are remote, volatile or controlled mainly by the customer.

How should the baseline be chosen

The parties should use the strongest feasible comparison method before deployment. Options include a contemporaneous control, phased rollout, matched units, difference-in-differences, a validated forecast or a frozen historical baseline with explicit limitations.

Who owns data created by the AI system

Ownership and permitted use depend on the data class, contracts and applicable law. The agreement should separately address customer data, vendor data, third-party data, prompts, outputs, logs, derived information and model-improvement rights.

How can the parties avoid double counting value

The measurement schedule should reconcile operational events to financial effects and apply adjustments in a fixed order. Related metrics such as time saved, cost avoided and margin gained should be connected through one value bridge.

What happens when the model changes

The change schedule should classify materiality. Material model, data, prompt, policy or workflow changes can require impact assessment, regression testing, baseline review and a new measurement period before outcome fees resume.

Can the vendor audit customer records

Only within the agreed rights and applicable privacy, confidentiality and security limits. The parties can use sampling, secure inspection, independent assurance or expert determination when direct access would expose protected information.

Does verified operational value determine the accounting treatment

No. Operational value, contractual consideration, accounting recognition and cash payment are separate. Qualified finance and accounting specialists should assess the relevant standards and specific terms.

References

- [1] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework 1.0. 26 January 2023. Accessed 14 September 2026. <https://www.nist.gov/itl/ai-risk-management-framework>
- [2] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST AI 600-1. July 2024. Accessed 14 September 2026. <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf>
- [3] UK Information Commissioner's Office. Contracts and third parties: artificial intelligence audit framework. Accessed 14 September 2026. <https://ico.org.uk/for-organisations/advice-and-services/audits/data-protection-audit-framework/toolkits/artificial-intelligence/contracts-and-third-parties/>
- [4] European Union. Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. 13 June 2024. Current consolidated version accessed 14 September 2026. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- [5] World Bank and Inter-American Development Bank. Impact Evaluation in Practice, Second Edition. 2016. Accessed 14 September 2026. <https://www.worldbank.org/en/programs/sief-trust-fund/publication/impact-evaluation-in-practice>
- [6] OECD. OECD AI Principles. Updated 2024. Accessed 14 September 2026. <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>
- [7] US Securities and Exchange Commission. SEC Charges Two Investment Advisers with Making False and Misleading Statements About Their Use of Artificial Intelligence. 18 March 2024. Accessed 14 September 2026. <https://www.sec.gov/newsroom/press-releases/2024-36>
- [8] US Government Accountability Office. Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities, GAO-21-519SP. 30 June 2021. Accessed 14 September 2026. <https://www.gao.gov/products/gao-21-519sp>
- [9] US Cybersecurity and Infrastructure Security Agency and UK National Cyber Security Centre. Guidelines for Secure AI System Development. 26 November 2023. Accessed 14 September 2026. <https://www.cisa.gov/news-events/alerts/2023/11/26/cisa-and-uk-ncsc-unveil-joint-guidelines-secure-ai-system-development>

- [10] IFRS Foundation. IFRS 15 Revenue from Contracts with Customers. Issued standards and supporting materials. Accessed 14 September 2026. <https://www.ifrs.org/issued-standards/list-of-standards/ifrs-15-revenue-from-contracts-with-customers/>
- [11] IFRS Foundation. IAS 38 Intangible Assets. Issued standards. Accessed 14 September 2026. <https://www.ifrs.org/issued-standards/list-of-standards/ias-38-intangible-assets/>
- [12] OECD. Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions. 18 September 2025. Accessed 14 September 2026. <https://doi.org/10.1787/795de142-en>

About the Author

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 closure.

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 and leads 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.

He is also 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.

<https://www.linkedin.com/in/ckadya/>

<https://www.matchpoint-partners.com/team/ck-adya.html>

This paper is part of a continuing series on the structure of private and alternative markets. The views expressed are the author's own. The paper is for information only, describes market structure in general terms, and does not constitute investment, legal, tax or regulatory advice or a recommendation in respect of any security, vehicle or counterparty.