MatchPoint Research Working Paper No. 2026/T19

The Economics of AI Adoption: Quantifying ROI in Financial Services

A confidence-weighted capital-allocation framework for family offices and GCC family businesses

Prepared by Matchpoint Partners
Author attribution pending CK approval
August 2026

ABSTRACT

Background. Financial institutions, family offices and owner-managed businesses are adopting AI rapidly, while reported returns use heterogeneous definitions and often omit realisation, control and integration costs. **Objective.** This paper develops a finance-specific method for selecting, measuring and funding AI adoption for A2 family-office CIOs and B4 GCC SME and family-business owners. **Approach.** The analysis reviews 40 primary, regulatory, standards, empirical and company sources available through 1 August 2026 and defines the accepted economic outcome as the value unit. **Findings.** Decision-grade ROI separates capacity, cash cost, collected contribution, expected loss and capital effects; includes full cost; and applies adoption, quality, realisation and attribution factors. **Implications.** Confidence-weighted NPV, explicit stage gates and a finance-owned value ledger convert AI adoption into controlled capital allocation. Attributed Matchpoint or client revenue, cost reduction and loss reduction remain USD 0 until approved observed evidence supports attribution.

Highlights

Unit. An accepted economic outcome passes the defined quality, risk, authority and business gates.

Benefits. Capacity, cash cost, collected contribution, expected loss and capital effects remain separate.

Costs. Integration, review, assurance, security, change, resilience and exit belong in total cost.

Evidence. External studies inform priors; local accepted outcomes determine capital allocation.

Economics. Confidence-weighted NPV and kill gates govern pilot, scale, redesign and retirement.

JEL Classification: C88, D24, G20, G23, M15, O31, O33

Keywords: artificial intelligence, AI adoption, return on investment, financial services, family office, GCC family business, productivity, value attribution, total cost of ownership, NPV, model risk, AI governance

1. INTRODUCTION

Financial institutions, investment offices and finance teams are moving from isolated artificial-intelligence experiments to portfolios of tools embedded in operating work. The adoption evidence is clear; the economic evidence remains uneven. In the Bank of England and Financial Conduct Authority's 2024 survey, 75% of 118 responding financial-services firms reported current AI use and another 10% planned adoption within three years. The median reported use-case count was expected to rise from nine to 21 [1]. Those figures establish diffusion. They do not establish return on investment.

The distinction matters because an AI output is not automatically an economic benefit. Time may be released without payroll falling. A recommendation may be faster without improving a decision. A chatbot may deflect contacts while creating rework elsewhere. A model may reduce false positives but require expensive monitoring, validation and data remediation. An individual may adopt a tool while the surrounding process, controls and incentives remain unchanged. Economic measurement must follow the whole operating system from source data to an accepted business outcome.

This paper develops a finance-specific method for quantifying that return. Its primary audience is A2 family-office CIOs and heads of alternatives. Its secondary audience is B4 GCC SME and family-business owners. A2 teams allocate capital, supervise managers, aggregate portfolio information and operate with fiduciary and governance duties. B4 leaders manage finance capacity, working capital, customer and supplier processes, succession and growth with smaller teams. Both need a practical answer to the same question: which AI adoption programme creates credible economic value after full cost, quality, risk and realisation are measured?

Current primary evidence supports disciplined optimism. DBS reported approximately SGD 1 billion of 2025 economic value from more than 2,000 AI and machine-learning models and more than 430 use cases [9]. Bank of America reported that approximately 18,000 technology employees using AI for coding saved more than 20% of core work, while nearly 150,000 active users generated more than 1.5 million prompts per week [10]. JPMorganChase reported an 80% reduction in time from manual research to insight for one asset-management capability, and separately reported more than double the transaction-screening volume with half the manual operator checks [11-12]. Its consumer-bank unit also reported a nearly 60% year-on-year increase in value from AI and machine learning [13]. These company disclosures describe their reporting settings; they are not benchmarks that a family office or GCC family business can import without local measurement.

Experimental evidence also requires narrow interpretation. Brynjolfsson, Li and Raymond found a 14% average productivity increase for customer-support agents using a generative-AI assistant, with larger gains among less experienced workers [19]. A randomised METR study found that experienced developers working on familiar open-source repositories took 19% longer with early-2025 AI tools [21]. The later METR update reported that selection effects impaired a follow-on estimate [22]. The productivity frontier is task-, user-, workflow- and time-specific.

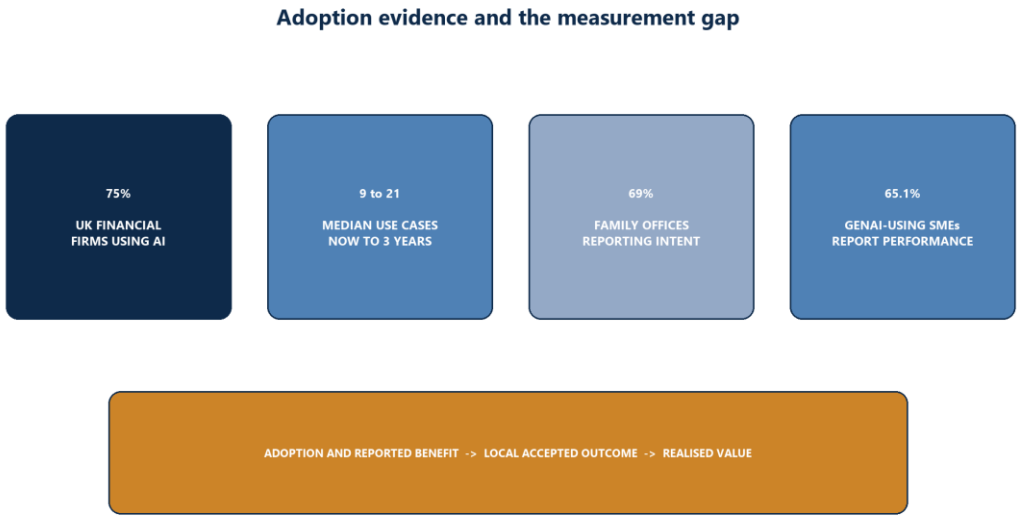

Figure 1. Adoption evidence and the measurement gap

Source: Bank of England and FCA [1]; UBS [14]; OECD [16]. Each percentage retains its source population and definition.

The central unit in this paper is the **accepted economic outcome**. An outcome is accepted when the defined work is complete, quality and risk gates pass, the authorised owner records acceptance and the benefit can be linked to a financial

or operating measure. Generated words, prompts, model calls, demonstrations, licence seats and hours notionally saved are intermediate quantities. They can explain adoption and cost; they do not prove realised value.

The framework advances six conclusions. First, build a use-case business case around an accepted outcome and a comparable baseline. Second, separate capacity released, cash cost removed, revenue collected, losses avoided and capital or working-capital effects. Third, include integration, review, assurance, security, change and exit costs in total cost of ownership. Fourth, apply adoption, quality, realisation and attribution factors before recognising value. Fifth, use a confidence-weighted NPV and explicit stop conditions to allocate capital. Sixth, keep attributed Matchpoint or client revenue, cost reduction and loss reduction at USD 0 until approved observed evidence supports attribution.

Decision	Required evidence	Initial authority
Fund an AI pilot	Baseline, task boundary, owner, cost range, risk tier and success threshold	Business sponsor and finance owner
Move from pilot to production	Representative evaluation, control evidence, user adoption and full-cost forecast	Process, technology and risk owners
Recognise productivity value	Comparable accepted outcomes and observed end-to-end effort	Finance and process owner
Recognise cash cost reduction	Approved and observed change in paid cost	CFO or delegated budget owner
Recognise revenue value	Incremental collected contribution with an attribution method	Commercial and finance owners
Recognise loss reduction	Credible counterfactual, exposure and statistical confidence	Risk and finance owners
Scale or retire	Risk-adjusted NPV, operating evidence and tested exit	Investment committee or equivalent

2. SCOPE, DEFINITIONS AND EVIDENCE BOUNDARIES

2.1 Artificial intelligence and adoption

AI adoption in this paper means purposeful use of a model-enabled capability inside an owned business process. It includes predictive machine learning, generative AI, document intelligence, copilots and bounded tool-using agents. A purchased licence without qualified use is availability. An employee prompt is activity. A completed workflow with accepted output is adoption. A financially realised change is value.

The distinction prevents four common category errors. Seat activation is not sustained use. Sustained use is not process change. Process change is not accepted quality. Accepted quality is not necessarily cash or collected revenue. A measurement system should record all four transitions because a stalled transition explains why technically capable tools fail to produce financial return.

Stage	Evidence	Economic status
Available	Approved tool, licence and access	Cost incurred
Active	Qualified user activity for defined work	Adoption signal
Embedded	Workflow, role and control change	Operating capability
Accepted	Output passes quality and authority gates	Candidate benefit
Realised	Approved financial or redeployment consequence observed	Recognised economic value

2.2 Return on investment

ROI is often quoted as a single ratio. A finance-grade case requires a family of measures. Net annual value compares annual realised benefits with annual recurring costs. Simple ROI divides net value by investment. Payback measures the time until cumulative net cash or approved economic benefits recover initial investment. NPV discounts future net cash flows and exit costs. Internal rate of return can supplement NPV when cash-flow patterns permit it. Each measure should use the same benefit definitions, cost perimeter, time horizon and risk adjustments.

Economic ROI and financial-statement accounting are separate. IAS 38 sets criteria for recognising and measuring intangible assets; research expenditure is expensed, while development expenditure meeting specified criteria is recognised as an intangible asset [40]. Management's economic model may include option value, released capacity and avoided delay. Qualified accountants determine the applicable accounting treatment. The paper does not prescribe capitalisation, amortisation or impairment conclusions.

2.3 Benefit taxonomy

The benefit ledger separates five classes because each requires different evidence.

Benefit class	Definition	Minimum evidence	Recognition boundary
Capacity released	Human time no longer required for the same accepted output	Comparable touch-time study and stable quality	Report separately until redeployment is approved and observed
Cash cost removed	Paid labour, vendor or operating spend ceases	Budget, invoice, payroll or contract evidence	Recognise after the cost change occurs
Revenue contribution	Incremental collected gross contribution caused by the capability	Test/control or credible attribution and collection evidence	Recognise contribution, not bookings or pipeline
Loss reduction	Reduction in fraud, credit, error or operational loss relative to a counterfactual	Exposure, loss definition, control group or validated model	Use confidence bounds and prevent double counting
Capital or working-capital effect	Change in funding, liquidity, inventory, receivables or capital use	Reconciled balances, causal operating change and finance approval	Value using the applicable funding or capital method

Customer satisfaction, decision speed, employee experience, resilience and compliance quality are important outcomes. Their monetary translation requires a documented bridge. A faster close can improve decision time; it does not become revenue unless an observed commercial mechanism and attribution exist. Lower error rates can reduce expected loss; they do not become cash savings without an appropriate loss baseline.

2.4 Cost taxonomy

Total cost of ownership includes initial and recurring components. Licence or model-inference spend is often visible. The material omissions are data readiness, integration, process redesign, evaluation, human review, cybersecurity, model risk, legal and privacy review, training, support, incident response and exit.

Cost layer	Initial examples	Recurring examples
Discovery and design	Process mapping, baseline, control and architecture design	Use-case portfolio review
Data	Ownership, remediation, labelling and migration	Quality monitoring, retention and lineage
Technology	Integration, retrieval, interfaces and environments	Licences, inference, hosting and observability
People	Training, workflow redesign and role changes	Review, exceptions, support and product ownership
Risk and assurance	Impact assessment, validation, threat modelling and legal review	Monitoring, testing, audit and regulatory change
Resilience and exit	Fallback, portability and termination design	Recovery tests, backups and exit reserve

2.5 Evidence classes

External research helps define priors and controls. Internal evidence determines whether a named programme creates value. Company disclosures are attributed to the reporting company and retain their reported scope. Survey responses describe respondents. Management scenarios demonstrate arithmetic only.

Class	Description	Permitted use
P1	Law, binding regulation or supervisory rule	Define an applicable obligation after scope review
P2	Official regulator, standard or public framework	Design governance, risk and measurement questions
P3	Primary empirical research	Bound an effect to its study design and population
P4	Issuer filing, annual report or company operating disclosure	Record what the company reported in its setting
I1	Reconciled internal financial or operating record	Measure a local baseline or realised result
I2	Approved evaluation, acceptance, cost and benefit record	Support a local investment decision
U	Unverified illustrative management assumption	Demonstrate formulae; never represent observed performance

2.6 Scope and professional boundaries

The paper covers business-case design, process measurement, financial modelling, portfolio selection, technology architecture, controls, procurement and adoption. It does not provide investment, legal, regulatory, accounting, audit, tax, privacy, cybersecurity, employment or valuation advice. Applicable obligations and professional judgements remain with qualified owners. Named-person authorship remains pending CK approval.

3. A2 AND B4 DECISION MAP

3.1 A2 family-office CIOs and heads of alternatives

Family offices can combine sophisticated investment portfolios with relatively compact teams. UBS's 2025 survey of 317 single family offices reported that 69% expected to use AI for financial reporting or data visualisation, 64% for text analysis and 62% for portfolio analysis over the following five years; 6% did not expect to use AI [14]. The survey describes intentions within its client sample. It does not establish realised ROI. UBS's 2026 survey of 307 family offices reported that 68% had formal financial-performance measurement processes, 60% used investment committees and 35% had a defined succession plan for the family office [15]. These governance characteristics provide useful adoption context; they do not measure AI returns.

A2 decision-makers should treat AI adoption as both an operating-capability investment and an investment-governance question. The office needs reliable information, clear delegation, data rights, confidentiality, service-provider control and continuity. The economic model should value scarce professional capacity while preserving the judgement that differentiates the office.

A2 decision	Candidate AI contribution	Accepted outcome	Principal value class
Is the portfolio pack ready?	Document extraction, normalisation and narrative draft	Reconciled, reviewed pack	Capacity and decision speed
Which manager reports need attention?	Evidence retrieval and exception triage	Owned exception queue	Capacity and risk information
Is a capital call consistent with terms?	Notice extraction and document comparison	Reviewed call packet	Capacity and loss prevention
How should the investment committee prepare?	Research retrieval and briefing assembly	Source-backed briefing	Professional capacity
Are exposures within policy?	Classification support and variance explanation	Deterministically calculated exposure pack	Risk and governance
Can the office scale without immediate hiring?	Workflow redesign and knowledge retrieval	Sustained accepted throughput	Redeployed or deferred capacity

The investment case should never count a highly paid professional's entire hourly cost whenever a task becomes faster. The value depends on what happens to the released time. If the team performs more manager diligence, improves client or principal service, defers an approved hire or reduces paid external work, the programme may realise value. If the time is absorbed without an agreed use, the result is capacity released and remains separate from cash benefit.

3.2 B4 GCC SME and family-business owners

OECD evidence shows the SME adoption gap and the danger of overstating benefits. In 2025, 20.2% of firms across available OECD countries reported AI use, including 17.4% of small firms and 52.0% of large firms [18]. In a survey of 5,232 SMEs across seven countries, 65.1% of generative-AI users reported improved employee performance, while the survey did not measure the magnitude of that improvement [16]. The evidence is informative for B4 leaders and is not UAE-specific.

B4 economics are often constrained by fragmented data, key-person dependence and limited change capacity. A low licence price can be overwhelmed by integration and management time. A compact team can also benefit materially from well-scoped tools because one repetitive process may consume a large share of scarce finance or commercial capacity.

B4 decision	Candidate AI contribution	Accepted outcome	Principal value class
Which receivables need action?	Evidence assembly and prioritised draft follow-up	Approved collection queue	Capacity and working capital
Can management reporting close earlier?	Variance explanation and pack assembly	Reviewed monthly pack	Capacity and decision speed

B4 decision	Candidate AI contribution	Accepted outcome	Principal value class
Which supplier or customer documents are incomplete?	Document classification and exception detection	Owned exception list	Capacity and loss avoidance
Can customer enquiries scale?	Assisted response and knowledge retrieval	Accepted response at required service level	Capacity and revenue retention
Which leads deserve sales attention?	Research, enrichment and next-action support	Qualified opportunity accepted by owner	Revenue contribution
Can a role or external service be deferred?	Cross-workflow automation with measured load	Sustained throughput and service quality	Deferred cash cost

3.3 Shared decision contract

Both ICPs should fix ten fields before spending materially: outcome, owner, population, baseline, benefit class, cost perimeter, quality gate, risk tier, evaluation design and stop condition. Missing fields create a business case that can expand after the fact and absorb unfavourable results.

Question	Weak answer	Decision-grade answer
Outcome	Improve productivity	Reduce end-to-end touch time per accepted monthly pack with unchanged quality
Baseline	The team is busy	Representative work population, touch time, cycle time, quality and cost
Benefit	Save hours	Capacity, cash, revenue, loss or capital benefit separately defined
Cost	Licence price	Initial and recurring full-cost ledger with owner
Quality	Users like it	Acceptance, error, rework, exceptions and material-harm thresholds
Attribution	AI helped	Test/control, phased rollout or documented contribution method
Realisation	Time saved	Approved redeployment, avoided spend or collected contribution
Risk	Enterprise vendor	Data, model, cyber, conduct, third-party and resilience assessment
Scale	Many possible use cases	Portfolio dependency, shared platform and marginal cost
Stop	Review later	Dated threshold that pauses, redesigns or retires the use case

4. MARKET, ADOPTION AND ECONOMIC EVIDENCE

4.1 Financial-services adoption

The Bank of England and FCA survey provides a useful regulated-finance baseline. Seventy-five per cent of respondents reported current AI use; 85% were using or planning to use it. Respondents expected the median use-case count to rise from nine to 21. Large UK and international banks reported much larger medians. The most highly rated current benefits were data and analytical insight, anti-money-laundering and fraud work, and cybersecurity. Operational efficiency, productivity and cost base were expected to show the largest increase in benefit over three years [1].

The survey also reports the control perimeter. Eighty-four per cent of respondents had an accountable person for their AI framework. Forty-six per cent of firms using or planning AI reported only partial understanding of implemented technologies. The top three providers accounted for 73% of named cloud providers and 44% of named model providers [1]. Data privacy, quality and security ranked among the largest current risks. These costs and dependencies belong in the economic model.

The FSB's 2024 report identifies operational efficiency, compliance, product customisation and analytics as potential benefits, alongside third-party concentration, market correlation, cyber, model, data-quality and governance vulnerabilities [3]. Its June 2026 sound-practices report was a consultation at the publication date of this paper and should be described as proposed consultation material [4].

4.2 What large financial institutions reported

Company disclosures illustrate several value mechanisms. They use different definitions and cannot be compared as a league table.

Reporting entity	Company-reported evidence	Measurement implication
DBS [9]	About SGD 1bn of 2025 economic value; 2,000-plus models and 430-plus use cases	Portfolio scale and value governance can matter more than one tool
DBS [9]	CodeBuddy time savings up to 20% on certain coding tasks; customer satisfaction rose 23% for DBS Joy users	Keep task and user population attached to every percentage
Bank of America [10]	Nearly 150,000 active users; 1.5m prompts weekly; coding users saved more than 20% of core work	Activity and task-time metrics should remain distinct from cash value
JPMorgan AWM [11]	80% reduction in time from manual research to insight for SpectrumIQ	Time-to-insight can be a valid capacity metric after output acceptance
JPMorgan CIB [12]	More than double transaction-screening volume with half manual checks	Throughput, control effort and quality can be measured together
JPMorgan CCB [13]	Nearly 60% year-on-year increase in value from AI/ML; 40% gross operations productivity objective by 2030	Management targets require later observed evidence and cost context

The largest organisations can spread platform, data, risk and change costs across many use cases. A smaller office may face higher cost per use case while benefiting from simpler architecture and faster decisions. The relevant benchmark is the marginal value and cost of the named local workflow, including shared platform economics.

4.3 Experimental and cross-sector evidence

The NBER customer-support study of 5,179 agents found an average 14% productivity increase, with gains concentrated among less experienced workers [19]. The treatment combined a model with a specific workflow, knowledge base and interaction environment. Transfer to investment research, financial reporting, compliance or owner-managed operations requires a local evaluation.

Stanford HAI's 2026 AI Index reports broad adoption and rapid technical progress while documenting uneven performance across tasks and benchmarks [23]. METR's 2025 developer experiment recorded a 19% slowdown among experienced participants in its setting [21]. Its 2026 update explained why selection effects impaired a later estimate [22]. These sources support an empirical rule: use an external effect size only as a prior for scenario design. Use local accepted outcomes to decide capital allocation.

OECD SME evidence reinforces the distinction between perceived benefit and measured magnitude. Sixty-five point one per cent of surveyed generative-AI users reported improved performance; the question did not measure how large the improvement was [16]. A separate OECD report links AI adoption to digital maturity and complementary organisational investments [17]. Tool access without process, data and skills investment can create a low-cost demonstration and a high-cost production gap.

4.4 The evidence-to-business-case bridge

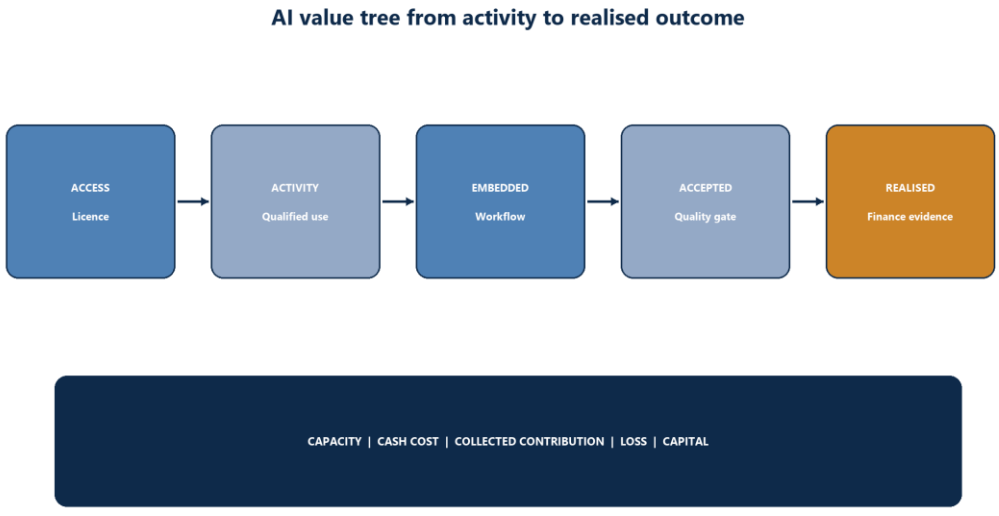

Figure 2. AI value tree from activity to realised economic outcome

Source: Matchpoint AI economics framework.

External evidence should inform four inputs: plausible use cases, measurement design, control requirements and scenario range. It should not fill local baseline, adoption, quality, realisation or attribution data. A bank disclosure can demonstrate that a mechanism has existed in one institution. A controlled study can establish an effect under a study design. Neither proves local cash savings.

External evidence	Appropriate use	Inappropriate use
Adoption survey	Prioritise common use cases and constraints	Claim local maturity or return
Company annual report	Understand reported mechanism and scale	Import the percentage as a forecast
Randomised study	Choose evaluation design and scenario range	Apply effect outside the studied task
Regulation or standard	Define required controls and evidence	Treat compliance as value without cost
Vendor documentation	Describe capability and limits	Treat capability as operating effectiveness

5. MEASUREMENT CONTRACT AND VALUE TREE

5.1 Use-case business case

Every use case receives a versioned business-case record. The record prevents portfolio summaries from combining incompatible definitions and allows later estimates to be compared with actual outcomes.

Field	Required content
Identity	Use-case ID, business process, version, sponsor and owner
Outcome	Accepted business output and decision supported
Population	Transactions, cases, users, entities, periods and exclusions
Baseline	Volume, touch time, cycle time, quality, loss and cost
Intervention	Model, tools, data, workflow and role changes
Benefits	Capacity, cash, revenue, loss and capital definitions
Costs	Initial, recurring, shared, allocated and exit costs
Quality	Acceptance, error, rework and material-harm metrics
Risk	Tier, controls, residual risk and owner
Evaluation	Design, sample, period, comparator and thresholds
Finance	Scenario, discount rate, NPV, payback and confidence
Decision	Fund, hold, scale, redesign or retire; date and authority

5.2 Accepted economic outcome

The denominator should include every attempted item in the defined population. A generated draft that is abandoned still consumes compute and attention. A failed run that is manually completed belongs in assisted-process cost. A corrected output is not first-pass accepted. An output delivered faster but after the business deadline may have no value.

Accepted outcome rate equals accepted outcomes divided by all eligible attempted outcomes. First-pass acceptance equals accepted outcomes requiring no material correction divided by reviewed outcomes. End-to-end touch time includes preparation, prompting, waiting, review, correction, exception handling, support and release. Cycle time runs from the defined business trigger to acceptance.

Metric	Numerator	Denominator	Control
Adoption	Qualified active users or workflows	Eligible users or workflows	Fixed activity definition
Accepted outcome rate	Accepted outputs	Eligible attempted outputs	Include failures and abandonments
First-pass acceptance	Accepted without material correction	Reviewed outputs	Fixed correction taxonomy
Productivity	Accepted outcomes	Total human hours	Comparable population and quality
Unit cost	Full process cost	Accepted outcomes	Include review, failure and support

Metric	Numerator	Denominator	Control
Timeliness	Outputs accepted within SLA	Eligible outputs	Fixed trigger and deadline
Material defect	Outputs with defined material error	Reviewed outputs	Independent review or sampling

5.3 Value recognition ladder

Capacity becomes economic value through a documented realisation path. The realisation factor is the observed proportion of released capacity that creates an approved consequence. It can include redeployed work that management values, an avoided hire tied to an approved staffing plan, reduced external spend or reduced overtime. Each type is reported separately.

Revenue follows a stricter ladder: AI-assisted action, qualified opportunity, customer decision, contracted revenue, collected revenue and gross contribution. Attribution can use randomisation, matched cohorts, phased rollout, difference-in-differences or a conservative contribution rule. Pipeline, bookings and collected contribution are separate measures.

Loss reduction requires an exposure, event definition and counterfactual. A lower false-positive rate can reduce review cost. A higher recall can reduce missed cases. A changed approval rate can affect revenue and credit loss simultaneously. Double counting is controlled through one benefit ledger and one owner.

5.4 Quality and risk as economic variables

Quality affects both numerator and denominator. A productivity gain that increases material error can create negative economic value. Risk controls also consume resources and reduce expected loss. The model should record gross operating value, control cost and residual expected loss separately.

Risk-adjusted annual value equals approved annual benefits minus recurring operating cost minus control and assurance cost minus residual expected loss. The residual loss estimate should use a documented probability, exposure and confidence range. Where estimation is weak, management can use a conservative reserve or keep the value at zero.

5.5 Value ledger

Ledger field	Rule
Baseline period	Fixed before intervention results are reviewed
Benefit owner	Named finance or business owner
Evidence class	P, I or U attached to every material input
Gross and net	Gross benefit displayed before all deductions
Capacity and cash	Never combined without an approved realisation rule
Revenue and contribution	Recognise collected gross contribution for ROI
Loss reduction	Use one counterfactual and confidence range
Shared cost	Allocate by documented driver and show portfolio total
Exit provision	Include data extraction, transition and termination
Attribution	Attributed Matchpoint or client value remains USD 0 without approved evidence

6. PORTFOLIO SELECTION AND STAGING

6.1 Use-case inventory

The strongest initial use cases combine material workload, stable scope, accessible data, reversible action and measurable quality. A use case can be economically attractive and operationally unready. It should then enter data or process remediation rather than an AI build.

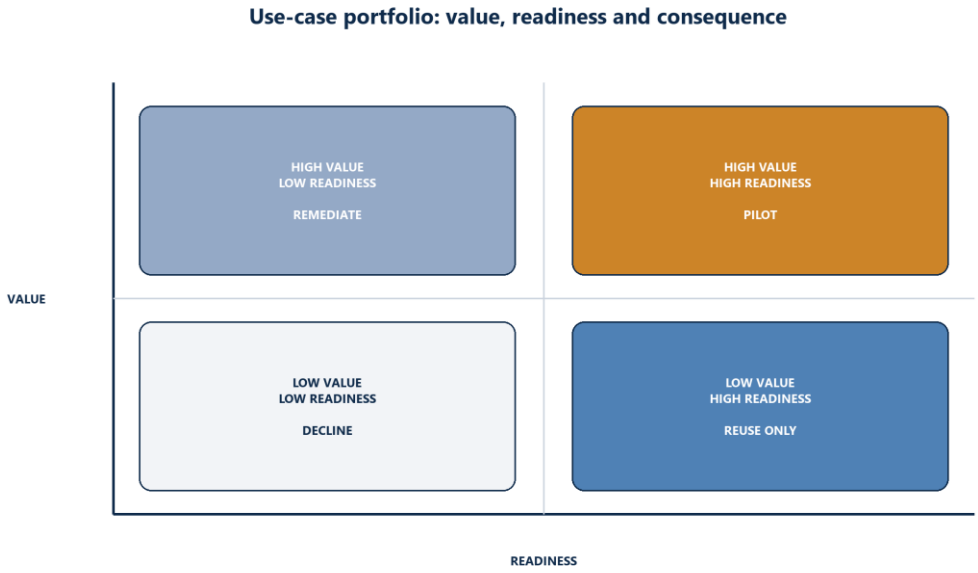

Figure 3. Use-case portfolio by value, readiness and consequence

Source: Matchpoint capital-allocation framework.

Dimension	Low score	High score
Addressable workload	Infrequent, low-cost activity	High-volume or scarce-professional activity
Outcome value	Convenience only	Cash, contribution, risk or strategic capacity
Data readiness	Unowned, fragmented and inaccessible	Owned, complete and controlled
Process stability	Policy and exceptions unresolved	Bounded task and known exception classes
Quality measurability	Subjective approval only	Observable acceptance and defect measures
Consequence	Irreversible or customer-harming action	Draft, recommendation or reversible preparation
Integration effort	Many legacy writes and custom interfaces	Approved reads and narrow tools
Adoption readiness	No owner, incentive or training	Named owner, workflow fit and user demand

6.2 Portfolio scoring

A three-axis score separates attractiveness, readiness and consequence. Attractiveness covers gross value and strategic option. Readiness covers data, process, skills and integration. Consequence covers financial, conduct, privacy, cyber and operational harm. A high-value, high-consequence use case may remain investable after enhanced control; it should not be allowed to outrank safer work solely on gross value.

Portfolio quadrant	Action
High value, high readiness, bounded consequence	Pilot with defined outcome and scale path
High value, low readiness	Fund remediation milestone before model build
Low value, high readiness	Use only if it proves a reusable platform component
Low value, low readiness	Decline or retire
High consequence at any value	Require independent risk case and restricted authority

6.3 Build, buy or configure

The choice is an economic architecture decision. Buying can reduce initial engineering and increase third-party dependency. Building can improve control and differentiation while increasing specialist cost and key-person risk. Configuring an enterprise platform can reuse security, identity and support while constraining model choice. The business case should include transition and exit.

Factor	Build	Buy	Configure platform
Differentiation	Highest potential	Depends on product	Moderate
Initial time	Usually longer	Usually shorter	Moderate
Data control	Design-dependent	Contract-dependent	Enterprise-policy dependent
Model portability	Can be designed	May be restricted	Platform-dependent
Specialist burden	High	Vendor-management burden	Shared enterprise burden
Exit cost	Code, people and infrastructure	Data and contract transition	Platform transition

6.4 Stage gates

The capital sequence is discovery, baseline, prototype, controlled pilot, shadow production, restricted production and portfolio scale. Each stage has a maximum spend, evidence requirement and termination rule. The gate prevents sunk cost from replacing evidence.

Gate	Evidence	Decision
G0 Problem	Outcome, owner and population exist	Baseline or stop
G1 Baseline	Volume, quality, effort and cost measured	Design or remediate
G2 Prototype	Capability and data path demonstrated	Pilot or stop
G3 Pilot	Representative acceptance, time and risk thresholds pass	Shadow or redesign
G4 Shadow	End-to-end operation and controls stable	Restricted production
G5 Production	Realisation and unit economics meet threshold	Scale, hold or retire
G6 Portfolio	Shared platform value exceeds shared cost	Expand portfolio

7. BASELINE AND CAUSAL MEASUREMENT

7.1 Baseline design

The baseline should represent the same accepted outcome, population and quality bar as the assisted workflow. It should cover routine and difficult cases, peak and normal periods, new and experienced users, and the full exception distribution. A short baseline may understate rare but costly events.

Baseline component	Measurement
Volume	Eligible items, completed items and backlog
Effort	Preparation, review, correction, exceptions and support
Quality	First-pass acceptance, defect severity and downstream corrections
Time	Trigger-to-acceptance cycle and SLA attainment
Cost	Loaded labour, vendor, system, control and loss cost
Outcome	Customer, investment, finance or risk measure supported
Context	Season, user tenure, product, entity and complexity

7.2 Evaluation designs

Randomised assignment offers strong causal evidence when operationally feasible. Phased rollout with matched comparison groups can be practical. An interrupted time series can help when a stable history exists. Pre/post comparison without controls is vulnerable to demand, staffing, seasonality, learning and selection changes.

Design	Strength	Principal limitation
Randomised controlled rollout	Strong causal attribution	Contamination and operational acceptability
Matched cohort	Practical for teams or cases	Residual selection difference

Design	Strength	Principal limitation
Difference-in-differences	Controls common time trends	Parallel-trend assumption
Stepped-wedge rollout	Every group eventually receives tool	Calendar and learning effects
Interrupted time series	Uses operational history	Concurrent changes
Pre/post only	Fast descriptive signal	Weak causal attribution

7.3 Productivity measurement

The Bank of England and FCA's 2023 feedback statement gives use-case-specific examples. Payment-matching benefits can be measured through processing volume, error reduction and customer satisfaction. AML applications can use precision and recall, alongside consumer, model and data metrics [2]. These examples support a balanced scorecard rather than a single speed metric.

Human hours per accepted outcome equals total eligible human touch time divided by accepted outcomes. Unit cost adds allocated technology, data, support, control and residual loss cost. The assisted workflow should be compared with a baseline at the same quality threshold. If quality improves, the model can separately value reduced rework or loss.

7.4 Revenue and relationship outcomes

Front-office AI can affect research coverage, meeting preparation, contact cadence, personalisation, pricing and service. JPMorganChase reported 1 million personalised AI-driven insights to front-office users through Connect Coach [11]. Activity should be followed through qualified action, customer response, conversion, collection and contribution.

Stage	Metric	Economic recognition
Suggestion	AI-generated insight	None
User action	Accepted next action	Adoption evidence
Customer response	Meeting, click or reply	Engagement evidence
Opportunity	Qualified pipeline	Forecast only
Contract	Signed revenue	Contracted amount, not yet collected
Collection	Cash received	Candidate revenue value
Contribution	Collected revenue less variable cost	ROI benefit after attribution

7.5 Risk and loss outcomes

Transaction screening, fraud detection, credit assessment and compliance triage require precision, recall, false-positive cost, false-negative exposure and review effort. JPMorganChase's reported transaction-screening result illustrates a combined throughput and manual-review measure [12]. A local model also needs downstream loss, alert quality, customer impact and control evidence.

The current US interagency model-risk guidance, SR 26-2, superseded SR 11-7 in April 2026. It emphasises a risk-based approach tailored to the bank's model profile, size and complexity, with effective development, validation, governance and control [27]. It is most relevant to specified US banking organisations; the principles can inform diligence without being represented as directly applicable to every A2 or B4 organisation.

8. FINANCIAL MODEL

8.1 Core equations

Gross annual capacity value equals eligible annual volume multiplied by baseline hours minus assisted hours per accepted outcome, multiplied by the loaded hourly value. Candidate realised capacity value multiplies gross capacity by the approved realisation factor. Cash cost, collected contribution, expected-loss and capital benefits are calculated separately.

Net annual value equals approved benefits minus recurring technology, data, people, control, assurance, resilience and exit-reserve costs. NPV equals initial investment as a negative cash flow plus discounted net annual values and the discounted terminal exit cost. Confidence-weighted NPV applies scenario or evidence probabilities approved before the investment decision.

Symbol	Definition
Q	Eligible accepted outcomes per year
Hb	Baseline human hours per accepted outcome
Ha	Assisted human hours per accepted outcome
W	Loaded hourly value
R	Approved realisation factor
C0	Initial implementation cost
Ct	Recurring full cost in year t
Bt	Approved total benefit in year t
r	Discount rate
X	Terminal transition or exit cost

Gross capacity value = Q x (Hb - Ha) x W.

Candidate realised capacity value = gross capacity value x R.

Net value in year t = Bt - Ct.

NPV = -C0 + sum from t equals 1 to T of (Bt - Ct) / (1 + r)^t - X / (1 + r)^T.

8.2 Adoption, quality, realisation and attribution factors

An estimate based on technical task time should pass four factors. Adoption is the eligible share actually using the workflow. Quality is the accepted share at the required threshold. Realisation is the share of released capacity or opportunity that produces an approved consequence. Attribution is the share credibly caused by the intervention.

Adjusted benefit equals theoretical gross benefit multiplied by adoption, quality, realisation and attribution factors. For some benefits, one factor may already be embedded in the measurement denominator. The model should document that choice and prevent double haircutting.

8.3 Cost allocation

Shared platform costs can make the first use case unattractive and the portfolio attractive. The investment committee should see both standalone and portfolio views. Initial platform cost is allocated only for decision support; the cash-flow statement retains total spend. Marginal use-case cost includes incremental data, integration, evaluation, support and compute.

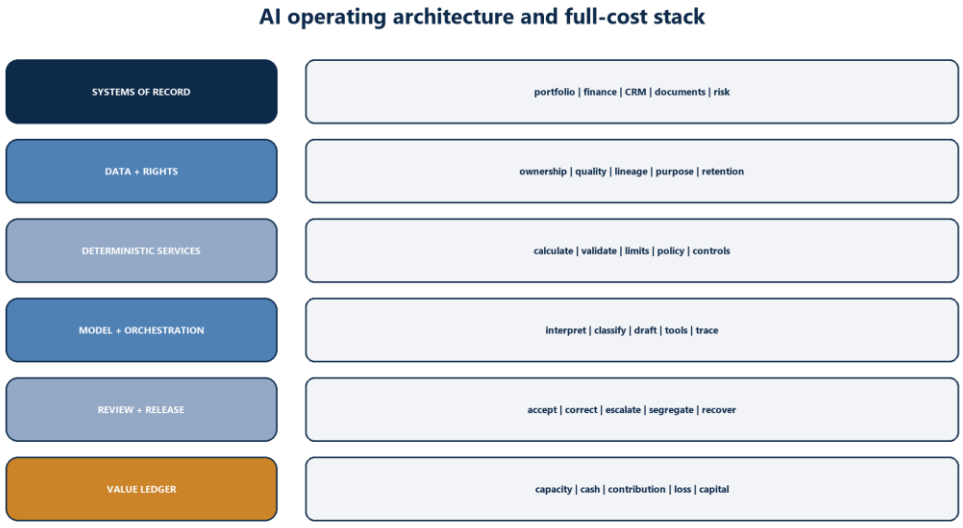

Figure 4. AI operating architecture and full-cost stack

Source: Matchpoint architecture; risk, security and agent context [3-7,24-31,38-39].

View	Purpose	Rule
Standalone	Decide whether the use case works independently	Include all required dedicated cost
Marginal	Decide whether to add a use case to an existing platform	Include truly incremental cost
Portfolio	Decide whether the combined programme creates value	Include all shared and use-case costs once
Cash	Budget and liquidity planning	Use contracted and expected cash flows
Economic	Include approved capacity and risk value	Display non-cash benefits separately

8.4 Confidence-weighted scenarios

Each scenario should vary no fewer than adoption, accepted time saving, realisation, recurring cost and implementation delay. Revenue or loss benefits require their own uncertainty. Probabilities are management estimates and are recorded before outcomes are known.

Scenario	Evidence posture	Capital implication
Downside	Lower adoption and quality; higher integration and review	Test survival and exit cost
Base	Observed pilot central estimate with conservative realisation	Primary allocation case
Upside	Sustained scale and reuse supported by evidence	Optional value; avoid funding solely on upside

8.5 Kill gates

A use case pauses when a material defect threshold is exceeded, data rights are uncertain, security control fails, unit cost exceeds the approved ceiling, user adoption remains below the dated threshold, realised value fails to cover the next funding stage or a vendor change invalidates evaluation evidence. Retirement is a capital-allocation outcome, not a project failure.

Trigger	Immediate action	Evidence to resume
Material quality breach	Stop release and preserve evidence	Root cause, correction and regression pass
Unauthorised data or action	Isolate access and invoke incident process	Rights, scope and security approval
Negative marginal value	Freeze expansion	Revised cost or observed benefit
Adoption below threshold	Diagnose workflow and incentives	Sustained qualified use
Control cost exceeds value	Redesign consequence or authority	Lower-risk design and new case
Exit or provider risk changes	Test portability and fallback	Approved transition evidence

9. ILLUSTRATIVE A2 FAMILY-OFFICE CASE

9.1 Scope and evidence boundary

The case models an AI-assisted portfolio-reporting workflow for an A2 family office. The accepted outcome is a reviewed reporting packet containing source lineage, reconciled metrics, approved definitions, exceptions and an accepted narrative. Investment judgement, valuation approval and release authority remain with named professionals.

[Unverified] Every numerical input in Sections 9 and 10 is an illustrative management assumption created solely to demonstrate the framework. The figures are not observed Matchpoint results, client results, market benchmarks or forecasts. Attributed revenue, cash cost reduction and loss reduction equal USD 0.

9.2 Illustrative baseline and intervention

Input	Base assumption	Evidence class
Accepted packets per year	600	U; unverified illustrative management assumption
Baseline human time per packet	2.60 hours	U; unverified illustrative management assumption
Assisted human time per packet	1.55 hours	U; unverified illustrative management assumption
Loaded capacity value	USD 120 per hour	U; unverified illustrative management assumption

Input	Base assumption	Evidence class
Capacity realisation factor	50%	U; unverified illustrative management assumption
Initial implementation cost	USD 30,000	U; unverified illustrative management assumption
Annual recurring full cost	USD 18,000	U; unverified illustrative management assumption
Year-one adoption ramp	60% of base benefit	U; unverified illustrative management assumption
Discount rate	12%	U; unverified illustrative management assumption
Year-three exit provision	USD 5,000	U; unverified illustrative management assumption

[Unverified] The baseline process uses 1,560 human hours each year; the assisted process uses 930 hours. The gross 630 hours of released capacity at USD 120 per hour produces USD 75,600 of candidate gross capacity value. Applying the illustrative 50% realisation factor produces USD 37,800 of candidate annual realised capacity value. This remains an economic scenario and does not constitute cash cost reduction.

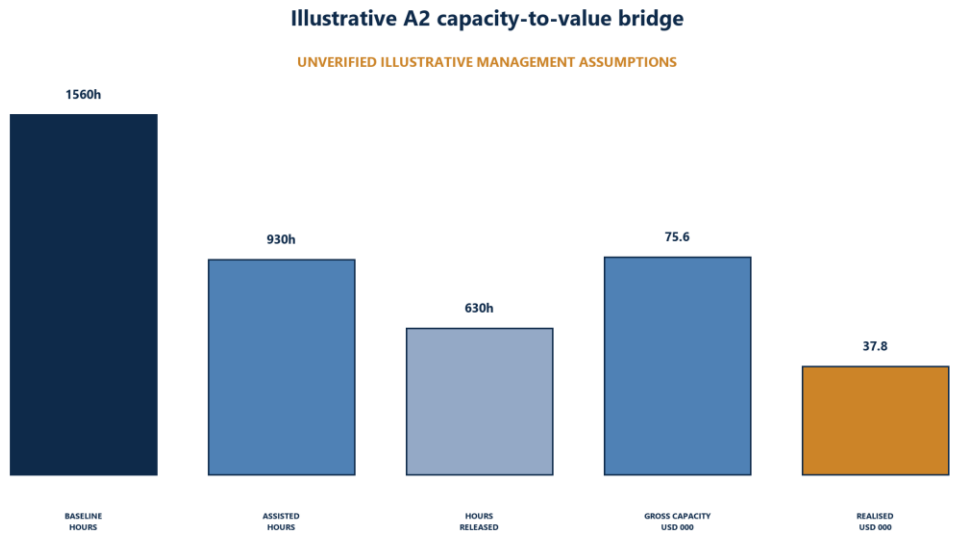

Figure 5. Illustrative A2 capacity-to-value bridge
Source: Unverified illustrative management assumptions; Matchpoint calculation.

9.3 Illustrative NPV

[Unverified] The base scenario records USD 22,680 of year-one benefit after the 60% ramp and USD 18,000 of recurring cost, giving USD 4,680 of year-one net value. Years two and three record USD 19,800 of net value before the USD 5,000 year-three exit provision. At a 12% discount rate, the three-year NPV is approximately USD 500. Undiscounted payback occurs early in year three before the terminal exit provision.

Scenario	Realisation factor	Year-one net value	Steady annual net value	Three-year NPV	Evidence
Downside	25%	USD -6,660	USD 900	approximately USD -38,100	U; unverified
Base	50%	USD 4,680	USD 19,800	approximately USD 500	U; unverified
Upside	70%	USD 13,752	USD 34,920	approximately USD 31,400	U; unverified

[Unverified] The sensitivity shows that realised use of released professional time drives more value than model cost in this illustrative case. A stronger business case could arise from platform reuse across more workflows, an observed avoided hire, approved external-spend reduction or collected revenue contribution. Each mechanism requires separate evidence and approval.

9.4 Measurement plan

Measure	Baseline	Pilot	Production gate
Population	All eligible packets and exception classes	Fixed representative sample	No unexplained denominator change
Quality	First-pass acceptance and material correction	Blind or independent review	Non-inferior quality at approved threshold
Effort	End-to-end touch time by role	Instrumented assisted touch time	Sustained hours per accepted packet
Adoption	Not applicable	Qualified use by assigned users	Sustained use without forced shadow work
Realisation	Current role and staffing plan	Named redeployment plan	Observed redeployment or cost consequence
Risk	Existing incident and control record	Adverse tests and rights review	Residual risk accepted by owner

9.5 A2 investment-committee questions

The investment committee should ask whether the reporting pack is truly comparable, whether released CIO or analyst time has an approved use, whether platform cost can serve additional workflows, whether data and service-provider rights are clear, and whether exit remains feasible. A small positive base NPV with a material downside should normally lead to staged funding and a strict scale gate.

10. ILLUSTRATIVE B4 GCC FAMILY-BUSINESS CASE

10.1 Scope

The B4 case models assisted receivables and finance-workflow packets. The system assembles invoice, delivery, payment and correspondence evidence; drafts an action; routes exceptions; and supports monthly management reporting. Named finance and commercial owners approve external communication, accounting treatment and escalation.

10.2 Illustrative baseline and economics

Input	Base assumption	Evidence class
Accepted work items per year	2,400	U; unverified illustrative management assumption
Baseline human time per item	0.55 hours	U; unverified illustrative management assumption
Assisted human time per item	0.32 hours	U; unverified illustrative management assumption
Loaded capacity value	USD 45 per hour	U; unverified illustrative management assumption
Capacity realisation factor	60%	U; unverified illustrative management assumption
Initial implementation cost	USD 8,000	U; unverified illustrative management assumption
Annual recurring full cost	USD 8,000	U; unverified illustrative management assumption
Year-one adoption ramp	55% of base benefit	U; unverified illustrative management assumption
Discount rate	12%	U; unverified illustrative management assumption
Year-three exit provision	USD 2,000	U; unverified illustrative management assumption

[Unverified] The illustrative workflow releases 552 hours each year. At USD 45 per hour, gross capacity value is USD 24,840. Applying the 60% realisation factor gives USD 14,904 of candidate annual realised capacity value. Working-capital, revenue, loss and cash-cost benefits remain USD 0 because the scenario supplies no observed causal evidence.

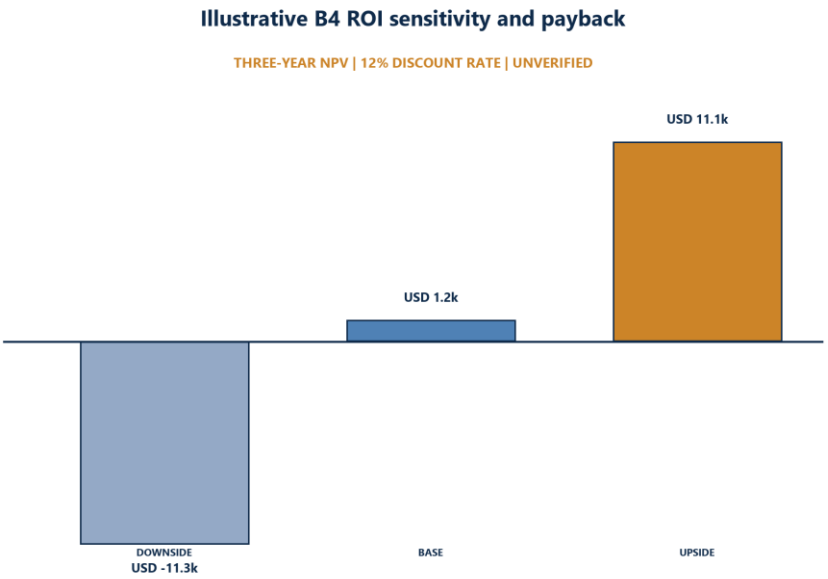

Figure 6. Illustrative B4 ROI sensitivity and payback

Source: Unverified illustrative management assumptions; Matchpoint calculation.

[Unverified] The base scenario produces USD 197 of year-one net value after ramp and USD 6,904 of steady annual net value. After the year-three USD 2,000 exit provision and a 12% discount rate, the illustrative three-year NPV is approximately USD 1,200. Undiscounted payback occurs early in year three.

Scenario	Realisation factor	Year-one net value	Steady annual net value	Three-year NPV	Evidence
Downside	35%	USD -3,218	USD 694	approximately USD -11,300	U; unverified
Base	60%	USD 197	USD 6,904	approximately USD 1,200	U; unverified
Upside	80%	USD 2,930	USD 11,872	approximately USD 11,100	U; unverified

10.3 Working-capital boundary

An AI-assisted receivables process may accelerate evidence collection and follow-up. A working-capital benefit should be measured from reconciled receivable balances and a credible comparator. The calculation should exclude changes caused by sales mix, customer quality, payment terms, disputes, seasonality and collection policy. Value can be estimated from the reduction in average receivables multiplied by the approved marginal funding rate. Principal cash collection is not revenue; it is balance-sheet conversion.

Working-capital field	Required evidence
Eligible receivables	Customer, invoice, currency, terms and exclusions
Intervention date	When the assisted workflow became effective
Comparator	Matched customer/invoice cohort or phased rollout
Outcome	Days to collect, overdue balance and dispute ageing
Funding value	Approved marginal cost of funds or liquidity method
Confounders	Terms, sales mix, write-offs, credits and seasonality
Attribution	Finance-approved causal or contribution method

10.4 B4 owner decisions

The owner should decide whether the workflow strengthens collections, finance capacity and management information without weakening customer relationships or data protection. The scenario's small base NPV means process fit and

realisation matter. A compact pilot should measure accepted actions and cash collection timing while retaining a manual fallback.

11. ARCHITECTURE, TOOL STACK AND COST DRIVERS

11.1 Economic architecture

The architecture links a business outcome to owned sources, deterministic services, bounded model work, review, release and the value ledger. OpenAI describes agents as systems combining models, tools and instructions [38]. Anthropic distinguishes workflows with predefined orchestration from agents that direct their own process and tool use [39]. The economic design should use the lowest sufficient autonomy because broader authority increases evaluation, security and control cost.

Layer	Economic role	Control
Systems of record	Supply authoritative financial and operating data	Ownership, completeness and access
Data and retrieval	Make approved evidence available	Purpose, lineage, retention and quality
Deterministic services	Calculate, match, validate and enforce limits	Versioned rules and independent tests
Model capability	Interpret, classify, draft and recommend	Representative evaluation and constraints
Orchestration	Sequence tasks, tools and exceptions	Typed contracts, stop conditions and trace
Review and release	Accept the business outcome	Named authority and segregation
Value ledger	Connect outcomes to cost and benefit	Finance approval and anti-double-counting

11.2 Model and tool selection

Model choice should follow the use-case quality threshold, latency, confidentiality, integration and cost requirements. A larger model can increase task performance and unit cost. A smaller or specialised model can lower cost and improve predictability. Routing several models can create operational complexity and evaluation burden. The portfolio business case should include regression testing when a provider or model changes.

Tool contracts should identify allowed inputs, outputs, rights, rate limits, timeouts, idempotency, approval and audit fields. Untrusted business content must remain data and should not become instructions. OWASP's 2025 LLM Top 10 identifies prompt injection, sensitive-information disclosure, supply-chain issues, improper output handling and excessive agency among relevant application risks [31]. NCSC guidance covers secure design, development, deployment and operation [30].

11.3 Cost per accepted outcome

Compute cost per generated output can understate operating cost. Cost per accepted outcome includes failed attempts, retries, retrieval, integrations, human review, exception resolution, support, monitoring and control. A cheaper model with lower acceptance can be more expensive end to end.

Unit-cost element	Allocation basis
Model and inference	All attempted calls, tokens or provider units
Retrieval and storage	Documents, queries, indexes and retention
Integration	Volume, connector or use-case activity
Human preparation and review	Observed role minutes
Failure and rework	All failed, abandoned and corrected items
Control and assurance	Risk tier, test population and review cycle
Support and operations	Tickets, incidents, releases and owner time
Shared platform	Documented cost driver with portfolio reconciliation

11.4 Data economics

Data remediation can create reusable value beyond one AI use case. The business case should distinguish minimum use-case data work from broader data-platform investment. BCBS 239 principles on risk-data aggregation and reporting

emphasise governance, architecture, accuracy, integrity, completeness, timeliness and adaptability for relevant banks [28]. A2 and B4 organisations can use these as diligence questions without representing them as universally binding.

11.5 Resilience and exit

BCBS operational-resilience principles emphasise the ability to deliver critical operations through disruption for banks within scope [29]. The FSB identifies critical third-party concentration as an AI-related vulnerability [3]. A full-cost design includes manual or deterministic fallback, last safe state, provider outage, model change, data export, contractual termination and replacement.

12. GOVERNANCE, CONTROL AND REGULATORY CONTEXT

12.1 Proportional governance

NIST AI RMF organises AI risk management through Govern, Map, Measure and Manage [24]. Its Generative AI Profile adds considerations specific to generative systems [25]. NIST CSF 2.0 provides a broader cybersecurity framework with Govern, Identify, Protect, Detect, Respond and Recover functions [26]. ISO/IEC 42001 specifies requirements for establishing, implementing, maintaining and continually improving an AI management system [33]. These frameworks can inform a proportional control system; certification and legal applicability require separate decisions.

Governance object	Minimum record
AI inventory	Use case, owner, model, tools, data, risk and status
Impact assessment	Stakeholders, rights, harms, benefits and controls
Business case	Baseline, full cost, scenarios and gates
Evaluation pack	Population, metrics, tests, results and limitations
Approval	Named roles, residual risk and permitted authority
Monitoring	Quality, drift, incidents, cost and value
Change record	Provider, model, data, prompt, tool and control changes
Exit record	Portability, retention, transition and termination

12.2 UAE financial-services context

The CBUAE's February 2026 guidance note applies to licensed financial institutions and insurance providers within its scope and focuses on consumer protection and responsible AI/ML adoption. It addresses governance and accountability, fairness, transparency, explainability, human oversight, data management and privacy [6]. The joint 2021 enabling-technologies guidelines expect a proportionate approach based on size, complexity, nature, risk and materiality, and address reliability, transparency, data governance, validation, human intervention and customer disclosure [7].

The CBUAE's 2025 Financial Stability Report describes AI's potential to improve inclusion and efficiency, alongside discrimination and resilience risks, and notes supervisory use of AI-enabled analytics [8]. A UAE family office or family business outside the licensed-financial-institution perimeter should confirm which rules apply to its entities and activities. The guidance can still inform its control design.

12.3 International financial stability and model risk

The FSB's 2024 assessment highlights third-party concentration, market correlation, cyber, model and data vulnerabilities [3]. Its 2026 sound-practices document remained a consultation at 1 August 2026 [4]. BCBS's 2024 digitalisation report considers AI, APIs, cloud and distributed-ledger technology across the banking value chain and supervision [5].

US SR 26-2 provides current interagency model-risk guidance and supersedes SR 11-7 [27]. A model inventory, materiality assessment, effective challenge, validation, monitoring and governance can be proportionately useful beyond direct scope. A tool marketed as generative AI may contain models, rules and human processes; governance should follow function and consequence.

12.4 Privacy and autonomous processing

The UAE Personal Data Protection Law provides the federal personal-data framework subject to scope and exceptions [36]. DIFC Data Protection Regulations include Regulation 10 on personal data processed through autonomous and semi-autonomous systems [37]. Cross-border family-office and financial-services workflows may also encounter the EU AI Act and DORA depending on entities, activities and counterparties [34-35]. Qualified legal and privacy owners determine applicability, notices, lawful basis, transfer, rights and records.

12.5 Responsible outcomes

The OECD AI Principles, updated in 2024, promote innovative and trustworthy AI respectful of human rights and democratic values [32]. Responsible outcomes have economic relevance. Bias, opaque decisioning, security failures and weak recourse can create customer harm, remediation cost and loss of trust. The value model should not monetise compliance as a benefit merely because controls exist; it should include control cost and residual risk.

13. PROCUREMENT AND OPERATING MODEL

13.1 Due diligence

Vendor due diligence should join technical, security, data, financial, legal, operational and exit questions. A proof of concept should use approved representative data and should not silently become production. Provider claims about accuracy or productivity should remain provider evidence until local evaluation confirms them.

Domain	Questions
Capability	Which task, population, language and evidence support performance?
Data	What is processed, retained, trained on, transferred and deleted?
Security	How are identity, encryption, secrets, logging and incidents managed?
Model	Which models, versions, routing and change notices apply?
Control	Can tools, actions, outputs and human approvals be constrained?
Resilience	What are service levels, dependencies, fallback and recovery evidence?
Economics	What are committed, usage, support, integration and exit costs?
Portability	Can prompts, evaluation sets, logs, data and outputs be exported?

13.2 Commercial structure

Pricing can be per seat, usage, workflow, transaction, model unit or outcome. The contract should align the price unit with the value unit while avoiding unbounded usage and perverse incentives. Outcome pricing requires an agreed outcome definition, baseline, attribution and audit right. A capped pilot with conversion gates can preserve option value.

13.3 Roles and decision rights

Role	Accountability
Business sponsor	Outcome, funding and realised value
Process owner	Workflow, acceptance and operating change
Finance owner	Baseline, cost, benefit, NPV and recognition
Data owner	Rights, quality, lineage and retention
Technology owner	Architecture, integration, operations and exit
Risk/privacy/security owners	Impact, controls, residual risk and incidents
Model or product owner	Evaluation, release, monitoring and change
User/reviewer	Qualified use, exception handling and acceptance

13.4 Change and adoption

The 2026 peer-reviewed jagged-frontier study involved 758 consultants. Within the tested capability frontier, AI users completed 12.2% more tasks and worked 25.1% faster with higher quality; on a task outside that frontier, AI users were 19 percentage points less likely to be correct [20]. Training should therefore teach task boundaries, evidence review, uncertainty, escalation and recovery. Prompt tips alone are insufficient.

Adoption is designed through workflow fit, incentives, role clarity, feedback and trust. Mandatory use can inflate activity and hide shadow work. Users should be able to reject output, classify a failure and see how issues are corrected. Managers should protect time for learning and avoid counting double work as productivity.

14. GATED ADOPTION ROADMAP

Figure 7. Gated 12-18 month AI economics roadmap

Source: Matchpoint adoption and capital-allocation framework.

Stage	Indicative period	Core output	Capital gate
0 Charter	Weeks 0-2	Outcome, owner, population and risk tier	Problem is material and owned
1 Baseline	Weeks 2-6	Volume, effort, quality, cost and outcome	Measurement is credible
2 Design	Weeks 5-10	Workflow, data, architecture, controls and case	Full-cost range is acceptable
3 Prototype	Weeks 8-14	Capability and integration evidence	Representative task is feasible
4 Pilot	Months 3-5	Comparative evaluation and user evidence	Quality, time and risk thresholds pass
5 Shadow	Months 5-8	End-to-end run with fallback	Operating controls are stable
6 Restricted production	Months 8-12	Accepted outcomes and value ledger	Marginal economics are positive
7 Portfolio scale	Months 12-18	Shared platform and use-case portfolio	Confidence-weighted portfolio NPV passes

14.1 First 30 days

The first month should produce a ranked use-case inventory, one selected outcome, a baseline design, an AI inventory entry, a preliminary impact assessment and a capped discovery budget. The team should also identify existing tools, duplicate licences and unsanctioned use.

14.2 Days 31-90

The next period should complete the baseline, representative evaluation set, architecture, vendor diligence, process redesign, control design and full-cost model. The investment decision should state downside, base and upside scenarios, a maximum loss and the next evidence gate.

14.3 Months four to twelve

The programme should progress from controlled pilot to shadow and restricted production only after gates pass. Measurement should include all failures, review and support. The finance owner should reconcile estimates with actual spend and approved benefits monthly or quarterly, depending on materiality.

14.4 Portfolio scale

Shared identity, retrieval, evaluation, observability and governance can lower marginal cost. Platform reuse should be demonstrated through live use cases rather than forecast counts. The portfolio review should retire redundant tools, consolidate contracts where concentration remains acceptable and preserve portability.

15. LIMITATIONS, RESEARCH AGENDA AND CONCLUSION

The public evidence has important limitations. Regulator surveys are self-reported and use heterogeneous definitions. Company disclosures use internal value methods that may not be fully comparable or externally assured as AI ROI measures. Controlled studies test particular models, tasks, users and periods. SME surveys describe reported experience and do not always measure magnitude. Rapid model and price changes shorten the life of technical estimates.

The illustrative A2 and B4 cases contain no observed client or Matchpoint data. Their assumptions, scenarios, NPVs and payback periods are unverified. They demonstrate arithmetic and decision gates. They do not forecast performance. The framework also does not determine the accounting, legal, regulatory, tax, privacy or employment treatment of an implementation.

Future research should publish consistent cost and benefit definitions, distinguish capacity from cash, report failed and abandoned outputs, compare model and workflow configurations, measure long-term adoption and skill effects, and study smaller financial institutions, family offices and GCC family businesses. Independent work is also needed on revenue attribution, rare-loss estimation, model-change economics and exit cost.

The practical conclusion is concise. AI adoption creates credible economic value when a named accepted outcome improves, the full operating cost is measured, quality and risk remain within approved bounds, and capacity or commercial effects are realised. An investment committee should fund evidence in stages, value the portfolio as well as the use case, and retire work that does not clear its gate. This method converts AI enthusiasm into a controlled capital-allocation process.

DECLARATIONS

Purpose. General research for professional audiences.
Author. Prepared by Matchpoint Partners. Named-person author attribution remains pending CK approval.
Evidence date. Sources were reviewed for this paper through 1 August 2026.
Commercial attribution. Attributed Matchpoint or client revenue, cash cost reduction and loss reduction are USD 0. No approved observed attribution evidence was supplied.
Illustrative assumptions. All worked volumes, times, rates, costs, realisation factors, discount rates, scenario probabilities, NPVs and payback periods are unverified illustrative management assumptions.
Professional boundary. This document is not investment, legal, regulatory, accounting, audit, tax, privacy, cybersecurity, employment, technology or valuation advice. It is not an offer, solicitation, recommendation or promise of results.

REFERENCES

[1] Bank of England and Financial Conduct Authority (2024). Artificial Intelligence in UK Financial Services - 2024. https://www.bankofengland.co.uk/report/2024/artificial-intelligence-in-uk-financial-services-2024
[2] Bank of England and Financial Conduct Authority (2023). FS2/23 - Artificial Intelligence and Machine Learning. https://www.bankofengland.co.uk/prudential-regulation/publication/2023/october/artificial-intelligence-and-machine-learning
[3] Financial Stability Board (2024). The Financial Stability Implications of Artificial Intelligence. https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/
[4] Financial Stability Board (2026). Sound Practices for Financial Institutions' Responsible AI Adoption. Consultation report at 1 August 2026. https://www.fsb.org/uploads/P100626.pdf
[5] Basel Committee on Banking Supervision (2024). Digitalisation of Finance. https://www.bis.org/bcbs/publ/d575.htm
[6] Central Bank of the United Arab Emirates (2026). Guidance Note on the Consumer Protection and Responsible Adoption and Use of Artificial Intelligence and Machine Learning by Licensed Financial Institutions in the U.A.E. https://rulebook.centralbank.ae/en/rulebook/guidance-note-consumer-protection-and-responsible-adoption-and-use-artificial-intelligence
[7] Central Bank of the United Arab Emirates, SCA, DFSA and FSRA (2021). Guidelines for Financial Institutions Adopting Enabling Technologies. https://rulebook.centralbank.ae/en/rulebook/guidelines-financial-institutions-adopting-enabling-technologies
[8] Central Bank of the United Arab Emirates (2026). Financial Stability Report 2025. https://www.centralbank.ae/media/kaqlwo0h/cbuae-fsr-report_2025_en.pdf
[9] DBS Group Holdings (2026). DBS Annual Report 2025 - CEO Reflections. https://www.dbs.com/annualreports/2025/ceo-reflections.html
[10] Bank of America Corporation (2026). 2025 Annual Report. https://www.sec.gov/Archives/edgar/data/70858/000007085826000209/bofaannualreport2025.pdf
[11] JPMorgan Chase & Co. (2026). 2025 Annual Report - Letter from the CEO of Asset and Wealth Management. https://www.jpmorganchase.com/ir/annual-report/2025/ar-ceo-letter-mary-callahan-erdoes
[12] JPMorgan Chase & Co. (2026). 2025 Annual Report - Letter from the Co-CEOs of Commercial and Investment Banking. https://www.jpmorganchase.com/ir/annual-report/2025/ar-ceo-letter-petno-rohrbaugh
[13] JPMorgan Chase & Co. (2026). 2025 Annual Report - Letter from the CEO of Consumer and Community Banking. https://www.jpmorganchase.com/ir/annual-report/2025/ar-ceo-letter-marianne-lake
[14] UBS (2025). Global Family Office Report 2025. https://advisors.ubs.com/mediahandler/media/708870/Global%20Family%20Office%20Report%20_2025_.pdf
[15] UBS (2026). Global Family Office Report 2026. https://www.ubs.com/content/dam/assets/wm/static/gfo/global-family-office-report-en-2026.pdf
[16] OECD (2025). Generative AI and the SME Workforce: New Survey Evidence. https://www.oecd.org/en/publications/generative-ai-and-the-sme-workforce_2d08b99d-en/full-report.html
[17] OECD (2025). AI Adoption by Small and Medium-Sized Enterprises. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/12/ai-adoption-by-small-and-medium-sized-enterprises_9c48eae6/426399c1-en.pdf
[18] OECD (2026). AI Use by Individuals Surges as Adoption by Firms Continues to Expand. https://www.oecd.org/en/about/news/announcements/2026/01/ai-use-by-individuals-surges-across-the-oecd-as-adoption-by-firms-continues-to-expand.html
[19] Brynjolfsson, E., Li, D. and Raymond, L. R. (2023). Generative AI at Work. NBER Working Paper 31161. https://www.nber.org/papers/w31161
[20] Dell'Acqua, F. et al. (2026). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality. Organization Science. https://www.hbs.edu/ris/Publication%20Files/dell-acqua-et-al-2026-navigating-the-jagged-technological-frontier_5c589c8c-fbb5-458f-b285-c944746cd717.pdf
[21] Becker, J., Rush, N., Barnes, B. and Rein, D. (2025). Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity. METR. https://metr.org/blog/2025-07-10-early-2025-ai-experienced-os-dev-study/
[22] Becker, J., Rush, N., Cunningham, T., Rein, D. and Mahamud, K. (2026). We Are Changing Our Developer Productivity Experiment Design. METR. https://metr.org/blog/2026-02-24-uplift-update/
[23] Stanford Institute for Human-Centered Artificial Intelligence (2026). AI Index Report 2026. https://hai.stanford.edu/ai-index/2026-ai-index-report
[24] National Institute of Standards and Technology (2023). Artificial Intelligence Risk Management Framework 1.0. https://www.nist.gov/itl/ai-risk-management-framework
[25] National Institute of Standards and Technology (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST AI 600-1. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

[26] National Institute of Standards and Technology (2024). Cybersecurity Framework 2.0. https://www.nist.gov/publications/nist-cybersecurity-framework-csf-20
[27] Board of Governors of the Federal Reserve System, Office of the Comptroller of the Currency and Federal Deposit Insurance Corporation (2026). SR 26-2: Revised Guidance on Model Risk Management. https://www.federalreserve.gov/supervisionreg/srletters/SR2602.htm
[28] Basel Committee on Banking Supervision (2013). Principles for Effective Risk Data Aggregation and Risk Reporting. https://www.bis.org/publ/bcbs239.htm
[29] Basel Committee on Banking Supervision (2021). Principles for Operational Resilience. https://www.bis.org/bcbs/publ/d516.htm
[30] UK National Cyber Security Centre and partners (2023). Guidelines for Secure AI System Development. https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development
[31] OWASP Foundation (2025). OWASP Top 10 for LLM Applications v2.0. https://owasp.org/www-project-top-10-for-large-language-model-applications/assets/PDF/OWASP-Top-10-for-LLMs-v2025.pdf
[32] OECD (2024). OECD AI Principles. https://oecd.ai/en/ai-principles
[33] International Organization for Standardization (2023). ISO/IEC 42001:2023 - Artificial Intelligence Management Systems. https://www.iso.org/standard/42001
[34] European Union (2024). Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32024R1689
[35] European Union (2022). Regulation (EU) 2022/2554 on Digital Operational Resilience for the Financial Sector. https://eur-lex.europa.eu/eli/reg/2022/2554/oj
[36] United Arab Emirates Government (2021). Federal Decree-Law No. 45 of 2021 Regarding the Protection of Personal Data. https://u.ae/en/about-the-uae/digital-uae/data/data-protection-laws
[37] Dubai International Financial Centre (current at 2026). Data Protection Regulations; Regulation 10: Personal Data Processed through Autonomous and Semi-Autonomous Systems. https://assets.difc.com/v1/media/edge/images/dubaiintern0078-difcexperie96c5-production-3253/media/project/difcexperiences/difc/difcwebsite/documents/laws--regulations/data-protection-regulation.pdf
[38] OpenAI (2025). A Practical Guide to Building Agents. https://openai.com/business/guides-and-resources/a-practical-guide-to-building-ai-agents/
[39] Anthropic (2024). Building Effective Agents. https://www.anthropic.com/engineering/building-effective-agents
[40] IFRS Foundation (current at 2026). IAS 38 Intangible Assets. https://www.ifrs.org/issued-standards/list-of-standards/ias-38-intangible-assets/

APPENDIX A. USE-CASE BUSINESS-CASE TEMPLATE

Field	Required content
Use case	ID, title, process, version and status
Owner	Sponsor, process owner, finance owner and risk owner
Outcome	Accepted output, decision and authority
Population	Users, items, entities, periods and exclusions
Baseline	Volume, effort, quality, cost, loss and outcome
Intervention	Model, tools, data, workflow and change
Benefits	Capacity, cash, contribution, loss and capital
Costs	Initial, recurring, shared and exit
Evaluation	Design, sample, metrics, thresholds and period
Risk	Tier, controls, residual risk and incidents
Finance	Scenarios, discount rate, NPV, payback and confidence
Gates	Fund, pause, scale, redesign and retire conditions

APPENDIX B. KPI DICTIONARY

KPI	Definition
Eligible volume	All items meeting the fixed population rule
Qualified adoption	Eligible users or workflows meeting the defined activity threshold
Accepted outcome rate	Accepted outcomes divided by all eligible attempted outcomes
First-pass acceptance	Outcomes accepted without material correction divided by reviewed outcomes
Human hours per accepted outcome	All eligible human touch time divided by accepted outcomes
Unit cost	Full process cost divided by accepted outcomes
Cycle time	Business trigger to accepted release
SLA attainment	Accepted outcomes within deadline divided by eligible outcomes
Material defect rate	Outcomes with a defined material defect divided by reviewed outcomes
Capacity realisation	Approved and observed value-bearing use divided by gross released capacity
Collected contribution	Attributed collected revenue less attributed variable cost
Expected-loss change	Exposure-weighted change relative to approved counterfactual

APPENDIX C. ILLUSTRATIVE FORMULAE

Gross capacity value = Q x (Hb - Ha) x W.
Candidate realised capacity value = gross capacity value x R.
Approved total benefit = approved realised capacity + cash cost reduction + collected contribution + expected-loss reduction + capital or working-capital benefit.
Net annual value = approved total benefit - recurring full cost.
NPV = -initial implementation cost + discounted annual net value - discounted terminal exit cost.
Adjusted theoretical benefit = theoretical gross benefit x adoption factor x quality factor x realisation factor x attribution factor, unless a factor is already embedded in the observed denominator.
Attributed Matchpoint or client revenue, cash cost reduction and loss reduction remain USD 0 until approved observed evidence exists.

APPENDIX D. MINIMUM CONTROL REGISTER

Objective	Example control	Evidence
Owned use case	Named sponsor, process, finance and risk owners	Approved inventory record
Complete population	Fixed eligibility and denominator reconciliation	Population manifest
Reliable quality	Representative evaluation and material-defect threshold	Evaluation pack
Supported output	Claim or value linked to controlled source	Source and transformation record
Restricted authority	Read, recommend, approve and release segregated	Access and approval logs
Protected data	Purpose, minimisation, retention and transfer controls	Data assessment and logs

Objective	Example control	Evidence
Secure operation	Threat model, testing, monitoring and incident process	Security evidence
Controlled change	Regression after model, prompt, tool or data change	Change and test record
Realised value	Finance-approved cost or benefit record	Value ledger
Recoverable exit	Fallback, export, transition and termination tested	Recovery and exit test

APPENDIX E. RED-FLAG REGISTER

Red flag	Why it matters	Immediate response
ROI uses hours saved as cash savings	Capacity has no observed financial consequence	Reclassify and set cash benefit to zero
Denominator excludes failed or abandoned work	Productivity and unit cost are overstated	Reconstruct the full eligible population
Revenue claim ends at activity or pipeline	Collection and contribution are unproven	Continue the attribution ladder
Loss reduction has no counterfactual	Avoided loss is speculative	Set value to zero or create an evaluation design
Licence cost is the only cost	Integration, review and control are omitted	Build the full-cost ledger
Model change bypasses regression	Operating evidence is stale	Pause and retest
Agent prepares and releases material action	Segregation and authority are weak	Restrict action and review incidents
Sensitive data enters an unapproved tool	Rights and confidentiality may be breached	Isolate, assess and remediate
User activity is forced while shadow work continues	Adoption and productivity are distorted	Measure total workflow effort
Portfolio forecast counts unapproved use cases	Shared-platform value is inflated	Value only gated, funded use cases
Business case has no exit provision	Downside cost is understated	Add portability and transition cost
Named professional judgement is displaced	Regulatory, fiduciary or professional boundary may fail	Stop and escalate to the authorised owner