

Agentic AI M&A and the Value of Completed Work

A Decision Framework for Workflow Evidence Autonomy Exceptions Synergies and Transaction Structure

Authored by Chennakeshav Adya

Independent Researcher

September 2026

Abstract

Agentic artificial intelligence changes the unit of analysis in technology transactions. Software seats, licences and model access remain relevant commercial inputs. They do not show whether an agent completes an economically useful task, whether the result is accepted, how much human intervention is required, or whether the workflow can operate within legal, security and customer constraints. Buyers and investors therefore need an evidence system that connects technical capability to completed work, contribution, cash and transferable control.

This paper develops a decision framework for mergers and acquisitions involving agentic AI companies and AI-enabled businesses. It defines an accepted work unit, separates assistance from bounded autonomy, maps exception and reversal costs, and converts production telemetry into cohort economics. The method also tests workflow ownership, customer acceptance, data and intellectual-property rights, identity and authorisation, model dependence, evaluation quality, human accountability, integration requirements and synergy delivery.

Four hypothetical operating cases illustrate the framework: a customer-resolution agent, a finance-close agent, a software-remediation agent and a regulated-document agent. Management assumptions produce annual revenue from USD 58 million to USD 145 million, contribution from USD 9 million to USD 62 million and adjusted contribution after exception and control costs from USD 4 million to USD 49 million. A separate probability-weighted illustration produces enterprise value of USD 615 million. These figures explain the method. They are not market observations, forecasts or valuation conclusions.

The analysis finds that completed work becomes a useful valuation unit when the buyer can reproduce the task definition, acceptance test, trace, exception policy, cost ledger and customer payment evidence. Autonomy should be treated as a controlled operating characteristic rather than a marketing label. Transaction value should be released against verified production outcomes, durable customer economics, transferable rights and tested post-close integration. Earn-outs, holdbacks and integration milestones can allocate uncertainty where evidence remains immature.

JEL Classification: G24, G34, L86, M15, O32, O33

Keywords: agentic AI M&A, completed work, AI agent valuation, workflow economics, autonomy, exception rates, AI due diligence, technology transactions

Introduction

Agentic AI products are increasingly presented as systems that can plan, use tools and act across enterprise workflows. The transaction question is narrower: what work is completed, accepted and paid for, under which controls, at what full cost, and with what degree of transferability? A buyer that values a target through seats or message volume can overstate value when usage creates retries, supervision or

unaccepted output. It can understate value when a small number of agents replaces costly queues and produces a measurable operational outcome.

The distinction matters because agentic work combines software, models, orchestration, permissions, data, human review and operating process. NIST launched an AI Agent Standards Initiative in 2026 around interoperability, security, identity and authorisation [7-10]. NIST's AI Risk Management Framework and Generative AI Profile provide broader governance structures [11-12]. Current standards activity confirms the importance of controls; it does not establish that any target's agent is safe, reliable or commercially valuable.

Task evaluation also requires care. METR defines a task-completion time horizon as the human task duration at which an agent is predicted to succeed at a stated reliability level, and warns that its suite is concentrated in software, machine-learning and cybersecurity tasks [13-15]. Production workflows are often less specified, more contextual and harder to score. A transaction team therefore needs target-specific evidence in addition to benchmarks.

This paper links operating telemetry to acquisition decisions. It asks how a buyer should define a work unit, measure acceptance, reconstruct cost, classify autonomy, price risk, identify synergies and design consideration. The aim is a diligence record that can survive investment-committee, board, accounting, legal, security and integration review.

1 Define the transaction decision and valuation perimeter

The diligence file should begin with the decision. A strategic buyer may seek a new product, lower operating cost, proprietary data, specialist talent, customer access or control of an orchestration layer. A financial sponsor may seek scalable recurring revenue, margin expansion and an exit route. A corporate acquirer of an AI-enabled business may be valuing an operating transformation rather than a software company. Each rationale requires different evidence and creates a different synergy case.

The perimeter should identify legal entities, repositories, model and tool dependencies, prompts, policies, evaluation datasets, connectors, credentials, customer contracts, data rights, workflow documentation and personnel. It should distinguish owned assets from open-source components, third-party models, customer configurations and partner-controlled services. A demonstration can combine these elements without establishing which rights transfer at closing.

The buyer should also define the unit of account. The target may sell software, a managed service, an outsourced outcome or a hybrid. Revenue may be priced by seat, token, transaction, resolution, project or minimum commitment. The valuation model should follow the actual contractual promise and operating burden. Completed work is useful when it aligns the technical record with customer acceptance and the economic obligation.

2 Define an accepted completed-work unit

A completed-work unit needs a bounded input, permitted actions, expected output, acceptance rule, time window and exception path. Examples include resolving an eligible support case, reconciling a defined account set, remediating a validated software defect or extracting specified clauses from an approved document population. The definition should state exclusions, customer dependencies and conditions that require human authority.

Acceptance must be observable. A case may be closed technically and reopened by the customer. A reconciliation may balance while using an unauthorised journal. A code change may pass a narrow test and fail a security review. A document may be summarised accurately while omitting a required citation. The accepted unit should therefore include quality, policy and outcome criteria rather than a system status alone.

The work ledger should retain the task identifier, cohort, model and workflow version, tools used, actions taken, elapsed time, human touches, exceptions, acceptance, reversal, customer charge and direct cost. It should preserve an auditable trace without retaining unnecessary personal or confidential information. This ledger provides the bridge from technical execution to revenue, contribution and liability.

3 Classify assistance autonomy and decision authority

Autonomy is not a single percentage. A useful classification distinguishes observation, recommendation, preparation, action with approval, bounded autonomous action and consequential decision. The same product may occupy several levels across a workflow. An agent can autonomously collect evidence, prepare a recommendation and wait for an authorised person to approve a payment, employment action or regulated conclusion.

The classification should record who defines the objective, who approves the plan, which tools are available, which data can be accessed, which actions are reversible and who owns the outcome. Identity and authorisation become transaction-critical when agents act across systems. NIST's current agent work identifies secure agent identity and authorisation as core adoption issues [7-10]. A buyer should test the target's implementation rather than rely on the existence of emerging standards.

Higher autonomy can reduce handling time and increase volume. It can also increase the cost of a faulty action. The valuation model should recognise the economic benefit only for workflows where permissions, controls, monitoring, escalation and recovery have been tested. Autonomy that cannot be bounded or audited may require additional control cost and a slower integration plan.

4 Build the workflow and tool-control map

The workflow map should show every system, data source, model, tool, human role and control from request to accepted outcome. It should identify where the agent plans, retrieves, transforms, decides, acts, verifies and records. It should also show which component can change without customer reacceptance and which component creates a contractual or regulatory dependency.

Tool permissions require specific diligence. The buyer should inventory credentials, scopes, service accounts, secrets, production and test separation, approval rules, transaction limits and emergency revocation. It should replay representative traces and verify that the agent selects the correct tool, supplies valid parameters and handles failures. Tool-call success is an intermediate measure; the accepted business outcome remains the economic measure.

The control map should include segregation of duties, restricted actions, rate limits, data-loss prevention, logging, monitoring, rollback and incident response. A target may have strong model evaluations and weak operational controls. The buyer should price the combined system that produces work, including the human and technical layers needed to operate it safely.

5 Design the production evaluation system

Production evaluation should begin with the customer's acceptance criteria. The test set should cover ordinary tasks, boundary cases, incomplete inputs, conflicting instructions, unavailable tools, policy restrictions and goal changes. It should include the consequences that matter to the workflow, such as money moved, access granted, customer promise made, code deployed or regulated conclusion recorded.

Trace-based evaluation can reveal whether an agent selected the right tool, followed a required handoff and complied with policy [16-18]. Outcome grading should be separated from process grading. A correct result reached through an unauthorised action remains a control failure. A compliant trace that produces an unusable result remains an economic failure. Both dimensions affect transaction value.

The target should maintain versioned datasets, rubrics, graders, human adjudication and regression thresholds. Management should reconcile evaluation results to production cohorts and investigate differences. A benchmark score has limited value when the production population, tool set, data, latency or human support differs materially from the test environment.

6 Measure acceptance exception reversal and recovery

Headline completion rate should be decomposed into accepted completion, exception, escalation, rejection, reversal and unresolved status. Exceptions include missing data, ambiguous objectives, unavailable systems, policy conflicts, model uncertainty and customer-specific rules. The exception taxonomy should be stable enough to support trend analysis and flexible enough to identify new failure modes.

Human intervention should be measured by purpose and duration. Review that protects a consequential decision is different from repeated repair of a weak workflow. The cost ledger should allocate triage, investigation, correction, approval, customer communication and recovery. Reopened cases and downstream remediation should be linked to the originating task where possible.

Recovery time is an operating and valuation measure. A system that fails infrequently but requires days of specialist intervention can create concentrated risk. The buyer should test detection, containment, rollback, customer notification and root-cause closure. Reserves, service credits and remediation cost belong in the transaction model when exceptions have financial consequences.

7 Convert telemetry into customer cohort economics

The customer ledger should connect contracted promise, eligible work volume, accepted units, price, implementation, inference, tool, data, review, support and collection. Revenue recognition follows the contract and applicable accounting requirements. Operational measurement should still reconcile each charge to the evidence that the promised service was provided [1-4].

Contribution per accepted unit is a useful starting point. It subtracts variable model, tool, infrastructure, human review, exception and partner costs from the related revenue. The buyer should also identify semi-variable customer engineering, evaluation, compliance and support costs. Shared research and platform expenditure should remain visible rather than being hidden through optimistic allocation.

Cohort analysis should track time to production, acceptance rate, expansion, price changes, contribution, collections, renewal and loss reasons. Minimum commitments can support cash while masking low utilisation. Usage growth can appear attractive while reducing contribution if tasks become harder or exceptions rise. Valuation should follow durable accepted work and cash rather than activity alone.

Pricing architecture should be tested against the cost driver. A per-seat price can work when seats correlate closely with eligible work and support burden. A per-outcome price transfers more operating risk to the provider and requires an agreed acceptance process. A consumption price can leave the customer exposed to retries and inefficient agent behaviour. Hybrid structures may combine a platform minimum with outcome or usage components. The buyer should model each contract under observed production behaviour and identify which party bears model-price, complexity and exception risk.

Cash conversion deserves its own cohort view. Implementation milestones, acceptance disputes, service credits, invoice timing and enterprise payment cycles can separate reported revenue from cash. A high-growth target may require working capital when supplier and payroll costs occur before customer acceptance and collection. The valuation model should reconcile booked revenue, billed revenue, receivables, deferred revenue and cash to the same customer and workflow records.

8 Reconstruct revenue quality and contractual obligations

Agentic AI contracts can contain pilots, implementation services, minimum commitments, usage fees, outcome fees and service-level credits. The diligence team should classify each component, verify enforceability and reconcile contract, invoice, revenue ledger, receivable and bank receipt. Announced customers and memoranda are evidence of commercial activity only to the extent of their executed terms and observed economics.

The contract should define the service, customer dependencies, acceptance, permitted use, data roles, model changes, support, security, audit, intellectual property, liability, termination and transition. A provider that promises a business outcome may carry a broader obligation than a software licensor. The valuation model should reflect the actual promise, including the cost of human operations required to deliver it.

Concentration requires a workflow view as well as a customer view. Several customers may depend on the same model, cloud, connector or implementation partner. A change to one shared dependency can affect multiple contracts. The buyer should test renewal economics under model repricing, stricter controls, lower minimums and customer demands for portability.

9 Reconcile model tool and infrastructure dependence

Many agentic products assemble third-party models, clouds, vector systems, observability services and enterprise connectors. This can accelerate product development while creating price, access, continuity and contractual dependencies. The buyer should inventory each dependency, its function, agreement, term, price, volume commitment, data treatment, change control and substitution path.

Portability should be tested through a representative workflow. A claimed multi-model architecture has value when the target can substitute a model while preserving acceptance, security, latency and cost. Switching may require prompt redesign, evaluation, customer approval and new safety controls. The model should include this transition cost and any period of lower performance.

Infrastructure economics should be reconstructed from invoices and telemetry. Cost per accepted unit should include failed attempts, retries, reasoning steps, tool calls, storage, retrieval, observability and review. A declining model price does not guarantee a declining workflow cost when agents attempt longer tasks or use more tools. The buyer should test both price and behaviour sensitivity.

10 Establish data intellectual-property and transfer rights

The rights register should cover source code, prompts, policies, workflow designs, evaluation sets, customer configurations, training and feedback data, synthetic data, trademarks, patents and documentation. It should identify creator, employer or contractor assignment, licence, permitted use, restrictions, sublicensing, change of control and termination. Open-source obligations and third-party terms should be mapped to each released component.

Customer data requires separate treatment. Access needed to operate a workflow may not create a transferable asset. Feedback may improve the product while remaining subject to confidentiality, privacy or use restrictions. The buyer should verify that production data is used only for authorised purposes and that deletion, retention and transition obligations can be performed after closing.

IFRS 3 requires an acquirer to recognise identifiable intangible assets separately from goodwill when the criteria are met, while IAS 38 addresses identifiable intangible assets and IFRS 13 addresses fair-value measurement [1-6]. The purchase-price allocation does not replace transaction diligence. Legal control, economic benefit, useful life, obsolescence and separability remain company-specific questions.

Technology obsolescence should be assessed at the workflow layer as well as the model layer. A prompt or routing method may be displaced quickly while customer integration, acceptance data, operating controls and domain process remain valuable. The diligence team should separate replaceable components from accumulated workflow evidence. It should estimate the cost, time and customer approval required to rebuild each component and the risk that a competitor can offer a comparable outcome with less friction.

Employee and contractor knowledge can be essential without being owned as a separable asset. The buyer should identify key maintainers, undocumented operating decisions, customer-specific knowledge and evaluation expertise. Retention and knowledge-transfer plans should align with the integration sequence. The valuation should not assume that source-code transfer alone preserves the system's capacity to deliver accepted work.

11 Test security identity authorisation and resilience

Agent security should be assessed across identity, authentication, authorisation, tool use, data access, supply chain, prompt injection, secrets, logging, monitoring and recovery. The target should demonstrate least-privilege design, credential rotation, environment separation, approval boundaries and rapid revocation. Security documentation should reconcile to tested production configuration.

Identity becomes more complex when several agents act through shared services. The buyer should be able to attribute an action to a user, agent, version, policy and credential. Delegation should be bounded by task, time, system and value where possible. A shared service account that obscures responsibility can weaken both control and customer trust.

Resilience includes model or cloud outage, tool unavailability, corrupted context, unexpected goal change and unsafe output. The diligence team should inspect incident records, recovery exercises, customer communication and remediation. A strong demonstration during normal operation does not prove that the system fails safely.

12 Preserve human accountability and customer authority

Human accountability should be designed into the workflow rather than added as a generic statement. Each consequential decision needs an authorised owner, information standard, escalation rule and retained record. The buyer should identify where a human approves, where a human reviews retrospectively and where an agent acts within a bounded policy.

The quality of review matters. A reviewer faced with excessive volume, poor explanations or automation bias may approve without meaningful scrutiny. Review effectiveness can be tested through sampled decisions, disagreement analysis, time spent, override frequency and downstream outcomes. Human touches that merely repair routine errors should be classified as operating cost.

Customer authority should also be explicit. An enterprise may authorise a target to prepare work while reserving final action. It may require notification before a model or workflow changes. The target's evidence should show how customer-specific controls are implemented and tested. Unrecorded operational exceptions can create post-close liability and integration delay.

13 Build the completed-work contribution bridge

The contribution bridge begins with billed or allocated revenue for accepted units. It subtracts model inference, tool charges, cloud, data, partner fees, implementation, review, support, service credits and exception remediation. It should show gross completion separately from accepted completion and identify cost generated by rejected or reversed work.

The bridge should then allocate workflow-specific engineering, evaluation, compliance and customer-success expense. Central research and platform costs remain outside unit contribution but inside company

cash flow. This structure helps the buyer distinguish a scalable workflow from a labour-intensive managed service presented as software.

Management should reconcile the bridge to the general ledger, payroll, supplier invoices and production telemetry. Variance analysis should explain changes in mix, complexity, model selection, latency, retry behaviour and human intervention. The transaction model can then test whether scale improves contribution or exposes a larger exception burden.

14 Value the company through evidence states

Traditional valuation methods remain relevant. Market multiples, discounted cash flow, precedent transactions and cost approaches can be used when their inputs reflect the target's economic reality [1-6, 19-22]. Agentic AI requires greater care with revenue quality, growth durability, contribution, capital needs, dependency risk and technological obsolescence.

An evidence-state model can supplement these methods. The first state demonstrates reproducible task completion. The second shows accepted production work and collected revenue. The third shows repeatable cohorts with durable contribution. The fourth shows transferable control, secure operations and strategic distribution. Probability and value assumptions should be documented and updated as evidence changes.

The model should avoid double counting. A forecast that already includes lower labour cost should not also add the full cost saving as a separate synergy. Data, technology and customer relationships may contribute through forecast cash flow and identifiable intangible assets. The valuation file should show where each source of value enters the analysis.

Comparable-company analysis should normalise business models before applying a multiple. An agent platform with usage revenue, an outcome-based managed service and an AI-enabled business-process operator can report similar growth with different delivery obligations and margins. Reported gross margin may use different treatments for inference, implementation and review. The transaction team should rebuild a common contribution measure and disclose remaining differences rather than force false comparability.

Discounted cash flow should include explicit investment in evaluation, security, customer integration, model transition and working capital. Terminal assumptions need a view of continuing renewal and obsolescence. A cost approach may inform replacement and remediation but may not capture customer access or future cash. Precedent transactions require careful reading because disclosed headline values rarely provide the workflow, rights and cost evidence needed for direct comparison.

15 Identify buyer-specific synergies

Revenue synergies may include access to the buyer's customers, embedded distribution, broader data rights, cross-selling or entry into new workflows. Cost synergies may include shared infrastructure, model purchasing, security, compliance, sales, support or elimination of duplicate tools. Capability synergies may shorten product development or improve the buyer's own operations.

Each synergy should have a baseline, action, owner, cost, timing, dependency and evidence gate. The buyer should distinguish synergy that requires target capability from value already present in the standalone plan. Customer consent, data-use restrictions, model changes and integration risk can delay or prevent delivery.

Completed-work metrics can make the synergy case testable. The baseline records eligible work, acceptance, time, cost and exceptions. The integration plan specifies the target workflow and control changes. The realised result is measured through accepted work, contribution and cash. This structure supports price discipline and post-close accountability.

16 Conduct customer and workflow diligence

Customer diligence should test why the workflow was purchased, how it was accepted, which alternatives were considered, who owns the budget, what remains manual and what would cause termination. The buyer should obtain permission and follow transaction protocols. Management-selected references should be supplemented with contract, usage, support and collection evidence.

The workflow file should identify implementation effort, customer data dependency, configuration, custom code, integration, evaluation, human roles and support. A target can show strong retention while carrying hidden services work. The buyer should compare contracted recurring revenue with the labour and specialist knowledge required to renew it.

Loss and non-conversion evidence is equally important. Failed pilots can reveal weak acceptance criteria, poor integration, security barriers, unclear ownership or insufficient economic benefit. The forecast should use observed conversion and renewal by comparable cohort. Management scenarios should remain clearly separate from observed results.

17 Design transaction structure around uncertainty

Transaction structure can allocate uncertainty where evidence is incomplete. Consideration may include cash at closing, rollover equity, retention, earn-out, holdback, escrow or contingent value. The measurement should be based on auditable outcomes that management can influence without encouraging unsafe deployment or short-term revenue recognition.

Useful milestones include accepted production work, signed minimum commitments, collected revenue, contribution after exception cost, customer retention, completed rights remediation, security closure and successful portability. Definitions should specify data sources, accounting policy, exclusions, governance and dispute resolution. A vague milestone can create post-close conflict.

Representations and covenants should address intellectual property, data use, open source, customer obligations, model and cloud dependencies, security incidents, evaluation records and material workflow changes. Specific indemnities or reserves may be appropriate for identified exposures. Legal, tax and accounting advice is required for the actual transaction.

Earn-out design should avoid incentives to increase gross volume while weakening acceptance or control. A balanced measure can combine collected revenue with minimum adjusted contribution, customer retention and defined control conditions. The buyer should protect ordinary integration rights while preventing changes that arbitrarily defeat the measure. Sellers need access to the underlying calculation and a clear review process. Both parties should model how customer concentration, pricing changes and platform migration affect the result.

Closing conditions should focus on matters necessary to transfer and operate the business. Examples include key assignments, customer and supplier consents, remediation of critical permissions, preservation of evaluation records and confirmation of transition support. Post-close covenants can address lower-severity work. The transaction file should show which exposure changes price, timing, structure or the decision to proceed.

18 Construct four hypothetical operating cases

The customer-resolution agent handles eligible service cases and charges by accepted resolution. It benefits from high volume but carries reopening, escalation and service-credit risk. The finance-close agent reconciles accounts and prepares evidence, with authorised people retaining approval over consequential entries. It has lower volume, higher control requirements and potentially stronger willingness to pay.

The software-remediation agent identifies, tests and proposes fixes within defined repositories. Acceptance requires tests, security review and deployment rules. The regulated-document agent extracts, compares and prepares findings from an approved document set, with citation and review requirements. Its economics depend on document quality, liability allocation and specialist review.

The cases use management assumptions solely to demonstrate the framework. A real target model requires contracts, invoices, collections, production telemetry, cost records, incident history, rights and customer evidence. The buyer should replace every assumption, assign an owner and retain the source.

19 Illustrative completed-work economics

The customer-resolution case assumes USD 145 million of revenue, USD 62 million of contribution and USD 49 million after exception and control costs. The finance-close case assumes USD 92 million of revenue, USD 36 million of contribution and USD 27 million after those costs. The software-remediation case assumes USD 78 million of revenue and USD 18 million of adjusted contribution.

The regulated-document case assumes USD 58 million of revenue, USD 9 million of contribution and USD 4 million after exception and control costs. Lower adjusted contribution reflects specialist review, customer-specific evaluation and liability controls. These figures are hypothetical management assumptions; they are not observations about any company or market.

Sensitivity should focus on accepted volume, price, model and tool cost, human intervention, exception, reversal, implementation, retention and collection. A small deterioration in acceptance can have a larger cash effect when rejected work still consumes inference, tools and staff. The model should show this operating leverage before applying a valuation multiple.

20 Integrate the target without destroying evidence

The first hundred days should preserve the evidence chain while consolidating systems. The buyer should freeze baseline definitions, retain traces and contracts, map credentials, identify critical personnel and agree customer-change protocols. It should avoid changing models, prompts, tools and controls simultaneously without a comparison plan.

Integration should proceed by workflow cohort. Each change needs a hypothesis, test, acceptance threshold, rollback and owner. Security and identity controls may need to be harmonised before broader deployment. Customer consent or reacceptance may be required when data location, model, subprocessor or operating process changes.

Synergy reporting should reconcile to the completed-work ledger and finance system. The integration team should report accepted units, exceptions, adjusted contribution, customer effects and cash. Product activity alone cannot establish realised transaction value.

The operating model should anticipate organisational change. Product, engineering, security, legal, finance, customer success and process owners may use different definitions of completion. Integration governance should agree one hierarchy of measures and preserve the underlying detail. A finance result that aggregates accepted units without exception cost can conflict with a security report that treats the same workflow as constrained. A reconciled data model helps the board evaluate value and risk together.

Customer communication should follow contractual and practical significance. Some customers may accept a change in ownership while requiring notice for a new model, subprocessor, hosting location or data use. Early mapping can prevent avoidable delay. Integration should protect service continuity and give customers evidence that controls, support and accountability remain effective.

21 Govern value after closing

Board reporting should combine commercial, technical, control and financial evidence. A useful scorecard includes accepted work, exception and reversal rates, human intervention, customer concentration, contribution, collections, model and tool dependency, security events, evaluation regressions and remediation status. Definitions should remain stable or show a documented bridge.

Management incentives should reward durable accepted outcomes, customer value, contribution and control performance. Volume-only targets can encourage low-quality deployment. Revenue-only targets can defer exception cost or create weak commitments. The governance system should preserve authority for consequential decisions and make material incidents visible.

The acquisition thesis should be reassessed at defined intervals. Evidence may support faster expansion, further investment, product consolidation or a narrower perimeter. An explicit evidence-state approach gives the board a disciplined way to change capital allocation as the operating record develops.

Conclusion

Agentic AI M&A requires a valuation bridge from technical activity to accepted economic work. Seats, tokens, tool calls and benchmark results describe parts of the system. Transaction value depends on whether the target can complete defined work, satisfy the acceptance standard, control exceptions, earn durable contribution, collect cash and transfer the relevant rights and operating capability.

Completed work is valuable as a diligence unit because it can connect trace, customer promise, cost and payment. The definition must include quality, policy and outcome. Autonomy should be classified by action and authority. Exception and recovery costs should remain inside the economics. Model, tool and infrastructure dependencies should be tested through contracts and portability.

The framework supports valuation, synergy analysis, transaction structure and post-close governance. It also creates disciplined boundaries for uncertainty. Hypothetical assumptions remain scenarios until the target's evidence replaces them. Consideration can then be linked to verified production, customer and control milestones rather than to broad claims about agent capability.

Completed-work evidence room

The evidence room should contain the workflow catalogue, accepted-work definitions, versioned traces, evaluation sets, exception taxonomy, customer acceptance, contracts, invoices, collections, cost ledger and incident record. Each schedule should identify owner, source system, period, customer cohort and reconciliation status. The transaction team should be able to reproduce sampled tasks and trace them to financial records.

Autonomy identity and control dossier

This dossier should record agent identities, credentials, permissions, tool scopes, approval boundaries, segregation of duties, monitoring, revocation, rollback and recovery. It should map each production workflow to the applicable customer and enterprise policies. Gaps should have a remediation owner, cost and completion test.

Rights dependencies and portability file

The file should map source code, data, prompts, evaluations, models, open-source components, clouds, tools, connectors and customer configurations. It should state ownership, licence, change-of-control treatment, restriction, renewal and substitution plan. Representative portability tests should record acceptance, latency, cost and customer impact.

Customer economics and synergy ledger

The ledger should reconcile eligible work, accepted work, exceptions, price, direct cost, adjusted contribution, invoice, collection and renewal by cohort. Synergies should have a baseline, action, owner, cost, timing and realised outcome. Duplicate recognition between standalone plan, purchase-price allocation and synergy value should be removed.

Transaction and integration control file

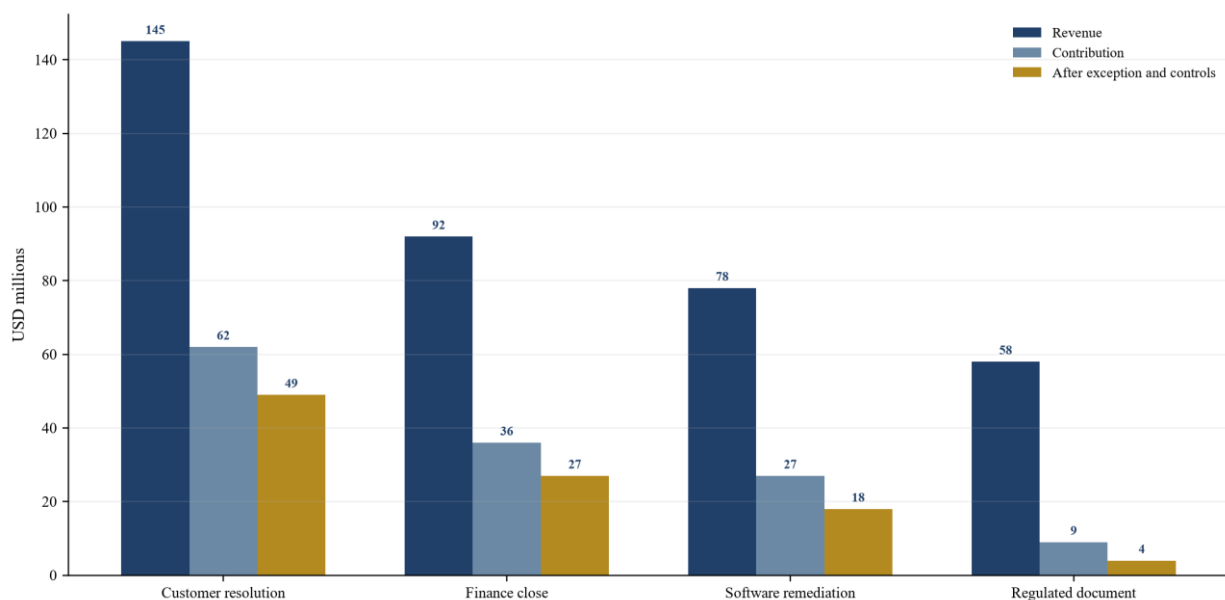
The file should contain consideration definitions, earn-out measures, holdbacks, representations, covenants, customer-change requirements, integration sequence and board reporting. It should preserve the pre-close baseline and define how post-close changes will be measured. Legal, accounting, tax, regulatory and security advisers should review the relevant components.

Figure 1 Agentic AI transaction value architecture



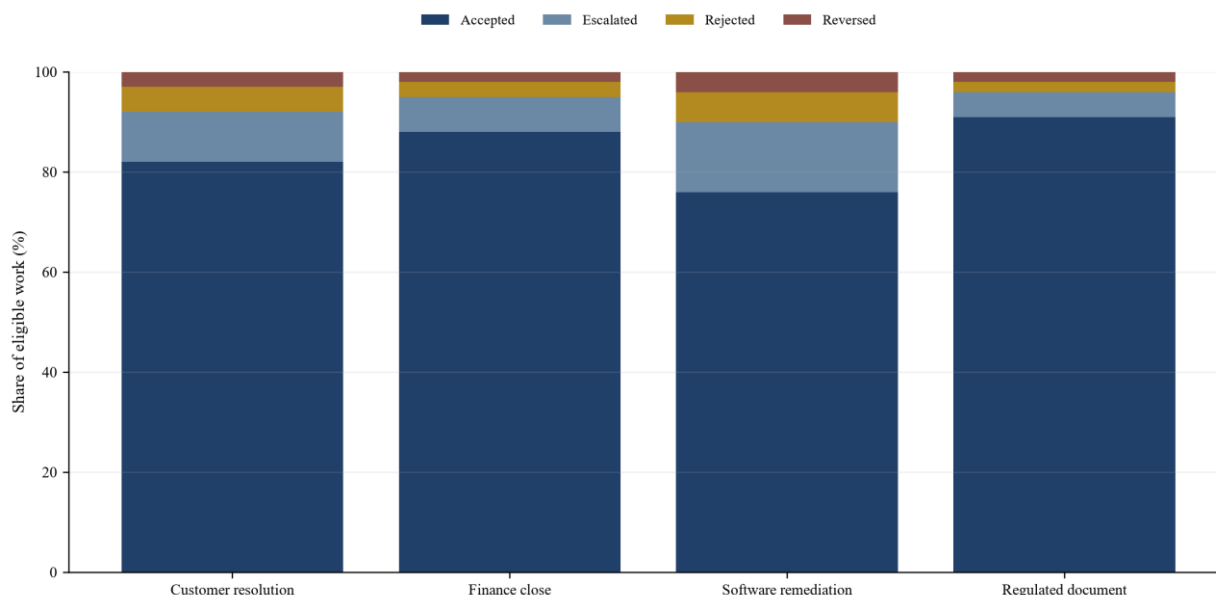
Proposed framework; conclusions require target-specific customer technical legal security and financial evidence.

Figure 2 Hypothetical annual revenue contribution and adjusted contribution



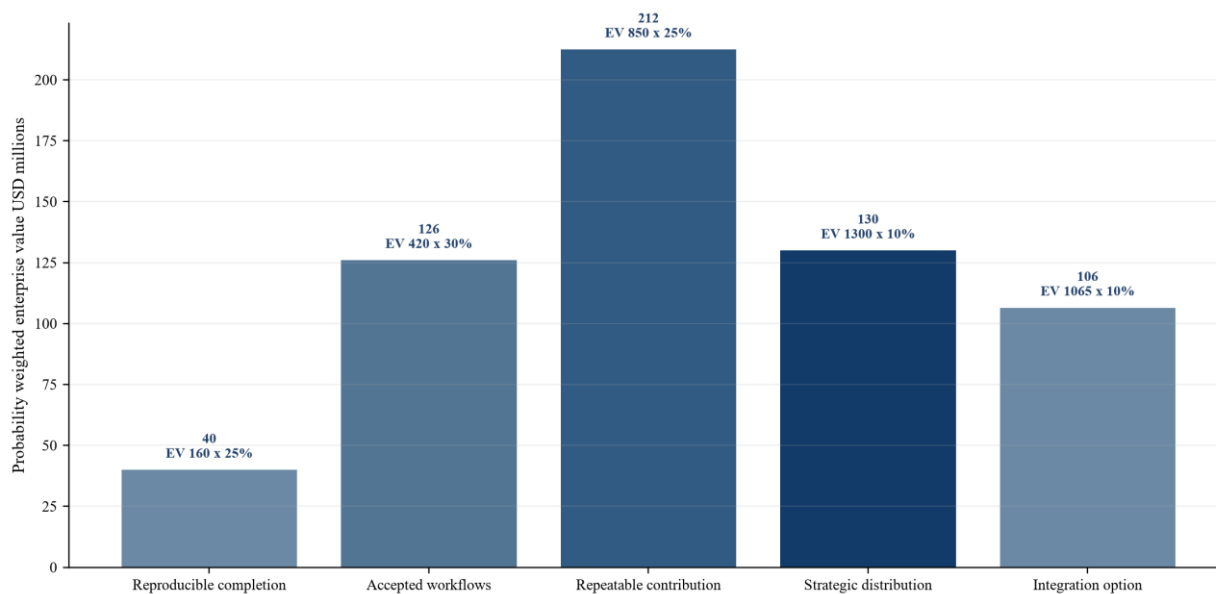
Management assumptions in USD millions; figures are not market observations forecasts or valuation conclusions.

Figure 3 Hypothetical acceptance and exception profile



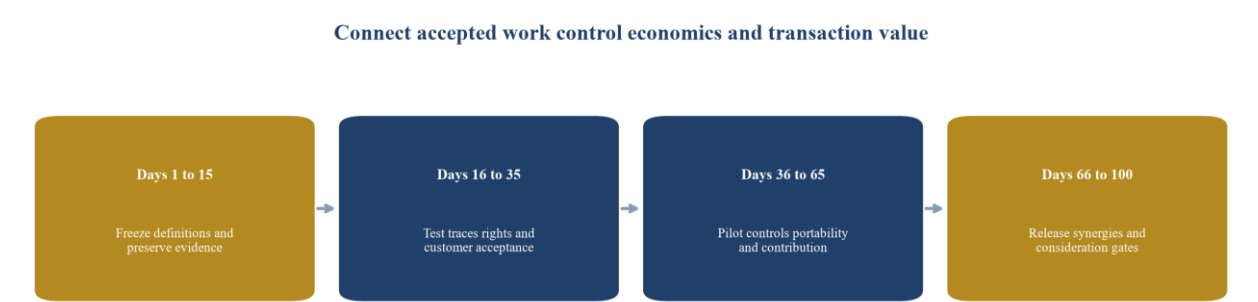
Management assumptions by completed-work case; percentages describe illustrative workflow outcomes.

Figure 4 Hypothetical probability weighted enterprise value by evidence state



Management assumptions in USD millions; total probability weighted value is USD 615 million.

Figure 5 First hundred days evidence and integration sequence



Proposed sequence; timing should follow customer security regulatory and integration requirements.

Table 1 Agentic AI transaction perimeter

Component	Required evidence	Valuation question	Principal risk
Workflow and acceptance	definitions traces evaluation customer sign-off	what work is completed and accepted	activity mistaken for outcome
Technology and rights	repositories assignments licences dependencies	what is owned and transferable	restricted or customer-controlled assets
Customer economics	contracts accepted units invoices collections renewals	which workflows create durable contribution	pilots or services mistaken for recurring software
Controls and resilience	identity permissions incidents recovery tests	can work operate safely at scale	unauthorised action or slow recovery
Integration and synergies	baseline owners costs milestones customer approvals	which buyer benefits are deliverable	double counting or delayed integration

Proposed diligence structure; requirements depend on the target customer and transaction.

Table 2 Completed-work evidence ladder

Evidence stage	Required record	Supported conclusion	Limitation
Demonstrated task	reproducible input trace output and grader	technical feasibility	may not reflect production
Accepted pilot	customer criteria sign-off exceptions and cost	workflow utility	commitment may be limited
Paid production	contract accepted units invoice and collection	commercial delivery	cohort may remain concentrated
Repeatable cohort	contribution renewal expansion and loss evidence	durable economics	dependencies may constrain scale
Transferable system	rights controls portability and integration test	transaction readiness	future change still requires governance

Each stage supports a different commercial or valuation conclusion.

Table 3 Autonomy and authority classification

Level	Agent role	Human authority	Evidence required
Observe	collect and organise evidence	defines purpose and access	source completeness and audit trail
Recommend	analyse and propose	reviews and decides	rubric rationale and disagreement record
Prepare action	populate transaction or change	approves before execution	validation approval and segregation
Bounded action	execute within explicit policy	sets limits and monitors	identity permission trace and rollback
Consequential decision	affects rights money safety or regulated outcome	retained by authorised person unless law and policy permit otherwise	decision record accountability and appeal

Proposed control map; actual boundaries require workflow-specific legal security and operational review.

Table 4 Hypothetical annual operating cases

Case	Revenue	Contribution	Exception and control cost	Adjusted contribution
Customer-resolution agent	145	62	13	49
Finance-close agent	92	36	9	27
Software-remediation agent	78	27	9	18
Regulated-document agent	58	9	5	4

Management assumptions in USD millions; figures are not market observations forecasts or valuation conclusions.

Table 5 Hypothetical customer-resolution contribution bridge

Item	USD millions	Evidence required
Revenue from accepted resolutions	145	contract acceptance invoice and collection ledger
Model tools and infrastructure	-31	workload telemetry and supplier invoices
Human review and exception operations	-27	time records workflow cases and payroll
Implementation partner and support	-25	project time supplier settlement and service evidence
Contribution	62	reconciled customer cohort schedule
Additional control remediation and service credits	-13	incident exception and credit records
Adjusted contribution	49	finance and workflow reconciliation

Management assumptions in USD millions; excludes central research sales administration financing and tax.

Table 6 Hypothetical valuation evidence states

Evidence state	Enterprise value	Probability	Weighted value
Reproducible task completion	160	25%	40
Accepted paid workflows	420	30%	126
Repeatable contribution across cohorts	850	25%	212.5
Strategic distribution and transferable control	1300	10%	130
Execution and integration option	1065	10%	106.5
Total		100%	615

Management assumptions in USD millions; this is not a valuation conclusion.

Table 7 Transaction and integration evidence gates

Gate	Required evidence	Transaction response	Post-close measure
Rights and dependencies	assignments licences component map and portability test	closing condition holdback or remediation	controlled releases and exceptions
Accepted customer work	production trace sign-off invoice and collection	base value after verification	acceptance contribution and cash
Controls and resilience	identity permissions incidents rollback and recovery test	reserve covenant or staged release	exceptions severity and closure
Repeatable economics	comparable cohorts renewal and adjusted contribution	earn-out or contingent value	durable contribution and retention
Synergy delivery	baseline action owner cost and customer approval	integration milestone	realised cash and customer outcome

Proposed framework; actual instruments require current legal tax accounting regulatory and financial advice.

Frequently asked questions

Why are software seats insufficient for valuing agentic AI

Seats measure access or users. Agentic AI economics depend on eligible work, accepted outcomes, exceptions, human intervention, full delivery cost and customer payment. Seats can remain a contractual measure while completed work provides the operating and valuation bridge.

What counts as completed work

A completed-work unit has a bounded input, permitted actions, expected output, acceptance rule, time window and exception path. It should satisfy quality, policy and outcome criteria and be traceable to customer and cost evidence.

How should autonomy be measured in an acquisition

Classify each workflow stage by action and decision authority: observation, recommendation, preparation, action with approval, bounded autonomous action or consequential decision. Verify identity, permissions, controls, monitoring, rollback and human accountability for the actual production configuration.

Which exception metrics matter most

Track escalation, rejection, reversal, reopening, unresolved work, recovery time, human effort, service credits and downstream remediation. The economic model should include the cost of unsuccessful attempts and control operations.

How should third-party model dependence affect value

Verify contract term, price, volume, data treatment, change control, continuity and substitution. Test a representative workflow on an alternative model and measure acceptance, latency, cost and customer impact before assigning portability value.

How should synergies be evidenced

Each synergy needs a baseline, action, owner, cost, timing, dependency and measurable result. Accepted work, adjusted contribution and cash provide stronger post-close measures than product activity alone.

Which evidence should affect consideration

Useful gates include transferable rights, accepted paid workflows, collected revenue, durable adjusted contribution, security closure, portability and customer-approved integration. Definitions and data sources should be written into the transaction mechanism.

Which Matchpoint practice is relevant to this work

This research connects to Matchpoint Partners' mergers and acquisitions, technology valuation, transaction diligence, strategy and execution, post-merger integration and capital advisory work for technology companies, strategic buyers and investors.

References

- [1] IFRS Foundation. IFRS 3 Business Combinations. <https://www.ifrs.org/issued-standards/list-of-standards/ifrs-3-business-combinations/>
- [2] IFRS Foundation. IFRS 13 Fair Value Measurement. <https://www.ifrs.org/issued-standards/list-of-standards/ifrs-13-fair-value-measurement/>
- [3] IFRS Foundation. IFRS 15 Revenue from Contracts with Customers. <https://www.ifrs.org/issued-standards/list-of-standards/ifrs-15-revenue-from-contracts-with-customers/>
- [4] IFRS Foundation. Educational material on measuring unquoted equity instruments within IFRS 13. <https://www.ifrs.org/content/dam/ifrs/supporting-implementation/ifrs-13/education-ifrs-13-eng.pdf>
- [5] IFRS Foundation. IAS 38 Intangible Assets. <https://www.ifrs.org/issued-standards/list-of-standards/ias-38-intangible-assets/>
- [6] IFRS Foundation. IAS 36 Impairment of Assets. <https://www.ifrs.org/issued-standards/list-of-standards/ias-36-impairment-of-assets/>
- [7] National Institute of Standards and Technology. AI Agent Standards Initiative. Updated 14 August 2026. <https://www.nist.gov/artificial-intelligence/ai-agent-standards-initiative>
- [8] National Institute of Standards and Technology. Announcing the AI Agent Standards Initiative. 17 February 2026. <https://www.nist.gov/news-events/news/2026/02/announcing-ai-agent-standards-initiative-interoperable-and-secure>
- [9] National Institute of Standards and Technology. Accelerating the Adoption of Software and AI Agent Identity and Authorization. 2026. <https://www.nccoe.nist.gov/projects/accelerating-adoption-software-and-ai-agent-identity-and-authorization>
- [10] National Institute of Standards and Technology. AI Agent Security Request for Information. 2026. <https://www.nist.gov/artificial-intelligence/ai-agent-standards-initiative>
- [11] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework. <https://www.nist.gov/itl/ai-risk-management-framework>
- [12] National Institute of Standards and Technology. Generative AI Profile NIST AI 600-1. <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf>
- [13] Model Evaluation and Threat Research. Task-Completion Time Horizons of Frontier AI Models. Updated 8 May 2026. <https://metr.org/time-horizons/>
- [14] Kwa T and others. Measuring AI Ability to Complete Long Tasks. 2025. <https://arxiv.org/abs/2503.14499>
- [15] Model Evaluation and Threat Research. How Does Time Horizon Vary Across Domains. 14 July 2025. <https://metr.org/blog/2025-07-14-how-does-time-horizon-vary-across-domains/>
- [16] OpenAI. Evaluate agent workflows. <https://developers.openai.com/api/docs/guides/agent-evals>

- [17] OpenAI. Trace grading for agents. <https://developers.openai.com/api/docs/guides/trace-grading>
- [18] OpenAI. Agents SDK tracing. <https://openai.github.io/openai-agents-python/tracing/>
- [19] International Valuation Standards Council. International Valuation Standards. <https://ivsc.org/standards/>
- [20] International Valuation Standards Council. Deciphering Technology. 28 June 2023. <https://ivsc.org/perspectives-paper-deciphering-technology/>
- [21] International Valuation Standards Council. Value and Data. 29 February 2024. <https://ivsc.org/perspectives-paper-value-and-data/>
- [22] International Valuation Standards Council. The Valuation Pulse AI and Technology in Valuation Sentiment Tracker. 8 May 2026. <https://ivsc.org/the-valuation-pulse-ai-technology-in-valuation-sentiment-tracker/>
- [23] OECD. OECD AI Principles. <https://oecd.ai/en/ai-principles>
- [24] OECD. Framework for the Classification of AI Systems. <https://oecd.ai/en/classification>
- [25] International Organization for Standardization. ISO IEC 42001 Artificial Intelligence Management Systems. <https://www.iso.org/standard/81230.html>
- [26] International Organization for Standardization. ISO IEC 23894 Artificial Intelligence Risk Management. <https://www.iso.org/standard/77304.html>
- [27] International Organization for Standardization. ISO IEC 5259 Data Quality for Analytics and Machine Learning. <https://www.iso.org/standard/81088.html>
- [28] International Organization for Standardization. ISO IEC 27001 Information Security Management Systems. <https://www.iso.org/standard/27001>
- [29] Cybersecurity and Infrastructure Security Agency. Secure by Design. <https://www.cisa.gov/securebydesign>
- [30] OWASP Foundation. Top 10 for Large Language Model Applications. <https://owasp.org/www-project-top-10-for-large-language-model-applications/>
- [31] OWASP Foundation. Agentic AI Threats and Mitigations. <https://genai.owasp.org/>
- [32] MITRE. ATLAS Adversarial Threat Landscape for AI Systems. <https://atlas.mitre.org/>
- [33] MLCommons. AI Safety Benchmark. <https://mlcommons.org/benchmarks/ai-safety/>
- [34] MLCommons. MLPerf Inference. <https://mlcommons.org/benchmarks/inference-datacenter/>
- [35] OpenAI. Frontier enterprise agent platform. <https://openai.com/business/frontier/>
- [36] OpenAI. Enterprise Signals. 2026. <https://openai.com/signals/enterprise-data/>
- [37] Anthropic. Economic Index report Cadences. 26 June 2026. <https://www.anthropic.com/research/economic-index-june-2026-report>
- [38] Anthropic. Responsible Scaling Policy. <https://www.anthropic.com/responsible-scaling-policy>
- [39] Google. Agents Companion security and governance guidance. <https://cloud.google.com/architecture/choose-design-pattern-agentic-ai-system>
- [40] Microsoft. Agentic AI architecture and responsible AI. <https://learn.microsoft.com/en-us/azure/architecture/ai-ml/guide/agentic-ai/agentic-ai-overview>
- [41] European Union. Regulation EU 2024 1689 Artificial Intelligence Act. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- [42] European Commission. AI Act implementation timeline. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- [43] United Kingdom Information Commissioner's Office. Guidance on AI and data protection. <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/>
- [44] United States Federal Trade Commission. Keep your AI claims in check. <https://www.ftc.gov/business-guidance/blog/2023/02/keep-your-ai-claims-check>
- [45] United States Securities and Exchange Commission. Artificial intelligence and investment fraud investor alert. <https://www.investor.gov/introduction-investing/general-resources/news-alerts/alerts-bulletins/investor-alerts/artificial-intelligence-fraud>
- [46] World Intellectual Property Organization. Artificial intelligence and intellectual property. <https://www.wipo.int/en/web/frontier-technologies/artificial-intelligence/index>
- [47] Linux Foundation. Model Context Protocol project. <https://www.linuxfoundation.org/press/linux-foundation-launches-new-agentic-ai-foundation>
- [48] MCPAgentBench. Real-world task benchmark for LLM agent tool use. 2025. <https://arxiv.org/abs/2512.24565>
- [49] OSWorld. Benchmarking multimodal agents for open-ended tasks in real computer environments. 2024. <https://arxiv.org/abs/2404.07972>
- [50] AgentChangeBench. Evaluation of goal-shift robustness in conversational AI. 2025. <https://arxiv.org/abs/2510.18170>

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 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.

<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.