MatchPoint Research Working Paper No. 2026/T20

AI Governance and Model Risk in Regulated Finance

An evidence-gated operating framework for family offices and institutional allocators

Prepared by Matchpoint Partners
Author attribution pending CK approval
August 2026

ABSTRACT

Background. AI is entering research, risk, compliance, operations and investment workflows while legal, supervisory and standards sources retain different scopes and authority. **Objective.** This paper develops an evidence-gated AI governance and model-risk framework for A2 family-office CIOs and A1 international institutional allocators. **Approach.** The analysis reviews 40 primary laws, regulators, standard setters and recognised standards sources available through 1 August 2026. **Findings.** A complete inventory, consequence-led materiality, explicit authority, system-level evaluation, independent challenge, third-party controls, monitoring and tested exit form one coherent operating model. **Implications.** Institutions can tailor governance to scale while preserving control objectives, jurisdictional status and decision evidence. All worked inputs are unverified illustrative management assumptions; attributed Matchpoint or client revenue, cash cost reduction and loss reduction remain USD 0 until approved observed evidence exists.

Highlights

Perimeter. One register covers models, AI systems, assistants, agents and embedded vendor capabilities.

Materiality. Consequence, authority, exposure and reversibility drive proportional control.

Evaluation. Predictive, generative, retrieval and agentic systems require behaviour-specific testing.

Dependency. Third-party and fourth-party services remain inside governance, resilience and exit.

Oversight. Board information should show tier, authority, incidents, exceptions, concentration and recovery.

JEL Classification: C53, C58, G20, G23, G28, G38, M15, O33

Keywords: artificial intelligence, AI governance, model risk, regulated finance, family office, institutional allocator, validation, generative AI, agentic AI, third-party risk, operational resilience, data governance

1. INTRODUCTION

Artificial intelligence has moved from isolated experimentation into investment research, client service, compliance, operations, risk analysis and technology delivery. In regulated finance, the central management question is whether each use can be governed in proportion to its consequence. The answer requires an inventory, a defensible classification, accountable decision rights, independent challenge, reliable evaluation, traceable operation and credible exit. A policy statement alone cannot provide that evidence.

The regulatory baseline changed materially in 2026. The US federal banking agencies issued revised model risk management guidance on 17 April 2026. The guidance supersedes SR 11-7, emphasises a risk-based and tailored approach, and is expected to be most relevant to Federal Reserve regulated banking organisations with more than USD 30 billion in total assets [1]. The OCC also stated that generative AI and agentic AI are outside the scope of that revised guidance because they are novel and rapidly evolving [2]. That scope boundary matters. It means an institution cannot assume that conventional model risk management alone answers every generative or agentic AI question.

In the UK, the April 2026 version of PRA Supervisory Statement SS1/23 applies model risk management principles across models used to inform business decisions, including vendor models and artificial intelligence where the principles apply [4]. The Bank of England, FCA and HM Treasury separately called for board understanding and active mitigation of frontier-AI cyber and operational-resilience risks in May 2026 [9]. In June 2026, the Financial Stability Board proposed 12 sound practices covering organisation-wide governance, lifecycle controls, cyber, information and communication technology, and third-party risk [10]. That FSB document remained a consultation at the 1 August 2026 evidence cut-off; its final report was scheduled for October 2026 [10].

The UAE position also became more specific. The Central Bank of the UAE issued consumer-protection guidance for responsible AI and machine-learning use by licensed financial institutions in February 2026 [16]. The DFSA published regulatory expectations on AI risk management in the DIFC in June 2026 [18]. These materials sit alongside the UAE regulators' 2021 enabling-technology guidelines, the UAE personal-data framework and DIFC rules for personal data processed by autonomous and semi-autonomous systems [17,22,23].

This paper converts those overlapping sources into an operating framework for A2 family-office CIOs and heads of alternatives and A1 international institutional allocators. The framework is designed for investment institutions that may be regulated directly, may allocate to regulated firms, may procure AI-enabled services, or may need institutional-grade governance because of fiduciary, contractual or reputational exposure. It distinguishes law, binding rules, supervisory guidance, consultation material, standards and voluntary practice. Applicability must be confirmed by qualified legal and compliance owners for the entity, activity and jurisdiction.

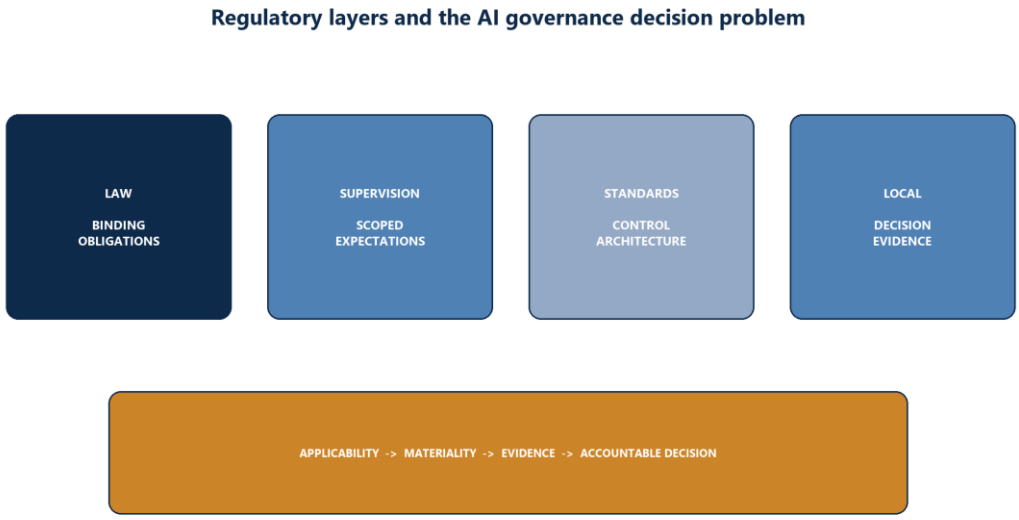

Figure 1. Regulatory layers and the AI governance decision problem

Source: Matchpoint synthesis of primary laws, regulators, standard setters and standards [1-40].

The contribution is practical. It defines a common inventory covering models, AI systems, assistants, agents and embedded vendor capabilities; a materiality method that separates decision consequence from technical novelty; a lifecycle with evidence gates; evaluation requirements for deterministic, predictive, generative and agentic systems; third-

party and concentration controls; and board information that supports actual challenge. Two worked operating cases are included. Every numerical input in those cases is explicitly unverified and illustrative. Attributed Matchpoint or client revenue, cash cost reduction and loss reduction remain USD 0 until approved observed evidence exists.

The paper makes five propositions.

Proposition	Decision implication
P1. Inventory completeness precedes risk classification.	Unrecorded use cannot be governed, monitored or retired.
P2. Materiality depends on consequence, authority and reversibility.	A familiar model can be high consequence; a novel model can be low consequence.
P3. Validation must match system behaviour.	Generative and agentic systems require evaluation beyond conventional point-estimate accuracy.
P4. Third-party AI remains the institution's governance problem.	Procurement evidence, contractual rights, monitoring and exit are control requirements.
P5. Board oversight needs decision evidence.	Counts, incidents, exceptions, residual risk and exit readiness are more useful than adoption narratives.

2. SCOPE, DEFINITIONS AND EVIDENCE BOUNDARIES

2.1 Scope

The subject is governance and model risk for AI used in regulated finance and institutionally governed investment activity. Covered uses include research, screening, portfolio analytics, valuation support, forecasting, financial crime controls, client communications, suitability support, operations, coding, document analysis and workflow agents. The paper addresses internally developed systems, configured third-party systems, embedded AI features and externally supplied model outputs.

The paper does not determine legal status, regulatory perimeter, fiduciary compliance, suitability, data-law compliance or accounting treatment for a particular institution. It does not certify a model, a vendor or an AI system. Each institution must establish the applicable requirements through its authorised legal, compliance, risk, information-security, privacy and business owners.

2.2 Working definitions

Definitions vary across rules and standards. The institution should preserve the applicable legal definition in its obligations register and maintain a broader operational perimeter for governance.

Term	Operational definition used in this paper
Model	A quantitative, statistical, economic or mathematical method that transforms inputs into estimates, decisions or information used in business decisions.
AI system	A machine-based system that infers outputs such as predictions, content, recommendations or decisions from inputs for explicit or implicit objectives.
Generative AI	AI that produces content, including text, code, images, audio or structured data.
Foundation model	A model trained broadly and capable of adaptation across multiple downstream tasks.
Agentic system	An AI-enabled system that plans or selects actions, invokes tools or services, and may continue across multiple steps.
Decision support	Output informs a human decision while documented authority remains with the human.
Automated decision	The system determines or executes an outcome within delegated authority.
Model risk	Potential adverse consequences from incorrect or misused model outputs and decisions [1].
AI risk	Risk arising from the design, development, procurement, deployment, operation, interaction, misuse or retirement of an AI system.
Materiality	The significance of a use based on consequence, exposure, scale, authority, rights impact, reversibility and interconnectedness.

The operational perimeter is intentionally inclusive. A spreadsheet formula, vendor score, retrieval assistant or workflow agent may not satisfy every legal definition of a model. It can still create consequential dependency. Registration routes the item to the correct governance owner; it does not predetermine regulatory classification.

2.3 Evidence classes

Evidence has different authority. Mixing these classes creates false certainty.

Class	Examples	Permitted use
Law and binding regulation	EU AI Act, DORA, GDPR, UAE data laws	Determine obligations after applicability is confirmed.
Supervisory rule or statement	PRA SS1/23, CBUAE rulebook guidance, DFSA rules	Translate supervisory expectations within stated scope.
Supervisory guidance	US interagency revised MRM guidance, FINMA guidance	Design risk-based practices while preserving non-binding status where stated.
Consultation	FSB June 2026 sound practices	Anticipate direction; label as proposed and monitor finalisation.
Consensus standard	ISO/IEC 42001 and 23894	Structure management systems and risk processes; confirm certification claims separately.
Voluntary framework	NIST AI RMF, OECD AI Principles	Create common language and operational outcomes.
Market evidence	Regulator surveys	Inform priors; preserve sample, date and population.
Local evidence	Inventory, tests, incidents, logs, approvals	Support the institution's own decision.

2.4 Evidence cut-off and current-state limitations

Sources were reviewed through 1 August 2026. The EU implementation timetable changed in July 2026, including extended dates for high-risk AI obligations [25]. Article 50 transparency obligations were scheduled to apply from 2 August 2026, one day after the evidence cut-off [26]. The FSB sound practices remained a consultation [10]. NIST stated that AI RMF 1.0 was being revised [29]. The obligations register must therefore record source date, current status, next review date and accountable legal owner.

2.5 Research method

The paper prioritises primary laws, regulators, standard setters and recognised standards bodies. It triangulates model risk, AI governance, operational resilience, cybersecurity, data protection and third-party risk. It does not treat a survey percentage as a universal adoption rate. It does not treat a voluntary standard as law. It does not interpret consultation language as final policy.

3. A2 AND A1 DECISION MAP

3.1 A2 family-office CIOs and heads of alternatives

The Topic Tracker defines A2 as investment professionals at single-family offices, multi-family offices, private-wealth organisations and external asset managers across the GCC, UK, Switzerland, Singapore and the EU. Their AI governance problem often spans a small internal team, confidential family information, external managers, private-market documents, bank and administrator data, and multiple outsourced technology providers.

A2 decision	Material risk question	Minimum evidence
Use AI for manager research	Can unsupported content influence selection or rejection?	Controlled corpus, citation coverage, reviewer protocol, exception log.
Summarise fund and deal documents	Can omitted or altered terms affect a commitment?	Clause-level traceability, material-term checklist, legal review boundary.
Generate investment-committee material	Can generated statements enter an approved record?	Source map, named preparer, reviewer, version and approval evidence.
Analyse portfolio data	Are data rights, lineage and reconciliation sufficient?	Data register, purpose, lineage, completeness and reconciliation tests.
Automate instructions or communications	Can the system bind, disclose or transmit value?	Authority limits, dual control, allow-listed tools, audit log and recovery.
Use embedded vendor AI	Is the capability visible and contractually governed?	Vendor inventory, change notice, data terms, security evidence and exit plan.

A2 institutions should avoid copying a bank's organisational scale. Proportionality can preserve the same control objectives with fewer committees. A named accountable executive, an independent challenger and a complete evidence pack may be more effective than multiple forums with unclear ownership.

3.2 A1 international institutional allocators

The Topic Tracker defines A1 as pensions, insurers, endowments and fund-of-funds in the US, EU and UK evaluating or scaling allocations to UAE and GCC private markets. Their AI governance problem adds delegation, public or beneficiary accountability, regulated outsourcing, committee process, manager oversight and cross-border data flows.

A1 decision	Material risk question	Minimum evidence
Screen managers or opportunities	Does AI change access, scoring or escalation?	Feature rationale, bias analysis where relevant, override and appeal process.
Produce due-diligence questions	Are gaps and adverse facts preserved?	Population coverage, source retention, reviewer sign-off and sampling.
Monitor funds and portfolios	Can model drift or data delay change risk signals?	Data timeliness, benchmark, drift limits, incident and escalation rules.
Support valuation or risk estimates	Is the method inside the model-risk perimeter?	Classification, development evidence, independent validation and limitations.
Outsource AI-enabled processing	Can the allocator audit, transition and continue service?	Contract, subcontractor map, service levels, testing, portability and exit.
Use AI in regulated reporting	Can provenance and control evidence withstand review?	Reconciliation, lineage, maker-checker approval, retention and reproducibility.

3.3 Shared governance outcome

Both ICPs need a compact decision contract.

Question	Required answer
What is the system?	Registered scope, version, owner, provider, data and interfaces.
What decision does it affect?	Use case, user, beneficiary, process and consequence.
What authority does it have?	Read, draft, recommend, approve, execute, communicate or transfer.
What can fail?	Error, bias, unsupported output, misuse, security, privacy, concentration and resilience.
How is it tested?	Representative evaluation, thresholds, challenger and decision record.
How is it operated?	Monitoring, exceptions, changes, incidents and user competence.
How is it stopped?	Kill switch, fallback, data export, transition and retirement evidence.

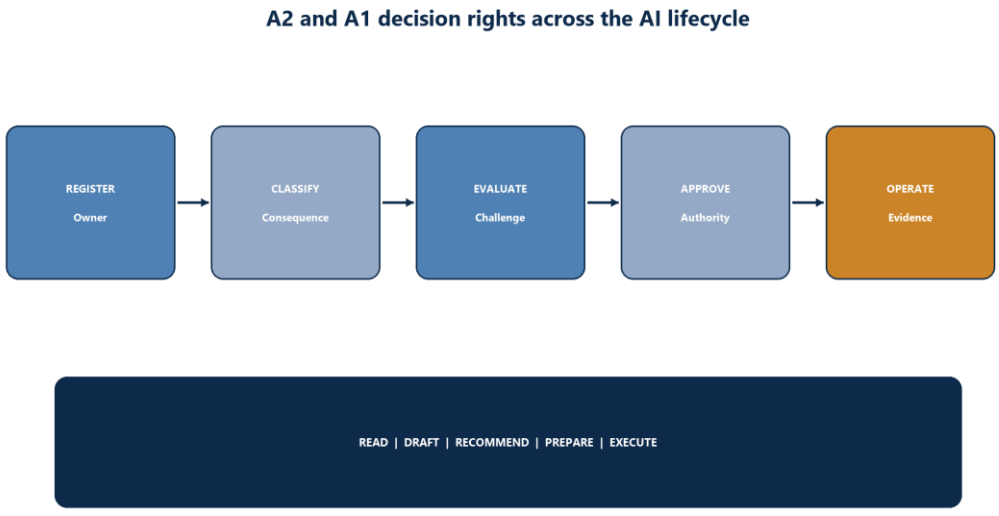

Figure 2. A2 and A1 decision rights across the AI lifecycle

Source: Matchpoint governance framework.

4. REGULATORY AND SUPERVISORY LANDSCAPE

4.1 United States banking model-risk guidance

The April 2026 interagency guidance describes model risk as the potential for adverse consequences from decisions based on incorrect or misused model outputs. It highlights effective development and use, validation, and governance and controls [1]. It expects tailoring to the banking organisation's model risk profile, size and complexity. Federal Reserve applicability is described as most relevant above USD 30 billion in total assets [1]. The OCC stated that the guidance is not an enforceable standard and that generative and agentic AI models are outside its scope [2].

This creates a two-part governance task. Conventional models should be mapped to the revised guidance where applicable. Generative and agentic systems need an adjacent AI-risk method that retains useful model-risk disciplines while adding content, interaction, tool-use, security, data-rights and autonomy controls. Institutions should record which regime or policy owns each control objective.

US guidance point	Operating translation
Risk-based tailoring	Set validation depth and frequency from materiality and exposure.
Development and use	Record purpose, design, data, assumptions, limitations and intended users.
Validation	Apply independent, effective challenge with scope matched to risk.
Governance and controls	Maintain board or senior oversight, policies, inventory, issues and audit.
Vendor products	Obtain enough information and testing to understand and manage dependency.
GenAI and agentic exclusion	Use a documented adjacent AI framework; do not imply coverage by the MRM guidance.

4.2 United Kingdom

PRA SS1/23 sets five principles: model identification and classification; governance; development, implementation and use; independent validation; and model-risk mitigants [4]. It applies to vendor models as well as internally developed models and addresses AI in modelling techniques where the general principles apply [4]. Bank and FCA materials also emphasise governance, accountability, data, model and third-party risks [5,6].

The May 2026 UK joint statement adds a distinct frontier-AI cyber-resilience concern. It calls for board and senior-management understanding, investment and resourcing, access management, network security, data protection, protective and detective capabilities, threat containment and response [9]. The July 2026 Financial Stability Report described four channels: core financial decision-making, financial markets, service-provider operational risk and the changing external cyber threat [8].

The governance implication is that AI risk cannot sit only within model validation. It needs coordination across business ownership, model risk, operational risk, information security, privacy, compliance, procurement and resilience.

4.3 UAE and DIFC

The CBUAE's February 2026 guidance focuses on consumer protection and market conduct for licensed financial institutions using AI and machine learning. It addresses transparency, bias, ethics, accountability, explainability and data privacy [16]. The joint UAE enabling-technology guidelines address AI, big-data analytics, cloud, application programming interfaces, biometrics and distributed ledger technology [17].

The DFSA's June 2026 SEO letter records regulatory expectations for AI risk management in the DIFC [18]. Its 2025 survey covered 661 authorised firms with 88 per cent participation. The DFSA reported that 52 per cent used AI, 60 per cent had some form of AI governance structure and 21 per cent lacked clear accountability or oversight even where use was critical [19]. Those figures describe that survey population and date; they are not universal adoption estimates.

DIFC Data Protection Regulation 10 addresses personal data processed through autonomous and semi-autonomous systems [22]. UAE Federal Decree-Law No. 45 of 2021 establishes the federal personal-data framework [23]. The institution's obligations register should identify the governing data regime, controller and processor roles, legal basis, sensitive-data constraints, transfer requirements, individual rights, retention and incident obligations.

4.4 European Union

Regulation (EU) 2024/1689 establishes a risk-based AI framework [24]. The July 2026 AI Omnibus changed elements of implementation and extended the start of high-risk AI rules to 2 December 2027, with rules for AI embedded in regulated physical products scheduled for 2 August 2028 [25]. Transparency obligations for specified AI interactions and generated or manipulated content were scheduled to apply from 2 August 2026 [26].

DORA requires financial entities in scope to identify and document ICT assets, dependencies and critical interconnections; manage ICT risk; govern third-party arrangements; test resilience; and maintain contractual and exit capabilities [27]. GDPR continues to govern personal-data processing and automated decision issues where applicable [28]. AI classification should therefore connect to existing ICT, outsourcing, privacy, conduct and resilience registers.

4.5 International coordination and standards

The FSB's June 2026 consultation proposes 12 sound practices. The first four address organisation-wide governance, the next six address lifecycle risk management and the final two address cyber, ICT and third-party risk [10]. The document states that the practices are not intended to create an international standard or prescribe whether institutions should adopt AI [10].

Basel Committee materials link digitalisation to governance, model risk, third-party dependency, operational risk and resilience [12-14]. BIS FSI research observed that most financial authorities address AI through existing technology-neutral frameworks while identifying gaps around governance, skills, model risk, data, third parties and new business models [15].

NIST AI RMF organises outcomes under Govern, Map, Measure and Manage [29]. Its Generative AI Profile identifies risks and suggested actions specific to or amplified by generative AI [30]. ISO/IEC 42001 specifies an AI management-system standard [33]. ISO/IEC 23894 provides AI-risk-management guidance [34]. These are useful control-design references. Their use does not establish legal compliance.

4.6 Applicability matrix

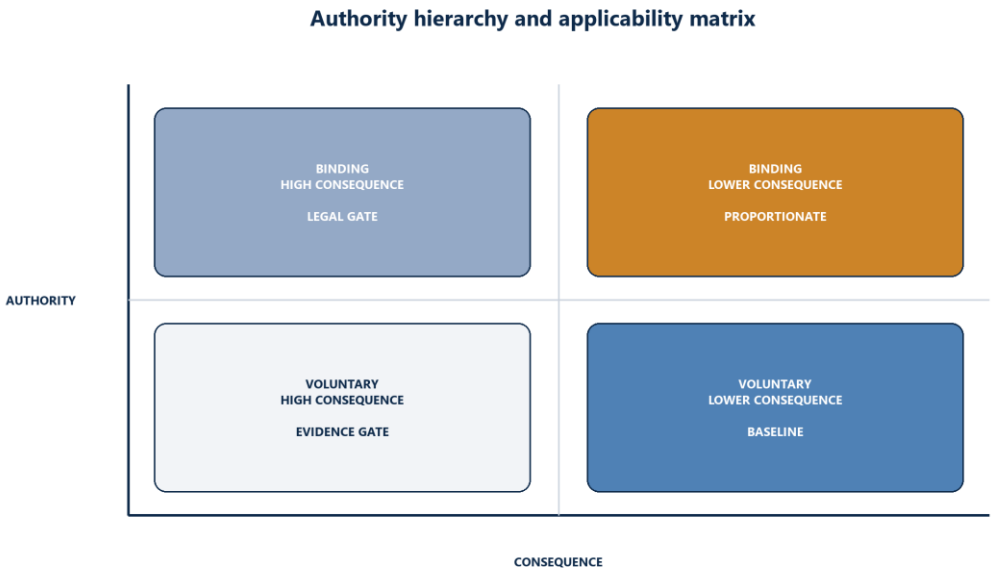

Figure 3. Authority hierarchy and applicability matrix

Source: Matchpoint synthesis [1-35]. Applicability requires qualified legal confirmation.

Source	Status at 1 Aug 2026	Illustrative relevance	Required owner action
EU AI Act and Omnibus	Binding EU law; staged application	EU providers, deployers or affected activity	Legal applicability and role assessment.
DORA	Binding EU regulation for in-scope financial entities	ICT and AI service dependency	Integrate AI services into ICT and third-party registers.
CBUAE AI guidance	Regulatory guidance for UAE licensed financial institutions	Consumer-facing or market-conduct AI	Map transparency, bias, accountability and privacy controls.
DFSA expectations	Supervisory communication for DIFC authorised firms	DIFC AI use	Evidence risk management and accountable oversight.
PRA SS1/23	Supervisory statement for in-scope banks	Models used in decisions	Map inventory, governance, development, validation and mitigants.
US interagency MRM	Supervisory guidance with stated scope	In-scope banking models	Tailor model-risk practices; preserve GenAI and agentic exclusion.

Source	Status at 1 Aug 2026	Illustrative relevance	Required owner action
FSB sound practices	Consultation	Directional cross-border practice	Track finalisation; use only as proposed practice.
NIST and ISO	Voluntary framework and standards	Control architecture	Adopt selected outcomes with an evidence map.

5. INVENTORY AND CLASSIFICATION

5.1 One perimeter, multiple taxonomies

An institution should maintain one discoverable AI and model register with linked classifications. Separate unconnected inventories for models, SaaS, end-user computing, privacy, vendors and operational processes create gaps. The master record should link to specialist registers rather than duplicate every field.

Inventory field	Purpose
Unique ID and version	Prevent ambiguity across deployments and changes.
Business purpose and process	Connect technology to an accountable outcome.
Owner, operator and users	Establish responsibility and competence.
Provider and hosting	Identify internal and third-party dependency.
Model or service components	Reveal foundation models, tools, retrieval and deterministic services.
Data sources and rights	Establish provenance, purpose, sensitivity and transfer.
Outputs and affected parties	Identify decision, conduct and rights exposure.
Authority level	Record whether the system reads, drafts, recommends, approves or acts.
Jurisdictions and entities	Route applicability assessment.
Materiality tier	Determine control depth and escalation.
Evaluation and approval	Link evidence, thresholds, conditions and expiry.
Monitoring and incidents	Maintain current operating risk.
Change history	Trigger proportionate re-evaluation.
Exit and fallback	Prove recoverability and portability.

5.2 Discovery routes

Inventory completeness needs multiple detection routes. Self-attestation alone is weak because embedded features and shadow use can be invisible to central teams.

Discovery route	Examples	Evidence
Procurement	Contracts, renewals, vendor questionnaires	Product, features, provider and data terms.
Identity and access	Application catalogue, single sign-on, privileged access	Users, roles and service access.
Network and cloud	Domains, application programming interfaces, cloud services	Service interaction and data flow.
Expense and card	Individual subscriptions	Shadow procurement candidates.
Code and model platforms	Repositories, registries, endpoints	Internally built and configured components.
Business process review	Research, investment, risk, compliance and operations	Embedded use and decisions.
User declaration	Periodic attestation and campaign	Unmanaged use with accountability.
Vendor change notice	Release notes and feature toggles	Newly embedded AI functionality.
Incident and support data	Service tickets, data events and control exceptions	Previously unregistered use.

5.3 Materiality model

Materiality should be assessed before controls are selected. The proposed score uses eight dimensions, each rated from zero to four. The score supports judgement; it does not replace it.

Dimension	Low end	High end
Decision consequence	Convenience or formatting	Capital, client, legal, regulatory or rights outcome
Authority	Read-only or draft	Execute, communicate, transfer or bind
Population and scale	Few internal users	Large or externally affected population
Data sensitivity	Public or synthetic	Personal, confidential, market-sensitive or restricted
Reversibility	Immediate correction	Irreversible or costly remediation
Explainability need	No material reliance	Decision must be reconstructed or explained
Dependency	Easy substitute	Concentrated provider or critical integration
Change velocity	Fixed and controlled	Provider-driven or adaptive behaviour

Proposed initial bands are 0-7 for Tier 1, 8-15 for Tier 2, 16-23 for Tier 3 and 24-32 for Tier 4. These bands are management design parameters, not regulatory thresholds. A mandatory override should elevate uses involving legal rights, material client communications, autonomous value movement, regulatory reporting, safety, credit or suitability decisions, or critical service dependencies.

5.4 Authority classification

Authority frequently drives risk more directly than model type.

Level	Capability	Default control position
A0	Search or retrieve approved information	Source controls and access logging.
A1	Draft or summarise	Human review before external or decision use.
A2	Recommend or rank	Defined decision owner, reason record and override.
A3	Prepare a transaction or instruction	Segregation, limit and independent release.
A4	Execute within a bounded mandate	Explicit delegation, allow list, dual controls where material, monitoring and kill switch.
A5	Self-directed planning across tools	Exceptional approval, sandboxing, least privilege, transaction limits, continuous monitoring and tested containment.

5.5 Classification output

The classification record should state facts, conclusions and uncertainty separately.

Record component	Example format
Facts	System X summarises investment documents using provider Y; data is hosted in region Z.
Legal classification	Legal owner conclusion, source, date and scope.
Model-risk classification	MRM owner conclusion and policy basis.
AI-risk tier	Score, overrides and rationale.
Conditions	Human review, approved corpus, prohibited data and authority limit.
Residual uncertainty	Provider training-data details unavailable; compensating restriction applied.

6. GOVERNANCE AND DECISION RIGHTS

6.1 Accountable ownership

Each registered use needs a business owner who owns the outcome and residual risk. Technology ownership does not substitute for business accountability. Model risk, compliance, privacy and security owners provide independent or specialist challenge within their mandates.

Role	Core accountability
Board or governing body	Risk appetite, oversight, major exposures and challenge.
Senior accountable executive	Organisation-wide framework, resources and escalation.
Business owner	Purpose, outcome, users, controls, residual risk and value.
System owner	Technical design, configuration, access, operation and change.
Model-risk owner	Classification, validation standards, issues and aggregation.
Compliance and legal	Applicability, conduct, regulatory obligations and legal risk.
Privacy and data	Purpose, rights, minimisation, transfers, retention and data governance.
Information security	Threat model, testing, access, logging and incident response.
Procurement and third-party risk	Due diligence, contract, concentration, monitoring and exit.
Internal audit	Independent assurance over design and operating effectiveness.

6.2 Proportional forum design

A2 institutions may use an existing investment, risk or operating committee with a standing AI agenda. A1 institutions may use a central AI risk committee with business-level approvals. The required outcome is clear authority and documented challenge.

Decision	Minimum forum	Required record
Register and classify	Business owner plus risk coordinator	Inventory record and tier rationale.
Approve low-risk internal use	Delegated owner	Conditions, expiry and monitoring.
Approve material use	Cross-functional risk forum	Evidence pack, challenge, decision and residual risk.
Accept significant exception	Senior accountable executive	Exception, compensating controls, duration and remediation.
Approve autonomous authority	Governing or expressly delegated forum	Authority boundary, limits, containment and accountability.
Retire or suspend	Business and system owners; risk notification	Trigger, service plan, evidence preservation and transition.

6.3 Policy architecture

The institution should avoid a stand-alone AI policy that conflicts with established obligations. A concise AI policy should define perimeter, principles, roles, prohibited uses, approval, monitoring, incidents and exceptions. Supporting standards should connect to model risk, data, privacy, cybersecurity, outsourcing, operational resilience, conduct, records, change and audit.

6.4 Risk appetite

Risk appetite should specify prohibited activities, quantitative or categorical limits, approval authorities and escalation. Examples include:

Appetite statement	Evidence measure
No unapproved confidential data enters public AI services.	Data-loss events, blocked attempts and approved exceptions.
No material investment decision relies solely on generated output.	Decision records with named human authority and source evidence.
No autonomous system moves value outside approved limits.	Tool permissions, transaction limits, release logs and incidents.
Tier 4 systems operate only within current approval.	Approval expiry, overdue conditions and deployment state.
Critical vendor dependency has a tested fallback.	Exit test date, recovery time and unresolved gaps.

6.5 Competence and challenge

Training should match role. General users need data, source, hallucination, phishing, prompt-injection and escalation awareness. Owners need classification, evidence and incident responsibilities. Validators need technical and domain competence. Boards need enough understanding to challenge risk concentration, autonomy, limitations and resourcing

[9,10]. Training completion is weak evidence on its own; scenario testing and observed behaviour provide stronger evidence.

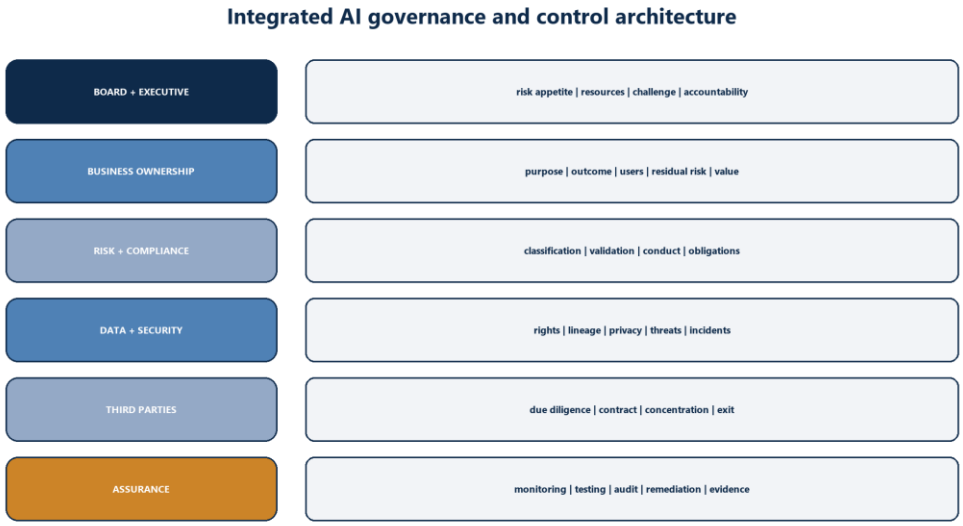

Figure 4. Integrated AI governance and control architecture

Source: Matchpoint control architecture; regulatory and standards context [1-37].

7. LIFECYCLE AND EVIDENCE GATES

7.1 Lifecycle

The proposed lifecycle contains ten stages: idea, registration, classification, design, development or procurement, evaluation, approval, controlled deployment, monitoring and retirement. Iteration returns to the appropriate earlier gate when the model, prompt, data, tool, purpose, population or provider changes.

Gate	Decision question	Minimum evidence
G0 Intake	Is the use permissible to explore?	Purpose, owner, data class and sandbox conditions.
G1 Classification	What governs the use?	Inventory, authority, materiality and applicability.
G2 Design	Are controls designed into the workflow?	Architecture, data flow, threat model and human role.
G3 Build or buy	Is the component understood enough?	Development record or vendor due diligence.
G4 Evaluation	Does it meet defined thresholds?	Representative tests, limitations and independent challenge.
G5 Approval	Is residual risk accepted by the right authority?	Decision pack, conditions, expiry and accountable sign-off.
G6 Deployment	Can release be controlled and reversed?	Access, change, fallback, logging and runbook.
G7 Operation	Does current evidence remain inside appetite?	Monitoring, incidents, exceptions, drift and user feedback.
G8 Change	Does the change require re-evaluation?	Change classification, regression and approval.
G9 Retirement	Can dependency end safely?	Export, archive, access removal, vendor termination and lessons.

7.2 Design evidence

The design pack should show the full system, including the user's task, data, model, retrieval, deterministic controls, tools, interfaces, human review and downstream action. A model card without workflow context is incomplete for a material AI use.

Design artefact	Required content
Purpose statement	Intended outcome, user, population, exclusions and prohibited use.
Architecture	Components, providers, versions, data flows, tools and boundaries.

Design artefact	Required content
Rights map	Data owners, legal basis, licences, confidentiality and transfers.
Decision map	Recommendations, approvals, execution and communication authority.
Failure analysis	Hazards, causes, controls, detection, consequence and recovery.
Threat model	Assets, actors, attack paths, controls and residual risk.
Evaluation plan	Population, metrics, thresholds, samples and challenger.
Operating model	Owners, monitoring, changes, incidents, fallbacks and support.
Exit plan	Portability, replacement, data deletion, continuity and cost.

7.3 Approval pack

The approval pack should answer one question: does the available evidence support the proposed bounded use? It should identify gaps rather than convert absence of evidence into a positive conclusion.

Section	Approval content
Executive decision	Approve, approve with conditions, pilot, redesign, pause or reject.
Applicability	Entity, jurisdiction, activity and authoritative sources.
Classification	Model, AI, authority, materiality and criticality.
Evidence	Design, data, evaluation, security, privacy and vendor results.
Limitations	Known gaps, unsupported populations and prohibited use.
Residual risk	Severity, likelihood, owner and acceptance authority.
Conditions	Action, owner, due date, evidence and consequence of breach.
Monitoring	Metrics, limits, frequency and escalation.
Expiry	Date or change trigger requiring renewal.

7.4 Change control

AI services can change through model versions, prompts, system instructions, retrieval content, tools, policies, safety filters, providers, data, user populations and hosting. The change standard should classify each change and define required regression.

Change class	Example	Default action
Routine	Non-material interface correction	Record and test affected function.
Moderate	Prompt or retrieval update	Targeted regression and owner approval.
Material	Model, provider, purpose, authority or population change	Reclassification, independent evaluation and approval.
Emergency	Security mitigation or service withdrawal	Contain, document, approve temporary state and complete retrospective review.

7.5 Retirement

Retirement should remove access and dependency while preserving required records. For critical systems, the institution should test data export, replacement operation and recovery before approving production. Vendor termination rights without an operational transition plan are weak evidence.

8. VALIDATION AND EVALUATION

8.1 Validation objective

Validation provides effective challenge over conceptual soundness, implementation, performance, use and limitations. The depth depends on materiality, complexity and exposure [1,4]. The validator should be independent enough to challenge the owner and competent across the relevant domain, data, model and workflow.

8.2 Evaluation by system type

System type	Core evaluation
Deterministic rules	Requirement coverage, boundary cases, reconciliation and change tests.
Predictive model	Discrimination, calibration, stability, sensitivity, bias where relevant and benchmark.
Generative assistant	Factual support, completeness, harmful output, robustness, privacy, source use and human acceptance.
Retrieval-augmented system	Retrieval recall, precision, permissions, freshness, citation fidelity and answer support.
Code generator	Correctness, security, dependency, licence, test coverage and reviewer competence.
Agentic workflow	Goal compliance, tool selection, authority, state, recovery, attack resistance and cumulative error.
Vendor black box	Outcome testing, limitations, change controls, monitoring and compensating restrictions.

8.3 Representative evaluation set

The evaluation population should reflect normal, difficult, adverse and prohibited cases. It should include material subgroups and periods, known failure modes, boundary conditions and out-of-distribution examples. Test-set governance should prevent contamination and untracked changes.

Evaluation stratum	Purpose
Routine	Establish normal performance and unit economics.
Complex	Test long, ambiguous, cross-document or exception-heavy cases.
Adverse	Test misleading, conflicting or malicious inputs.
Sensitive	Test confidentiality, personal data and restricted information.
Authority boundary	Test attempted prohibited action or escalation.
Temporal	Test current and older regimes, terms or market states.
Subgroup	Test legally or operationally relevant populations.
Recovery	Test timeout, tool failure, corrupted state and provider outage.

8.4 Metrics

A single accuracy percentage is insufficient for most material AI uses.

Metric	Definition
Material factual-support rate	Outputs with every material claim supported by an approved source divided by reviewed outputs.
Material omission rate	Outputs omitting at least one required material item divided by reviewed outputs.
Unsupported-action rate	Attempts to act outside delegated authority divided by relevant tests.
First-pass acceptance	Outputs accepted without material correction divided by reviewed outputs.
Severe-defect rate	Outputs with a defect meeting the severity definition divided by reviewed outputs.
Escalation precision	Correct escalations divided by all escalations.
Escalation recall	Correct escalations divided by all cases requiring escalation.
Citation fidelity	Citations that support the associated claim divided by citations reviewed.
Recovery success	Failure scenarios restored within defined limits divided by scenarios tested.
Human override rate	Decisions changed by authorised reviewers divided by AI-assisted decisions.

Thresholds must be tied to consequence. A drafting assistant and an automated client instruction should not share the same tolerance. For zero-tolerance event types, the evaluation objective should be absence in the tested population plus strong preventative and detective controls; it should not be described as proof that the event cannot occur.

8.5 Generative AI evaluation

The NIST Generative AI Profile identifies risks including confabulation, data privacy, harmful bias, human-AI configuration, information integrity, information security, intellectual property and value-chain dependency [30]. Evaluation should cover the system as configured, including prompts, retrieval, tools, safety controls and workflow. Provider benchmark scores cannot establish fitness for a local investment or regulated-finance use.

8.6 Agentic evaluation

Agentic systems add sequence, state, tool and authority risk. A step may be individually acceptable while the sequence produces an unauthorised outcome. Evaluation should record each action, tool input, tool output, permission decision, human intervention and final state.

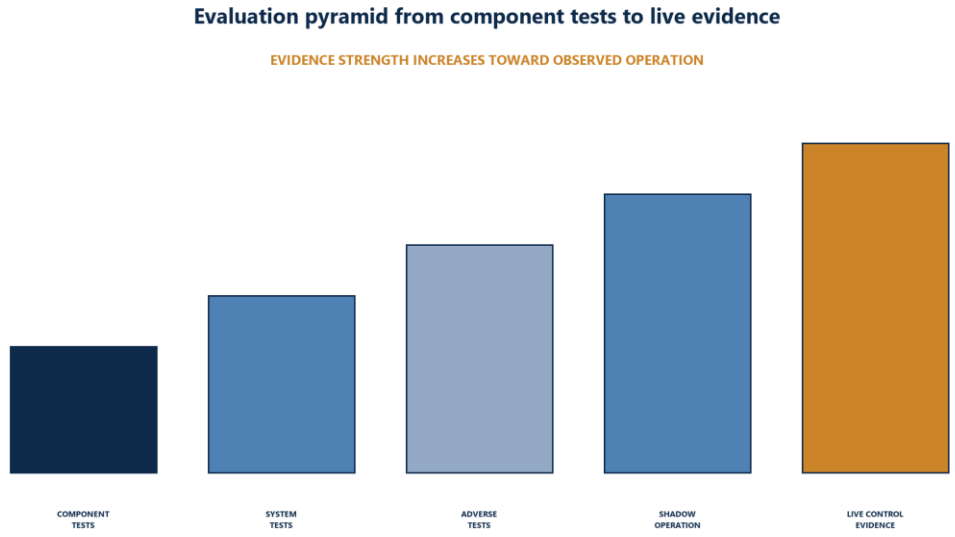

Figure 5. Evaluation pyramid from component tests to live control evidence
Source: Matchpoint evaluation framework; model-risk and AI-risk context [1,4,29-32,36].

Agent test	Question
Goal adherence	Does the system remain within the approved objective?
Tool allow list	Does it select only approved tools for the task?
Parameter limits	Are amounts, recipients, destinations and data bounded?
Prompt injection	Can external content override system policy or exfiltrate data?
Cumulative error	Do small intermediate errors create a material final outcome?
State integrity	Can stale or corrupted state change action?
Human checkpoint	Is approval requested at the right stage with complete information?
Kill and recovery	Can operation stop safely and resume from a known state?

8.7 Independent challenge record

The validation report should state scope, methods, evidence, limitations, findings, severity, conditions and conclusion. A qualified conclusion can support a bounded pilot. Missing evidence should remain a limitation or issue. Management acceptance should name the accountable owner and authority.

9. DATA, PRIVACY, SECURITY AND THIRD PARTIES

9.1 Data governance

AI controls inherit from the data lifecycle. Each material use should establish provenance, ownership, permission, purpose, quality, lineage, minimisation, retention, transfer, deletion and incident response.

Data question	Evidence
May the institution use the data for this purpose?	Contract, legal basis, consent or other authorised basis.
May the provider process or retain it?	Data-processing terms, configuration and verified service behaviour.
Is the corpus complete and current?	Source register, reconciliation, refresh and exception logs.
Can users access only authorised material?	Identity, permissions, retrieval filters and adversarial tests.
Can output reveal sensitive input?	Privacy testing, logging, redaction and incident process.
Can the data be exported and deleted?	Portability test, deletion evidence and contract.

9.2 Secure development and operation

NIST CSF 2.0 and the NCSC secure-AI guidance provide governance and lifecycle security references [31,32]. OWASP's LLM application risk work provides a technical threat catalogue [36]. The institution should select controls from its applicable security framework and record their implementation.

Threat	Illustrative control objectives
Prompt injection	Treat external content as untrusted; isolate instructions; restrict tools and data.
Sensitive disclosure	Minimise data, enforce access, filter output, monitor and respond.
Supply-chain compromise	Verify components, provenance, dependencies, updates and provider controls.
Data or model poisoning	Control contribution, validate sources, detect anomalies and preserve rollback.
Excessive agency	Least privilege, transaction limits, approvals, allow lists and kill switch.
Insecure output use	Validate before execution, encode safely and segregate duties.
Model theft or extraction	Access controls, rate limits, monitoring and response.
Denial or provider outage	Capacity plan, fallback, recovery and alternate process.

9.3 Third-party AI

Financial institutions remain responsible for risks created by services they procure or rely on. DORA, Basel materials, FSB work, FINMA guidance and UAE regulator materials all reinforce the importance of third-party dependency, governance and resilience [10,12,15-17,27,37].

Due-diligence domain	Evidence request
Service definition	Components, versions, hosting, subprocessors and support.
Data	Collection, use, retention, training, location, transfer and deletion.
Model governance	Development, evaluation, limitations, monitoring and change.
Security	Control assurance, incident history, testing and notification.
Resilience	Availability, recovery, capacity, continuity and dependency map.
Auditability	Logs, evidence access, audit and supervisory cooperation.
Legal and rights	Licence, intellectual property, confidentiality and prohibited use.
Change	Notice, version control, regression evidence and customer options.
Exit	Export, assistance, deletion, transition time and fees.

9.4 Concentration and fourth parties

Foundation models, cloud infrastructure, data services and specialist tooling can create common dependencies across apparently different applications. The inventory should identify ultimate providers and material subprocessors. Scenario analysis should examine simultaneous failure, degraded service, unilateral term change, region outage, model withdrawal and loss of audit access.

9.5 Contract control schedule

The contract should define service scope, approved data use, confidentiality, security, incident notice, business continuity, audit, regulator access where applicable, subcontracting, model changes, performance, intellectual property, record retention, deletion, exit assistance and liability allocation. Contract language must be reviewed by qualified counsel. A right without operational testing remains incomplete evidence.

10. HUMAN OVERSIGHT, CONDUCT AND ACCOUNTABILITY

10.1 Meaningful human oversight

Human review is meaningful when the reviewer has competence, time, authority, information and an effective ability to challenge or stop the outcome. A mandatory click with no supporting evidence is weak control.

Review condition	Evidence
Competence	Role-specific training, experience and assessment.
Information	Source material, system output, limitations and uncertainty.
Time	Workflow capacity and service-level design.
Authority	Ability to reject, amend, escalate or stop.
Independence	Separation where consequence requires it.
Traceability	Named reviewer, time, action, reason and final outcome.
Feedback	Corrections routed into monitoring and improvement.

10.2 Automation bias and workload

Oversight can fail when users defer to polished output, review too many items, or lack access to source evidence. Monitoring should measure correction, override, review time, exception and disagreement patterns. A very low override rate may indicate high quality; it may also indicate weak challenge. Investigation requires sampled evidence.

10.3 Fairness and affected parties

Where AI affects access, pricing, suitability, employment, credit, insurance or other consequential treatment, the institution should identify legally and operationally relevant groups, evaluate differential outcomes, document legitimate objectives, provide appropriate explanation and maintain contest or escalation routes. Applicable law determines specific obligations. OECD principles, NIST outcomes, CBUAE guidance and FINMA expectations provide supporting governance references [16,29,35,37].

10.4 Records and explanations

The record should preserve the input, relevant source version, system configuration, output, reviewer action, final decision and reason at the level required for reconstruction. Explanation should be appropriate to the audience and decision. Technical interpretability does not automatically satisfy client or legal explanation requirements.

10.5 Prohibited and restricted uses

Each institution should define prohibited uses based on law, mandate and risk appetite. Restricted uses can operate only under stated conditions. Examples for management consideration include unapproved confidential-data entry, unsupervised legal or investment advice, autonomous transfer of value beyond limits, fabricated source attribution, covert client interaction, and use of personal data outside approved purpose. These are proposed policy examples; legal owners must establish the actual prohibitions.

11. MONITORING, INCIDENTS AND ASSURANCE

11.1 Monitoring design

Monitoring should connect system behaviour to risk appetite and approval conditions.

Monitoring domain	Example measure	Escalation trigger
Performance	Material support, omission and severe-defect rates	Threshold breach or adverse trend.
Use	Volume, users, processes and out-of-scope attempts	Unapproved population or purpose.

Monitoring domain	Example measure	Escalation trigger
Authority	Tool calls, limits, approvals and blocked actions	Any unauthorised or unexplained action.
Data	Freshness, completeness, permissions and sensitive events	Stale critical source or access failure.
Change	Provider, model, prompt, retrieval and configuration changes	Material change without approval.
Security	Injection, exfiltration, vulnerabilities and anomalous access	Defined severity event.
Conduct	Complaints, overrides, affected outcomes and explanations	Material client or rights impact.
Resilience	Availability, latency, recovery and fallback	Critical-service limit breach.
Vendor	Service, assurance, incidents and concentration	Evidence lapse or control deterioration.

11.2 Incident taxonomy

An AI incident can be a model, data, security, privacy, conduct, operational, legal or third-party event. The incident process should avoid routing everything to a specialist AI team. The severity assessment should consider actual and plausible consequence, exposure, containment, regulatory notification, affected parties and recurrence.

Incident stage	Required action
Detect	Preserve alert, output, input, logs, versions and context.
Contain	Stop or restrict the use, tools, access or data flow.
Assess	Determine consequence, population, obligations and root causes.
Notify	Follow legal, regulatory, contractual and internal escalation.
Remediate	Correct outcome, system, control and affected records.
Recover	Restore within approved conditions and validate stability.
Learn	Update evaluation, monitoring, training, risk and vendor management.

11.3 Control assurance

First-line testing should confirm controls operate. Second-line monitoring should challenge classification, approval, issues and aggregate risk. Internal audit should assess framework design and operating effectiveness using a risk-based plan. External assurance may support vendor or technical evidence, but the institution must assess scope, period, exclusions and relevance.

11.4 Aggregate risk

Portfolio reporting should identify concentration by provider, model, cloud, data source, business process, geography and authority level. Individually low-tier assistants can become material when they share a provider, process a large volume, or embed common error into multiple decisions.

12. ILLUSTRATIVE A2 OPERATING CASE

12.1 Case definition

[Unverified] A family-office investment team is assumed to use a retrieval-augmented assistant to summarise private-market fund documents, compare material terms and draft investment-committee questions. The system has no authority to approve investments, communicate with managers or transmit instructions. All numerical values and operating assumptions in this section are unverified illustrations.

Item	[Unverified] illustrative assumption
Users	12 investment and operations professionals
Annual document sets	180
Typical source length	120-450 pages
Output	Source-linked summary, term table and question draft
Authority	A1 draft only

Item	[Unverified] illustrative assumption
Data	Confidential manager and fund documents
Provider	Hosted foundation model plus controlled retrieval
Decision	Human investment team and investment committee retain authority

12.2 Classification

The proposed materiality assessment is Tier 3 because the use handles confidential data and can influence investment analysis, while output remains draft and reversible. A mandatory override would apply if the output entered committee records without review, generated legal conclusions, or initiated external communications.

12.3 Evaluation design

Test set	[Unverified] illustrative size	Acceptance focus
Routine fund documents	120	Material-term support and completeness
Complex structures	40	Waterfalls, side letters, conflicts and exceptions
Conflicting sources	25	Source hierarchy and escalation
Access-control cases	25	Retrieval isolation across entities and deals
Prompt-injection cases	30	Instruction separation and data protection
Historical documents	20	Version and effective-date handling

[Unverified] Illustrative proposed thresholds are 100 per cent support for defined material terms in the tested set, zero cross-deal access failures, zero autonomous actions, at least 98 per cent citation fidelity and documented escalation for every unresolved source conflict. These values are proposed management assumptions and require approval.

12.4 Control design

Control objective	Proposed control
Complete sources	Reconcile uploaded corpus to the authorised deal-room manifest.
Correct access	Enforce matter-level permissions before retrieval.
Traceable output	Require clause-level citation for material terms.
Human authority	Block external communication and record named review.
Legal boundary	Route legal interpretation and unresolved terms to authorised counsel.
Change	Regress after model, retrieval, prompt or provider change.
Incident	Preserve input, corpus, version, output, reviewer and downstream use.
Exit	Export corpus metadata, evaluation set, logs and approved outputs.

12.5 Economic boundary

[Unverified] The use may release analyst capacity. Released time is not a cash saving unless an approved financial consequence is observed. Attributed Matchpoint or client revenue, cash cost reduction and loss reduction remain USD 0. A value ledger can record accepted outcomes, review time, rework and full operating cost without converting theoretical hours into realised benefit.

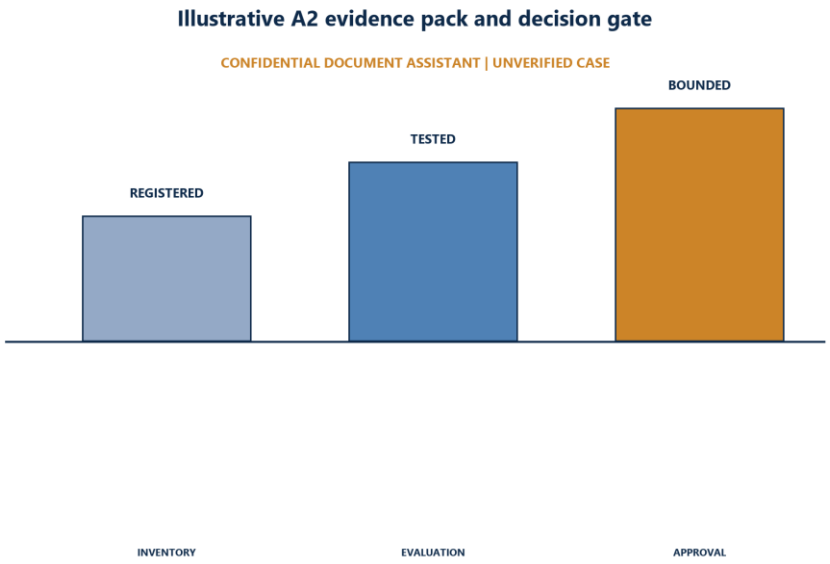

Figure 6. Illustrative A2 evidence pack and decision gate

Source: Unverified illustrative management assumptions; Matchpoint framework.

13. ILLUSTRATIVE A1 OPERATING CASE

13.1 Case definition

[Unverified] An institutional allocator is assumed to use a predictive and generative system to monitor external managers, flag reporting anomalies and draft due-diligence follow-up questions. The model can prioritise review; it cannot change an allocation, issue a breach notice or communicate externally. All numerical inputs and thresholds in this section are unverified illustrations.

Item	[Unverified] illustrative assumption
External managers	85
Funds and mandates	140
Reporting frequency	Monthly or quarterly by mandate
Inputs	Manager reports, administrator data and internal exposure records
Output	Risk flags, supporting evidence and draft questions
Authority	A2 recommend or rank
Material consequence	Review priority and potential escalation

13.2 Model and workflow boundary

The predictive component detects anomalies against approved features and benchmarks. The generative component explains the flag using controlled records and drafts questions. A deterministic rules layer enforces exposure limits and required escalations. The investment-risk owner decides whether to escalate. This separation allows different evaluation methods and reduces authority assigned to probabilistic output.

13.3 Evaluation

Layer	Proposed evidence
Data	Reconciliation to administrator and internal books; lateness and correction history.
Predictive	Back-testing, stability, sensitivity, benchmark, false-positive and missed-event review.
Generative	Claim support, omission, source conflict, explanation and prohibited-content tests.
Workflow	Correct routing, owner decision, override, service levels and record retention.
Resilience	Provider outage, stale data, fallback report and recovery test.

Layer	Proposed evidence
Vendor	Change notice, model dependency, audit evidence, subprocessors and exit.

[Unverified] An illustrative pilot uses 18 months of historical records, 60 adjudicated events and a three-month shadow period. These values are management assumptions. The institution would need to assess whether the sample supports the intended claims and material subgroups.

13.4 Decision and escalation

System output	Human decision	Prohibited system action
Low-confidence anomaly	Request analyst review	Suppress or close without review
Supported material anomaly	Escalate to investment risk	Notify manager externally
Conflicting source data	Route for reconciliation	Select one source silently
Limit-rule breach	Immediate mandatory escalation	Alter threshold or exposure record
Provider or model change	Suspend affected output pending regression	Continue under stale approval

13.5 Portfolio and third-party aggregation

The allocator should aggregate common providers across manager monitoring, research, reporting and client-service tools. A single foundation-model or cloud dependency may support multiple nominal vendors. The board pack should show this look-through concentration and the tested recovery path.

13.6 Economic boundary

[Unverified] No financial benefit is attributed in the illustrative case. A later evaluation could measure review throughput, accepted flags, material omissions, false escalation, time to resolution and operating cost. Revenue, cash-cost and loss claims require approved observed evidence and a defensible counterfactual.

14. BOARD PACK, ROADMAP AND IMPLEMENTATION

14.1 Board information

Board reporting should support direction and challenge. It should separate exposure from activity and report exceptions plainly.

Board view	Decision evidence
Portfolio	Uses by tier, authority, business, provider and status.
Consequence	Client, capital, regulatory, rights and critical-service exposure.
Assurance	Evaluation current, overdue validation, audit and open findings.
Risk appetite	Breaches, exceptions, conditions and remediation.
Incidents	Severity, affected population, cause, containment and recurrence.
Change	Material provider, model, purpose and authority changes.
Concentration	Foundation model, cloud, data and fourth-party dependency.
Resilience	Fallback coverage, exit tests, recovery gaps and critical services.
People	Competence, capacity, challenge and key-person dependency.
Value	Approved observed benefits, full cost and unverified hypotheses separated.

14.2 Twelve-month roadmap

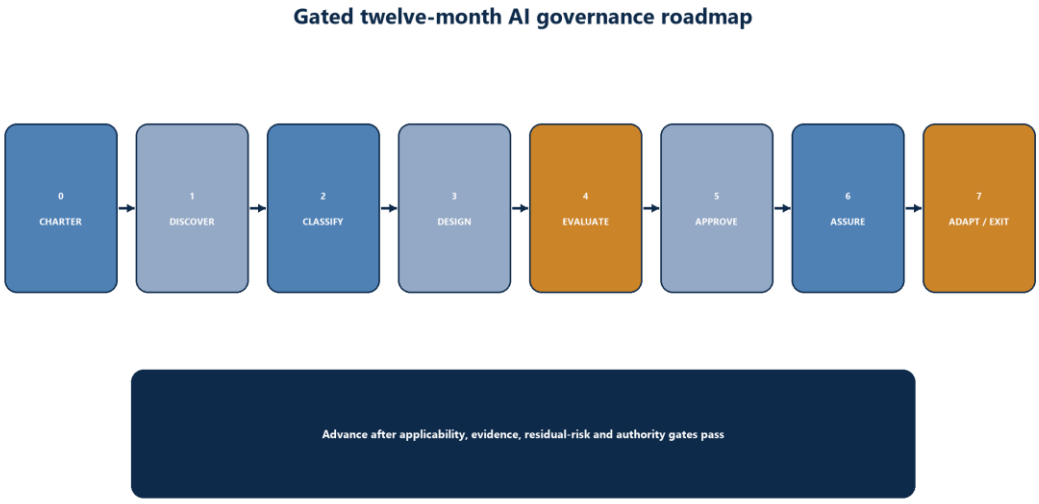

Figure 7. Gated twelve-month AI governance roadmap

Source: Matchpoint implementation framework.

Period	Core outcomes	Gate
Days 0-30	Sponsor, perimeter, interim prohibited uses, discovery and incident route	Framework charter approved.
Days 31-60	Inventory, authority map, materiality method and obligations register	High-risk unknowns escalated.
Days 61-90	Lifecycle, evidence templates, vendor controls and monitoring minimum	Pilot governance approved.
Months 4-6	Evaluate material uses, remediate access, logging, data and contracts	Tier 3 and 4 decisions current.
Months 7-9	Concentration analysis, resilience and exit tests, board reporting	Critical dependencies tested.
Months 10-12	Independent assurance, lessons, policy refresh and next-year plan	Residual risks accepted or remediated.

14.3 First 30 days

The first month should establish authority and visibility. Name the accountable executive, issue interim data and authority restrictions, create a single intake route, reconcile known applications and subscriptions, and connect AI events to existing incident management. Prioritise uses that can affect clients, capital, reporting, rights, confidential data or movement of value.

14.4 Days 31-90

Build the obligations register from applicable legal and supervisory sources. Complete classification for material uses. Create evaluation, approval, monitoring, change and retirement templates. Review critical vendors and embedded AI features. Establish an exception process with expiry and compensating controls.

14.5 Months four to twelve

Evaluate and approve or suspend material uses. Test access isolation, failure recovery and vendor exit. Reconcile provider concentration. Implement board metrics and independent assurance. Refresh the framework when the FSB consultation is finalised, NIST updates AI RMF, or jurisdictional rules change.

14.6 Control maturity model

Level	Description	Evidence
0 Uncontrolled	Uses are unknown or unmanaged	Incidents and ad hoc declarations.
1 Visible	Inventory and interim restrictions exist	Registered use and accountable owner.

Level	Description	Evidence
2 Defined	Classification, lifecycle and standards exist	Approved policy, templates and roles.
3 Operating	Material controls are implemented and monitored	Tests, approvals, logs and issues.
4 Assured	Independent challenge and portfolio evidence operate	Validation, audit, board challenge and remediation.
5 Adaptive	Changes, incidents and external developments update controls promptly	Measured learning, timely refresh and tested resilience.

15. LIMITATIONS, RESEARCH AGENDA AND CONCLUSION

This paper synthesises cross-jurisdictional materials; it does not establish applicability for a particular entity. Requirements can overlap and change. The FSB sound practices were consultative at the evidence cut-off. The EU timetable changed shortly before publication. The US revised model-risk guidance expressly excludes generative and agentic AI from scope. Standards and voluntary frameworks do not establish legal compliance.

The operating framework requires empirical testing. Future research should compare classification consistency across institutions; measure validation reliability for generative and agentic workflows; evaluate whether citation and human-review controls reduce material errors; assess fourth-party concentration; and test exit plans under provider withdrawal, model change and cyber disruption. Institutions should publish or share incident taxonomies and evaluation methods where confidentiality permits.

The central conclusion is that controlled AI adoption is an evidence problem. A complete inventory shows exposure. Materiality and authority determine proportional control. Evaluation establishes bounded fitness. Governance assigns the decision. Monitoring keeps approval current. Incident and exit capability contain failure. For A2 family offices and A1 institutional allocators, this structure supports innovation while preserving investment authority, confidentiality, traceability and resilience.

Attributed Matchpoint or client revenue, cash cost reduction and loss reduction remain USD 0 until approved observed evidence supports attribution.

DECLARATIONS

Author attribution. Prepared by Matchpoint Partners. Named-person author attribution is pending CK approval.
Evidence cut-off. Sources were reviewed through 1 August 2026.
Scenario status. Every worked case, threshold, volume, time, rate and economic input labelled unverified is an illustrative management assumption. It is not observed client or Matchpoint evidence.
Financial attribution. Attributed Matchpoint or client revenue, cash cost reduction and loss reduction remain USD 0 until approved observed evidence exists.
Professional boundaries. Research paper only. This document is not investment, legal, regulatory, accounting, audit, tax, privacy, cybersecurity, employment, valuation or technology advice.
Conflicts and funding. No external funding or client-specific data was used. Matchpoint Partners has a commercial interest in advisory work related to the subject.

REFERENCES

[1] Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation and Office of the Comptroller of the Currency (2026). SR 26-2: Revised Guidance on Model Risk Management. https://www.federalreserve.gov/supervisionreg/srletters/SR2602.htm
[2] Office of the Comptroller of the Currency (2026). OCC Issues Updated Model Risk Management Guidance. https://occ.gov/news-issuances/news-releases/2026/nr-occ-2026-29.html
[3] Federal Deposit Insurance Corporation (2026). FIL-16-2026: Agencies Revise the Interagency Model Risk Management Guidance. https://www.fdic.gov/news/financial-institution-letters
[4] Prudential Regulation Authority (2026). SS1/23: Model Risk Management Principles for Banks, April 2026. https://www.bankofengland.co.uk/prudential-regulation/publication/2023/may/model-risk-management-principles-for-banks-ss
[5] 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
[6] 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
[7] Bank of England and Financial Conduct Authority (2026). The Bank of England and FCA's 2026 AI Survey. https://www.bankofengland.co.uk/prudential-regulation/regulatory-digest/2026/june-2026
[8] Bank of England (2026). Financial Stability Report, July 2026. https://www.bankofengland.co.uk/financial-stability-report/2026/july-2026
[9] Bank of England, Financial Conduct Authority and HM Treasury (2026). Joint Statement on Frontier AI Models and Cyber Resilience. https://www.bankofengland.co.uk/news/2026/may/boe-fca-and-hm-treasury-joint-statement-on-frontier-ai-models-and-cyber-resilience
[10] Financial Stability Board (2026). Sound Practices for Financial Institutions' Responsible AI Adoption: Consultation Report. https://www.fsb.org/2026/06/fsb-consults-on-sound-practices-for-the-responsible-adoption-of-artificial-intelligence-ai/
[11] Financial Stability Board (2024). The Financial Stability Implications of Artificial Intelligence. https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/
[12] Basel Committee on Banking Supervision (2024). Digitalisation of Finance. https://www.bis.org/bcbs/publ/d575.htm
[13] Basel Committee on Banking Supervision (2021). Principles for Operational Resilience. https://www.bis.org/bcbs/publ/d516.htm
[14] Basel Committee on Banking Supervision (2013). Principles for Effective Risk Data Aggregation and Risk Reporting. https://www.bis.org/publ/bcbs239.htm
[15] Financial Stability Institute (2024). Regulating AI in the Financial Sector: Recent Developments and Main Challenges. https://www.bis.org/fsi/publ/insights63.htm
[16] 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

[17] Central Bank of the UAE, Securities and Commodities Authority, DFSA and FSRA (2021). Guidelines for Financial Institutions Adopting Enabling Technologies. https://www.adgm.com/media/announcements/uae-regulatory-authorities-jointly-issue-guidelines-for-financial-institutions
[18] Dubai Financial Services Authority (2026). DFSA Regulatory Expectations on Artificial Intelligence Risk Management in the DIFC. https://www.dfsa.ae/your-resources/publications-reports/seo-letters-1
[19] Dubai Financial Services Authority (2025). Artificial Intelligence Survey 2025. https://www.dfsa.ae/news/new-dfsa-ai-survey-generative-ai-adoption-has-nearly-tripled-within-difc-last-12-months-governance-continues-develop
[20] Dubai Financial Services Authority (2025). Cyber and Artificial Intelligence Risk in Financial Services: Strengthening Oversight Through International Dialogue. https://www.dfsa.ae/news/new-dfsa-report-explores-regulatory-insights-cybersecurity-artificial-intelligence-and-quantum-risks
[21] Financial Services Regulatory Authority of Abu Dhabi Global Market (current at 2026). Supplementary Guidance: Digital Investment Management. https://assets.adgm.com/download/assets/supplementary-guidance-authorisation-of-digital-investment-management-robo-advisory-activities.pdf/f2ae6f36606c11ef946bcecc50de2f68
[22] 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
[23] 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
[24] 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
[25] European Commission (2026). AI Omnibus Enters into Force. https://digital-strategy.ec.europa.eu/en/news/ai-omnibus-enters-force
[26] European Commission (2026). Guidelines on Transparency Obligations for Providers and Deployers of Certain AI Systems. https://digital-strategy.ec.europa.eu/en/news/commission-publishes-guidelines-transparency-obligations-providers-and-deployers-certain-ai-systems
[27] 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
[28] European Union (2016). Regulation (EU) 2016/679, General Data Protection Regulation. https://eur-lex.europa.eu/eli/reg/2016/679/oj
[29] National Institute of Standards and Technology (2023). Artificial Intelligence Risk Management Framework 1.0. https://www.nist.gov/itl/ai-risk-management-framework
[30] 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
[31] National Institute of Standards and Technology (2024). Cybersecurity Framework 2.0. https://www.nist.gov/publications/nist-cybersecurity-framework-csf-20
[32] 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
[33] International Organization for Standardization (2023). ISO/IEC 42001:2023, Artificial Intelligence Management Systems. https://www.iso.org/standard/42001
[34] International Organization for Standardization (2023). ISO/IEC 23894:2023, Artificial Intelligence: Guidance on Risk Management. https://www.iso.org/standard/77304.html
[35] Organisation for Economic Co-operation and Development (2024). OECD AI Principles. https://oecd.ai/en/ai-principles
[36] 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
[37] Swiss Financial Market Supervisory Authority (2024). FINMA Guidance 08/2024: Governance and Risk Management When Using Artificial Intelligence. https://www.finma.ch/en/news/2024/12/20241218-mm-finma-am-08-24/
[38] Swiss Financial Market Supervisory Authority (2025). FINMA Survey: Artificial Intelligence Gaining Traction at Swiss Financial Institutions. https://www.finma.ch/en/news/2025/04/20250424-mm-umfrage-ki/
[39] Hong Kong Monetary Authority (2024). Research Paper on Generative Artificial Intelligence in the Financial Services Sector. https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2024/20240927e1.pdf
[40] International Organization of Securities Commissions (2025). AI Use Cases in Capital Markets. https://www.iosco.org/library/pubdocs/pdf/IOSCOPD788.pdf

APPENDIX A. MINIMUM AI AND MODEL REGISTER

Field group	Required fields
Identity	ID, name, version, status, owner, operator and provider
Purpose	Use case, process, user, population, output and excluded use
Components	Model, retrieval, tools, deterministic services, interfaces and hosting
Data	Sources, rights, sensitivity, location, transfer, retention and deletion
Decision	Affected outcome, authority level, reviewer, execution and communication
Classification	Legal, model, AI, materiality, criticality and rationale
Evidence	Design, evaluation, validation, security, privacy and third-party reviews
Approval	Decision, authority, conditions, expiry and residual risk
Operation	Monitoring, changes, incidents, issues, exceptions and service levels
Exit	Fallback, portability, transition, deletion, archive and termination

APPENDIX B. MATERIALITY QUESTIONNAIRE

Question	Response evidence
Can output affect capital, client, legal, regulatory or rights outcomes?	Decision and consequence map
Can the system approve, execute, communicate, bind or transfer?	Tool and authority map
What population and value are exposed?	Volume, users, assets and transaction limits
What data enters or can be retrieved?	Data-flow and rights record
Can the outcome be reversed promptly?	Recovery and remediation analysis
Must the decision be explained or reconstructed?	Record and explanation requirement
Which providers and fourth parties are critical?	Dependency and concentration map
How can behaviour change?	Version, configuration, data and provider change map

APPENDIX C. VALIDATION REPORT OUTLINE

Section	Required content
Mandate	Validator, independence, scope, period and exclusions
System	Purpose, architecture, components, data, users and authority
Classification	Materiality, applicable standards and validation depth
Methods	Review, reproduction, benchmark, testing, sampling and limitations
Results	Metrics, thresholds, subgroup, adverse and recovery outcomes
Findings	Severity, evidence, consequence, owner and remediation
Conclusion	Fit, fit with conditions, pilot only, redesign, pause or reject
Follow-up	Conditions, monitoring, expiry and revalidation triggers

APPENDIX D. BOARD DASHBOARD DATA DICTIONARY

Metric	Definition
Registered uses	Active unique AI and model records by approved inventory rule
Tier 3 and 4 exposure	Material uses active, pilot, suspended or overdue
Authority exposure	Uses at A3-A5 and aggregate transaction or communication limits
Current evaluation	Material uses with unexpired approval and required tests complete
Open severe issues	Issues meeting approved severity definition and not closed
Exceptions	Active risk-appetite or policy exceptions with owner and expiry

Metric	Definition
Incidents	Events by severity, consequence, cause and recurrence
Provider concentration	Uses and critical processes dependent on common providers
Exit readiness	Critical services with current successful fallback or exit test
Observed value	Finance-approved realised benefit separated from hypotheses

APPENDIX E. RED-FLAG REGISTER

Red flag	Why it matters	Immediate response
Material AI use has no inventory ID	Ownership and evidence cannot be reconciled	Register, contain and classify
AI policy uses one definition for every regime	Legal and operational perimeters may diverge	Add linked classifications and applicability owners
GenAI is assumed covered by US revised MRM guidance	OCC stated generative and agentic AI are outside scope	Apply adjacent AI-risk controls and preserve the boundary
Consultation language is reported as final	Authority is overstated	Correct status and monitor finalisation
Vendor benchmark substitutes for local evaluation	Local purpose, data and workflow remain untested	Build representative system-level evaluation
Human review is an unrecorded click	Effective challenge cannot be shown	Define competence, evidence, authority and traceability
Agent has broad tools and credentials	Excessive agency increases consequence	Apply least privilege, allow lists, limits and checkpoints
Provider can change models without notice	Approval evidence can become stale	Contract for notice and detect version change
Critical AI has no tested fallback	Service and decision continuity are unproven	Restrict deployment and test recovery
Board pack reports only use-case counts	Consequence and concentration remain hidden	Add tier, authority, provider, incident and exit views
Hours saved are reported as cash benefit	Financial realisation is unproven	Reclassify as capacity until finance-approved evidence exists
Source citations are present but unsupported	Apparent traceability can mislead reviewers	Test citation fidelity and material claim support