EXECUTIVE EDITION

AMAT

Asset Management AI Transformation Guide

A practitioner-built framework for adopting AI safely, compliantly, and at scale

2026 Release Preview Version

(Preliminary Prototype, Subject to Change)

Notice to Readers

This is a preview prototype version of the 2026 release. It is currently being shared for review, feedback, and evaluation purposes. The content, structure, and data contained herein are not final and may be updated before the official final publication.

INDUSTRY VOICES ON THE AMAT GUIDE

"Today, most of the AI conversation is still about automating what we already understand. The real breakthrough will come from AI that discovers what was too complex, too fragmented, or too invisible for humans to see. The AMAT Guide challenges investment leaders to look beyond automation and prepare for that future."

Mike Hunstad, President of Northern Trust Asset Management (NTAM)

"Most firms have adopted AI; few have changed how they operate. Pillar 1 of the AMAT Guide is clear about that difference, and gives leadership teams concrete moves to close it, from naming who owns delivery to setting a target maturity."

Evan Fire, COO at Pzena Investment Management

"The AMAT Guide helps institutions connect business ambition with the data, operating practices and governance needed to make that transition. For financial institutions, that foundation is essential to earning trust, scaling adoption and shaping the next phase of the business."

Akshay Chand, Partner and Chief Technology Officer at MKP Capital Management

"The AMAT Guide treats strategy and control as one discipline: decide where AI should create value, then be able to prove the controls behind it ran. That is the test boards, supervisors and clients now apply."

Declan Sheehy, Managing Director, One Blackwater Consultancy

"The AMAT Guide keeps the human in charge without pretending the machine is optional. Every action names a decision owner and a stop authority, which is exactly what a fund board needs to see."

Christophe Santer, Director, Mangis Bay

The full AMAT Guide
All eight pillars, every article and the evidence behind them.
amatguide.org/amat-guide



INTRODUCTION

by Junchen (William) Wu, PhD, CFA
Founder, AMAT

Why this guide exists

Most asset managers now use AI. Few have changed how they operate because of it. In a 2026 survey of 352 financial-services firms, about four in five reported adopting AI at some level, while about one in seven regarded their deployment as transformational to strategy and competitive advantage (self-reported). Pillar 1 treats that gap as the central problem, and the guide is written to help close it.

AMAT stands for Asset Management AI Transformation. It began as closed-door conversations among the people accountable for AI at asset managers and hedge funds, in New York in May 2026 and then in Chicago, London and San Francisco. The AMAT Circle that grew out of those sessions brings together C-suite and senior AI leaders from institutions that together manage more than US$38 trillion.

It is an industry guide because asset management is a fiduciary business, and trust in a fiduciary business is a shared balance sheet. When one firm's agent sends a wrong figure to a client, compliance officers everywhere grow more cautious and regulators write tighter rules. The standards that will govern AI in investment management are being drafted now, in several rooms at once, and they will be shaped by whoever arrives with evidence. Moving together does not mean moving at the pace of the slowest. It means not paying a hundred times for the same lesson.

What AI-native means

A firm exists to overcome the cognitive limits of individual people. No one can follow 3,000 issuers, hold 400 positions, check every order against every mandate and remember why the committee decided as it did in 2019. So the industry built analysts, portfolio managers, traders, risk, compliance and operations, and wired them together with meetings, memos and morning notes. The firm is a workaround for the bandwidth of one human mind.

AI that makes each workaround faster is useful, and it is not yet a transformation, because the person is still the bottleneck. Pillar 1 draws the line between bolting AI onto existing processes and embedding it in how the firm makes decisions. The second changes the org chart. Roles built around assembling context are rebuilt around judgment and exceptions, decision rights are written down because an agent needs to know where its authority ends, and the line between human judgment and machine capability becomes a design decision rather than an accident of who owned the spreadsheet.

The most natural mistake is to take every process the firm runs today and make it faster. Done perfectly, that produces a very efficient museum, a flawless reproduction of how asset management worked before a model could read every filing itself. The better question is what you would design if AI were assumed.

Where firms get stuck

Firms that make progress meet the same obstacles, and each has a home in the guide. Assistants that cannot see each other, and data access wired up one analyst at a time, are the subject of Pillars 3 and 4. Dependence on a single model or vendor is covered in Pillar 4, including how to keep the exit open (P04-ST11), and in Pillar 7. AI cost that scales with usage rather than seats is examined in P01-ST08 and Pillar 8, which also deals with pilots that never reach production. Across all of them, the asset that lasts is the firm's own context: what it knew, what it weighed, why it acted and what it learned. Models will be rented and swapped as capability shifts. How that context is captured, governed and kept under the firm's control links Pillars 3, 4 and 7.

The next phase raises the stakes. Agents will deal with other agents: a research agent asking a custodian's agent for point-in-time holdings, a client's agent requesting attribution and expecting a sourced answer in minutes. In an EY survey from early 2025 of 100 wealth and asset management firms, 78 percent were exploring agentic AI and 7 percent had deployed it. The guide's answers are concrete. Place every agentic workflow on an explicit autonomy spectrum, write down decision rights for every material decision, watch override rates (a rate near zero signals rubber-stamping, not trust), and name a human who can stop any live workflow. Autonomy should be earned while the stakes are still small enough to learn from.

Who is behind the guide, and how to use this edition

The content comes from published evidence and from the practitioners who lead and review each pillar. Its goal: one shared, evidence-based reference for the industry's move to AI-native, so that no firm pays full price for a lesson another firm has already learned. It sells nothing. Where the evidence is strong, it says so. Where the evidence is thin, as it still is on realized AI returns, it says that too, rather than borrowing confidence it does not have.

This Executive Edition is the short version for leaders. For each of the eight pillars it gives a one-page summary and the priority moves, then the recommended actions for every sub-topic, with who acts, the top risk to watch and what to measure. It carries no citations by design. Every figure is referenced in the digital edition, and each page has a QR code to the full article. Do not read it front to back. Find the pillar that owns your current problem, read its summary, and follow the code when you need the proof.

This is a first edition, which is a polite way of saying it is wrong in places we have not found yet. Register at amatguide.org/register, argue with us, and send back what you learn. We built this guide so the industry could stop learning in private.

Junchen (William) Wu, PhD, CFA
Founder, AMAT
November 2026

HOW TO USE THIS EDITION

This Executive Edition distils the AMAT Guide to what leaders need to act on: for each of the eight pillars, a one-page executive summary and the priority moves, then the recommended actions for every sub-topic.

The full guide, built from over 550 sources, is online. Each page here carries a QR code to the matching section, where you will find the complete article with its evidence, case examples, risks and controls, guidance by firm size, and figures.

About citations. To keep this edition short it carries no citations. Every figure and claim is fully referenced in the digital edition, where each article ends with its Citation Pack.

A note on sub-topic numbers. Sub-topic numbers are permanent identifiers shared with the digital guide, so every QR code and cross-reference stays valid as the guide develops. You may notice that a few numbers are missing within some pillars. During development, closely related topics were consolidated into a single, deeper sub-topic, and the original number was retired rather than reused. Nothing has been dropped: the content now sits in the sub-topic that absorbed it. New sub-topics will be added to the digital edition as the field develops.

CONTENTS

P01	Strategy, Vision, Leadership	6
P02	AI Use Cases	32
P03	Data, Knowledge, and Decision Intelligence	64
P04	Technology Architecture, Infrastructure, and Model Layer	90
P05	Governance, Risk, Regulations, and Controls	128
P06	People, Talent, Culture, and Change Management	170
P07	Assurance, Audit, Testing, and Third-Party Oversight	202
P08	Scaling, Value Realization, and Industry Transformation	238

The full AMAT Guide
All eight pillars, every article and the evidence behind them.
amatguide.org/amat-guide

AMAT

Pillar 1 Executive Summary

For: Board and Executive; Senior Leadership · Also relevant to: Strategy; Technology

THE ARGUMENT

The industry's AI challenge is less about adoption than transformation. Most financial-services firms now use AI in some form, but few have changed how they actually operate. Advantage is more likely to come from becoming AI-native, with AI embedded in the operating model, the strategic logic, and the competitive identity of the firm, rather than from adding tools to a business still built for a pre-AI world. Pillar 1 sets that destination and the case for moving toward it deliberately.

14%
of 352 surveyed financial-services firms see AI as transformational, though 81% adopt it (2026, self-reported)

74%
of AI economic value captured by about 20% of companies, in one 2026 cross-industry study

95%
of alternative-investment managers use gen AI; under 1 in 5 enterprises across industries have scaled it (2025)

WHY NOW

In a 2026 survey of 352 financial-services firms, around four in five report adopting AI, but only about one in seven treat it as transformational to strategy. That gap is where advantage is being won and lost. Value is already concentrating, and margin pressure is structural: PwC projects that about one in six asset and wealth managers will be bought or shuttered by 2027, twice the historical rate of turnover. The window to build a durable position is narrowing while many firms are still running pilots.

RISKS TO AVOID

Mistaking adoption for transformation (running pilots while the operating model stays pre-AI). **Cost-only framing without senior sponsorship** (under-investing in alpha, resilience, and client experience). **Delay** (value and data advantages compound, and consolidation pressure is rising).

Bottom line. Pillar 1 sets the destination: an AI-native firm where AI is embedded in the operating model, the strategy, and the competitive identity. Firms that treat this as transformation, sponsor it from the top with clear delivery ownership, and move deliberately are best placed to capture a disproportionate share of the value; the rest risk fee pressure and consolidation.

Chapter Leads, Pillar 1: Evan Fire; Declan Sheehy

Explore Pillar 1 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P01-summary

PRIORITY MOVES FOR LEADERSHIP

1. **Have the CEO sponsor the AI ambition and name an operating executive to own delivery**, and frame it across all four value drivers, not cost alone.
2. **Baseline where the firm sits on the AI-native maturity curve**, and set an explicit target state with a named, accountable owner.
3. **Redesign priority workflows front-to-back around AI** (zero-based), rather than automating existing processes as they stand.
4. **Map material decisions to a human-AI decision-rights spectrum**; keep final investment, fiduciary, and client decisions human-led.
5. **Treat responsible AI as a strategic principle embedded from the start**, with a prohibited tier in the AI policy and trust as a source of differentiation.
6. **Break out of pilot mode in the right order**: back the teams who must change how they work with a top-down mandate and business-owned delivery; classify before scaling and align before automating.

IN THIS PILLAR

P01-ST01	Definition of an AI-Native Asset Manager	8
P01-ST02	Strategic Ambition: Efficiency, Alpha, Resilience and Client Experience	11
P01-ST03	Front-to-Back Transformation Vision: Designing the AI-Native Operating Model	14
P01-ST04	The Human-Plus-AI Operating Model	17
P01-ST05	Competitive Advantage and Differentiation: Building a Moat AI Cannot Copy	20
P01-ST06	Responsible Innovation as a Strategic Principle	23
P01-ST07	From Experimentation to Enterprise Transformation	26
P01-ST08	How AI Changes the Economics of Asset Management	29

Explore Pillar 1 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P01-summary

P01-ST01

Definition of an AI-Native Asset Manager

For: Senior Leadership; Board and Executive

Most financial-services firms now use AI in some form, yet only a small minority have been reorganized around it. The test of an AI-native firm is whether AI has changed how its decisions are formed, tested and executed, across investment and equally in operations, distribution and the middle office. Functional adoption bolts AI onto existing processes to make them faster; systemic adoption embeds it in the decision architecture, and only the second produces a smarter organization.

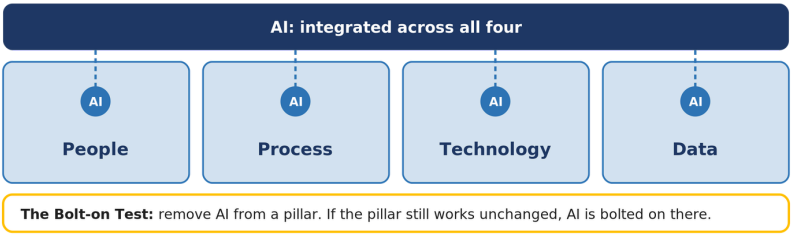
81%
of surveyed financial-services firms adopt AI at some level (n=352, 2026)

14%
regard their AI deployment as transformational to strategy and competitive advantage

WHERE FIRMS GO WRONG

Functional and systemic adoption are easy to confuse from the inside, because both involve real spending, real vendors and visible activity. Loose language makes it worse: "we use AI" can mean anything from task automation to models that shape live portfolios. Many firms that would describe themselves as well advanced are running a portfolio of functional pilots around an unchanged core.

The AI-native operating model



AMAT framework, drawing on Evan Fire, AMAT chapter lead contribution, 2026.

Takeaway. Treat "AI-native" as a position to diagnose and improve rather than a badge to claim: distinguish functional from systemic adoption, baseline honestly, name a target and an owner, and re-baseline as the frontier moves.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST01

RECOMMENDED ACTIONS

1. **Classify every material AI use, including vendor-enabled features, as functional or systemic** using one shared definition, and record which decisions each one actually changes.
2. **Baseline the firm's AI maturity** against a single named maturity framework within one quarter, scored per function, not as a single headline level.
3. **Start by reporting breadth and depth of AI use to the board annually**; build toward more frequent reporting once the measures are reliable, so shallow breadth is never mistaken for depth.
4. **Define the target AI-native end-state, name an executive sponsor for it and a business owner for each live AI workflow who can stop it**, supported by a cross-functional AI council spanning technology, investment, and operations, so accountability is not diffused.
5. **Re-baseline maturity annually**, treating "AI-native" as a moving target that rises as capability compounds.
6. **Reframe the leadership question** from "what can AI do for our existing process?" to "what would we design if AI were assumed?", reserving redesign for the workflows where it is worth the disruption.
7. **Prioritize by decision impact before scaling**: direct AI investment where it changes decisions and outcomes, not where deployment is broadest or the technology most impressive.

SEQUENCE AND DEPENDENCIES

1. **Agree one shared taxonomy** first; without it, baselines are not comparable across functions and the board metric becomes noise.
2. **Classify and baseline** AI use honestly before setting targets, and be willing to publish any gap against the firm's own narrative internally.
3. **Set targets** only after classification and baselining, because a destination set without a starting point produces slogans rather than plans.
4. **Decide redesign versus augment** workflow by workflow; redesign is costly and disruptive, augmentation elsewhere captures most value for a fraction of the cost.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns honest classification of its AI uses, including vendor features, and the redesign-versus-augment decision per workflow.
Second Line	Maintains the shared AI taxonomy and maturity scale, and checks that stated maturity matches how AI actually shapes decisions.
Board and Executive	Sets the target end-state, names a sponsor, gives each live workflow an owner, reviews breadth and depth at least annually.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST01

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Mistaking functional adoption for transformation**	Classify every AI use as functional or systemic against one definition; report breadth and depth to the board regularly.
HIGH **Execution gap: adoption without embedding**	Set a target end-state with a named sponsor, give each live AI workflow a named owner, and re-baseline maturity regularly.
MEDIUM **No shared AI taxonomy**	Adopt one firm-wide taxonomy and maturity scale before scaling investment, and baseline against it by function.
MEDIUM **Model and vendor concentration**	Document key model, supplier and data dependencies with named owners, and confirm each provider's export, fallback and disable options.

METRICS SCORECARD

AREA	LEADING	LAGGING
Breadth of use	Share of staff using AI in daily work (active in the last 30 days)	Share of roles whose standard workflow includes an AI step
Depth of integration	Core workflows running with AI in production, not pilot, by function	Before-and-after change in those workflows (cycle time, error rate, capacity released)
Decision impact	Share of production AI workflows with a documented decision or outcome they change	Change in that decision's outcome over time (e.g., time to decision, exceptions avoided)
Maturity position	Priority functions scored against the chosen framework, including people and data	Year-on-year movement in scored level

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Simpler model can be redesigned around AI; the trap is adopting a vendor's definition. Focus on one or two edge-defining workflows.

MID-SIZED

Widest range of outcomes. Trap is fragmentation: many pilots, no consistent ownership or baselines. Shared taxonomy and board reporting matter most.

LARGE AND GLOBAL

Can mistake AI spending for depth of integration, and legacy operating models make redesign hardest. Strong IT leadership, modularity and reuse help.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P01-ST01

AMAT

P01-ST02

Strategic Ambition: Efficiency, Alpha, Resilience and Client Experience

For: Senior Leadership; Board and Executive

An AI-native operating model is a means to an end, and the board owns the prior question of what the firm's AI ambition is for. Ambition should be measured against four value drivers: efficiency, alpha, resilience and client experience. AI leaders are more likely to frame it as a growth and reinvention agenda than as a cost-reduction program. Where firms go wrong is consistent: ambition is set too low, framed around efficiency, sponsored too far below the C-suite, and often never translated into concrete objectives.

74%
of 500 CEOs believe they could lose their job within two years without measurable AI outcomes

1 in 10
organizations name the CEO as primary leader of the AI agenda (separate survey)

WHY IT MATTERS NOW

The pressure to have an answer on AI now comes from clients, allocators, competitors and boards, and chief executives are judged on measurable outcomes rather than pilots sponsored. A firm that treats AI as a way to shave cost from today's model is preparing for the last environment rather than the next one.



Takeaway. Set ambition across all four value drivers, sponsor it from the CEO with a named delivery owner, and convert it into measurable objectives with risk and spending guardrails. None of these is a technology decision; all of them are the board's to make.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST02

RECOMMENDED ACTIONS

1. Start by aligning the AI strategy explicitly with the firm's business and technology strategy and tying each AI initiative to a stated commercial objective; build toward AI as an integrated layer supporting those objectives, not a parallel initiative.
2. Frame ambition across all four value drivers and set the weighting with a four-question test. Start by recording an indicative allocation of AI spend and capacity across the drivers; build toward a split of the AI budget. Make growth and business-model reinvention the organizing logic, not cost reduction alone.
3. Have the CEO sponsor the AI agenda and set its weighting with the board, and name an operating executive who owns delivery, with a standing board cadence.
4. Define the firm's AI ambition and competitive posture with key stakeholders before selecting or prioritizing use cases.
5. Set measurable AI objectives across productivity, client experience and risk, and fix explicit risk-appetite, governance and spending guardrails alongside them, with an expected return for each material AI investment.
6. Run a dual track. Start by capturing near-term value on the platforms the firm has today and fixing the weakest capability first; build toward the data and governance foundations that let AI scale.

SEQUENCE AND DEPENDENCIES

1. **Run two tracks together**: deliver near-term value while building the data and governance foundations that let it scale.
2. **Anchor to posture and objectives first**, not a use-case list; use-case-first programs drift toward fragmented effort and uneven maturity.
3. **Build the narrative** for why AI matters strategically; many firms lack one, which starves ambition of alignment and momentum.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Translates ambition into value driver by driver and into client-facing products; turns internal efficiency gains into client-facing capability.
Second Line	Drafts the risk appetite and disclosure guardrails, and owns the AI-washing and conduct exposure that overstated ambition creates.
Board and Executive	Sets the ambition and four-driver weighting; CEO sponsorship and board oversight of a named delivery owner; accountable for AI governance.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST02

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Over-reliance on market tailwinds**	Weight ambition toward resilience and durable commercial capability, not efficiency alone; use AI to strengthen distribution, risk identification and pricing.
HIGH **Ambition-execution gap**	Set measurable objectives across productivity, client experience and risk, with a named owner and board cadence.
MEDIUM **Cost-only framing ceiling**	Make growth and business-model reinvention the organizing logic; require every AI initiative to state which value driver it advances.
MEDIUM **Client-experience quality gap**	Design around complex, high-value interactions, not only simple queries; measure satisfaction on AI-assisted journeys, not just volume.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Value-driver balance	Allocation of AI spend and capacity across the four drivers, and share of initiatives mapped to each	Outcomes against each initiative's stated driver (e.g., client retention, mandate wins) alongside cost savings realized
Strategic ownership	CEO sponsor and named delivery owner in place; AI a standing board agenda item	Board review cadence maintained over the year
Client experience	AI-enabled client-experience initiatives live	Client satisfaction or NPS on AI-assisted journeys
Financial discipline	Share of material AI investments with an approved envelope and expected return	Realized versus expected return, and AI run-cost against plan

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Capacity is the constraint. Weight all four drivers, fund one, and lean on platform and technology partners rather than building broadly.

MID-SIZED

Some firms can carry several drivers; others fund the clearest one or two first. Convert ambition into measurable objectives before widening.

LARGE AND GLOBAL

Scale allows ambition across all four drivers. The risk is diffusion, so set and enforce the four-driver weighting from the center.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P01-ST02

P01-ST03

Front-to-Back Transformation Vision: Designing the AI-Native Operating Model

For: Operations; Technology; Senior Leadership

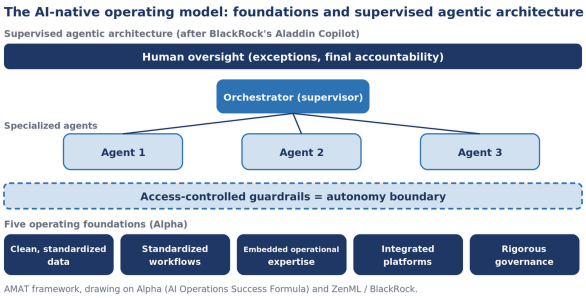
The strategic ambition depends on one architectural decision: redesign the operating model front to back around AI, or bolt AI onto the processes the firm already runs. The evidence, much of it cross-industry, is that the largest gains come from redesigned processes while AI on a legacy workflow tends to plateau. The reference design combines Alpha's five operating foundations with a supervised agentic architecture, and adds a written boundary between where AI acts and where a named human intervenes. It is AMAT's synthesis, not one firm's tested model. Two disciplines follow: let the operating model lead the technology build, and recalibrate the target as it moves.

59%

of firms classed as AI leaders say they develop and communicate a top-down vision for AI transformation (ThoughtLab survey)

WHERE FIRMS GO WRONG

The dominant failure is not slow adoption but the shape adoption takes: point solutions are bolted onto workflows designed for a pre-AI world, work at first, then flatten because the constraint was never the tool. Beneath it sits the habit of automating the existing process rather than reinventing it, and fragmentation makes it worse, since an orchestration layer over a different core system in each geography has nothing coherent to orchestrate.



AMAT framework, drawing on Alpha (AI Operations Success Formula) and ZenML / BlackRock.

Takeaway. Front-to-back transformation is a choice about architecture. Name an accountable owner, set a concrete vision, and run the processes up for funding through the automate-versus-reinvent gate before the next budget cycle commits money to the wrong version of the future.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P01-ST03

RECOMMENDED ACTIONS

1. **Set and communicate a top-down front-to-back vision** that states where AI sits in the operating model and what changes in how work is done.
2. **Classify core processes** as "automate as-is" or "reinvent around AI" before committing budget, starting with the domains up for funding this cycle.
3. **Let the operating model lead the platform**: test each release against the target model and standardize processes across geographies before scaling.
4. **Redesign one priority domain end to end** before deploying agents into it, then add domains as measured results justify.
5. **Build the orchestration layer one redesigned domain at a time**, and state in writing which decisions are AI-autonomous and which need human intervention.
6. **Recalibrate the target operating model** at each annual strategy cycle.
7. **Name a single accountable owner** for the front-to-back operating model at board or executive level.
8. **Assign run-state ownership for every live AI workflow**, with a register that includes vendor-enabled features and re-review before any material change.

SEQUENCE AND DEPENDENCIES

1. **Build data foundations in order**: quality, then standardization, then pilots, concentrating investment on priority domains rather than comprehensive cleanup.
2. **Define who owns each value stream** before platform work starts; retrofitting role clarity after the technology is live generates costly customizations.
3. **Apply one shared classification rule**: reinvent where a process is complex, cross-functional, exception-prone and performance-critical; automate where it is standard and low-variability.
4. **Start run-state ownership at go-live**: a register of live workflows, a named reviewer for each and a change gate, budgeted as staffing.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns AI-assisted decisions in redesigned processes, handles routed exceptions, approves changes at the gate, decides when to suspend or retire.
Second Line	Sets the risk standard for the human-versus-AI boundary and change gate, challenges boundaries, and tests that changes passed the gate.
Third Line	Provides independent assurance over redesigned processes and the decision boundary, testing that the orchestration layer's controls operate as designed.
Board and Executive	Owns the front-to-back vision, sees the operating model as the binding constraint, and names an accountable owner with a mandate.
Cross-Functional	Treats the integrating layer and cross-function feedback loops as shared infrastructure that no single function owns.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST03

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Bolt-on ceiling**	Mandate zero-based redesign of priority domains before agent deployment; gate funding on an explicit automate-versus-reinvent classification.
HIGH **Sequencing failure: platform before model**	Keep the operating model in the lead: standardize processes before the platform scales and treat data foundations as the first dependency.
HIGH **Undefined autonomy boundary**	Define the decision boundary in writing in the governance layer; run an exception-based model with explicit human review points.
HIGH **Uncontrolled change to live workflows**	Register every live workflow with an accountable reviewer; require re-review before any material change to a model, parameter, prompt or skill is used.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Process reinvention depth	Share of priority domains with a recorded pre-redesign baseline and a redesigned workflow live	Cost, cycle time and error or exception rate per redesigned workflow against its baseline
Straight-through processing	STP rate on redesigned processes	Cost and cycle time per transaction
Platform consolidation	Number of core systems in the estate	Count of cross-system interfaces
Recalibration cadence	Operating model recalibrated each annual strategy cycle	Time since last target-state reassessment

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Buy rather than build. Check current vendors can support one end-to-end workflow without duplicating data or controls; focus redesign on one domain.

MID-SIZED

Trap: scaling the platform before the target operating model is defined. Keep the model in the lead and design orchestration across vendors.

LARGE AND GLOBAL

Legacy fragmentation and interface complexity dominate. Standardize the global operating model before the platform scales, or pay later in customizations.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P01-ST03

AMAT

P01-ST04

The Human-Plus-AI Operating Model

For: Senior Leadership; Investment Teams; Operations

The human-plus-AI operating model is a design decision about where authority sits, not a technology rollout. So far the industry has not chosen autonomy: in a February 2026 survey of 131 asset managers, just five percent gave AI autonomous or semi-autonomous authority over investment decisions. Machines earn their place on breadth, speed and consistency; people earn theirs on judgment under uncertainty and on accountability that cannot be delegated. Firms that allocate decision rights well find AI compounds scarce human judgment, and those that allocate them badly find it quietly erodes value. The practical work is to map decision rights, redesign roles around activation and exception management, and instrument the model so oversight does not become a rubber stamp.

5%

of 131 surveyed asset managers entrusted AI with autonomous or semi-autonomous authority over investment decisions (February 2026)

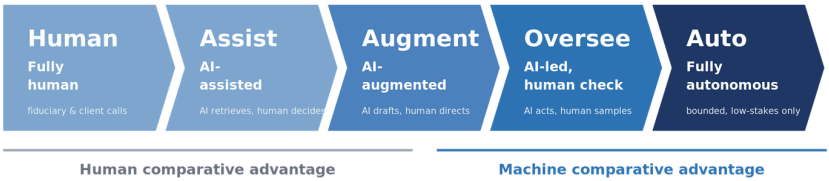
2.8 times

more likely to have increased decisions made without human intervention among companies capturing most AI value (cross-industry, 1,217 executives)

WHERE THE BOUNDARY GOES WRONG

The characteristic failure is a boundary drawn in the wrong place, which loses either judgment or breadth and consistency. Most firms bolt AI onto the existing org chart and leave roles untouched, so the efficiency never fully arrives. Where automation is pushed into judgment, the cost shows up as an accountability gap, and as routine cognitive work moves to machines, firms risk deskilling the very capabilities the human check depends on.

Decision-rights spectrum: allocating authority between people and machines



AMAT framework, drawing on AllianceBernstein (2026).

Takeaway. The boundary between people and machines is not a fixed settlement. Draw it deliberately, keep human judgment on the calls that carry consequence, instrument oversight so it stays honest, and redraw it as capability changes.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P01-ST04

RECOMMENDED ACTIONS

1. **Map decision rights** for the material decisions where AI is in use, vendor-enabled features included, recording who decides, who oversees and the escalation trigger, and build toward a map of every material decision, reviewed at least annually.
2. **Make human sign-off non-negotiable for high-stakes outputs.** Marketing claims, investor communications, legal conclusions and final investment decisions stay human-owned, with AI assistive only.
3. **Scope AI to bounded tasks** with clearly defined objectives, and avoid open-ended deployment that raises error risk and weakens oversight.
4. **Redesign roles around activation and exception management**, starting where AI is already in daily use, naming who handles exceptions and redirecting a measured share of the time saved to review and higher-judgment work.
5. **Instrument the model.** Track the autonomy-authorization rate, the override rate and high-stakes sign-off compliance, and investigate an override rate at or near zero as a possible sign of rubber-stamping.
6. **Put an anti-deskilling policy in place:** continual fact-checking of machine output, rotation through non-automated reasoning and deliberate retention of core skills.
7. **Design learning into the system.** Capture every human override or correction and feed it back into model and process design, with every change versioned, reviewable and reversible.

SEQUENCE AND DEPENDENCIES

1. **Start from the decision-rights map** and the target division of labor, not from a list of tools to deploy.
2. **Advance decision-rights mapping, role redesign and skills-building together**; deploying capability ahead of the operating model leaves tools nobody is clearly accountable for.
3. **Redirect freed capacity** into higher-judgment work, and build enough trust that people use the tools within scope.
4. **Run a human-machine feedback loop**: humans guide, correct and validate the machine, and corrections flow back into design under version control.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the decision-rights boundary inside the workflow, operates AI within scope, manages exceptions and embeds human judgment at defined points.
Second Line	Sets and validates oversight and escalation thresholds, enforces human sign-off on high-stakes outputs, blocks unadjudicated autonomy in regulated workflows.
Third Line	Audits that the boundary as practiced matches the design, and that override and exception logging is real, not cosmetic.
Board and Executive	Owns the design question and governance accountability, personal under regimes such as SM&CR, and approves the enterprise-wide autonomy posture.
Cross-Functional	Runs the human-machine feedback loop and the cross-functional role redesign, so the operating model changes together rather than in silos.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST04

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Mis-drawn boundary**	Maintain an explicit, reviewed decision-rights map and classify each task by comparative advantage; on a boundary error, halt or demote to full human review first.
HIGH **Accountability gap**	Require human sign-off for high-stakes outputs and keep AI scoped to bounded tasks.
HIGH **Autonomy in regulated work**	Restrict autonomy in regulated and high-stakes workflows; require adjudication of model disagreements.
HIGH **Client trust delegated**	Keep relationship and advice decisions human-led on the decision-rights map.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Decision-rights coverage	Share of material decisions with an assigned tier and owner	Audit findings of unmapped or drifted decisions
Autonomy authorization	Decisions granted AI-led or autonomous authority per cycle	Incidents traced to autonomous decisions
High-stakes sign-off	Share of high-stakes outputs with documented human sign-off	Conduct or regulatory findings
Skills retention	Participation in non-automated reasoning and rotation	Measured decline in core skills

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Limited oversight capacity and vendor reliance. Concentrate human judgment on irreversible, client-facing and fiduciary decisions and set the autonomy spectrum conservatively.

MID-SIZED

Building in-house capability alongside vendors. Formalize the decision-rights map before scaling, so local autonomy choices do not proliferate inconsistently.

LARGE AND GLOBAL

Many desks and jurisdictions; inconsistent boundaries are the dominant risk. Govern the spectrum centrally and standardize escalation thresholds across the enterprise.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST04

AMAT

P01-ST05

Competitive Advantage and Differentiation: Building a Moat AI Cannot Copy

For: Senior Leadership; Board and Executive

AI does not by itself create competitive advantage in asset management. Frontier models are becoming shared infrastructure that every firm can rent. Durable advantage comes from what a firm builds around the model: proprietary data and a closed learning loop, the house judgment it encodes into its systems, the governance discipline that lets it scale with confidence, and the strategic clarity to choose where it will compete. In one 2026 study, roughly 74% of AI economic value was captured by about 20% of companies, and in one working paper the return gap between higher and lower AI-adoption US mutual funds widened from 2008 to 2023, though it measures adoption, not these capabilities.

74%
of AI economic value captured by about 20% of companies (one 2026 study)

11%
of asset and wealth managers qualify as "future-fit", strong on both strategic clarity and execution

WHERE FIRMS GO WRONG

The central error is mistaking the purchase of AI for the building of advantage. The same generative tools are available to every participant, so AI is largely a leveler that raises the industry baseline and shortens the half-life of an information edge. Most firms also adopt AI functionally rather than systemically, which is easier to buy and to show a board but yields faster tools, not a smarter organization. Only about 11% of asset and wealth managers qualify as future-fit.

Table stakes versus durable differentiation

Dimension	Table stakes (commoditizing)	Durable differentiation
Foundation models	Same models rented by all	Judgment encoded around them
Data	Vendor and public datasets	Proprietary data, closed learning loop
AI applications	Drafting, summaries, basic analytics	Cross-function orchestration, memory
Source of edge	Faster tools (functional)	Smarter organization (systemic)
Client experience	Generic AI features	Trust and relationship quality

AMAT framework, drawing on CFA Institute (2025), DiligenceVault (2026) and AllianceBernstein (2026).

Takeaway. Decide where the firm will compete, fund the few capabilities that make that choice defensible, and stop paying for the rest. The studies suggest the leader-laggard gap is widening and the differentiating work is slow and path-dependent, so the cost of delay is rising.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST05

RECOMMENDED ACTIONS

1. **Decide your competitive archetype** this planning cycle and align capital, talent, product and partnerships to it.
2. **Require each AI investment to name the edge it amplifies** (informational, analytical or behavioral), and decline those that only improve a commoditizing capability.
3. **Assign an owner for the learning and memory layer** and capture investment reasoning, not just outcomes, building toward a funded, closed learning loop.
4. **Codify prompts and workflows as owned assets**, building toward governed proprietary datasets and durable AI intellectual property.
5. **Review each material agentic workflow as an investment case** (outcome, run cost, human review, option to stop), building toward allocating agent capacity to measurable outcomes.
6. **Start early with agentic AI in one or two workflows** that matter to the value proposition, then embed it more widely.
7. **Protect cognitive and epistemic diversity** in investment teams as a source of non-consensus views.

SEQUENCE AND DEPENDENCIES

1. **Make an early, sustained commitment**: the closed learning loop and institutional-memory layer are path-dependent and cannot be purchased or replicated from scratch.
2. **Establish the data and operational foundation first**, with disciplined model versioning, testing and drift monitoring, so models can be deployed and maintained reliably.
3. **Layer judgment-encoding work on top** once the foundation is in place; it compounds on that base.
4. **Expect a back-loaded payoff** and protect the initiative from cost-cutting pressure in the interim.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns moat-building: matching AI to the firm's edge, capturing reasoning into the learning loop, encoding house judgment into owned tools.
Second Line	Turns governance into competitive advantage rather than drag; explainable, well-governed AI supports client acquisition and valuations.
Third Line	Tests durability claims: ownership of workflow, prompt and data assets, evidence of human challenge and rollback, and AI-linked performance claims.
Board and Executive	Owns the archetype decision and the capital allocation behind it, and guards against assuming scale alone will win.
Cross-Functional	Protects epistemic diversity across investment, data and technology so uniform AI adoption does not erode capacity for non-consensus insight.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST05

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Leader-laggard divergence**	Commit to systemic, governed integration with board-level ownership and a dated roadmap; treat governance as an accelerant, not a brake.
HIGH **Capability concentration lockout**	Concentrate limited resources on one or two differentiating capabilities; use shared platforms and partnerships to reach parity on table-stakes capabilities.
HIGH **Alpha decay from homogenization**	Differentiate data and modeling deliberately; monitor signal decay and crowding; do not assume a bought model confers a lasting edge.
MEDIUM **Diffuse mid-tier investment**	Concentrate investment on one or two differentiating capabilities on a shared platform; buy table-stakes capability.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
AI integration maturity	Share of priority workflows systemically integrated (not just tool-enabled), governed and owned	Revenue growth relative to peers
Data and learning moat	Proprietary-data coverage of core process; reasoning-capture rate on closed positions	Persistence of AI-linked performance edge; share of sampled decisions amended after documented challenge
Strategic clarity	Share of AI spend mapped to the named archetype and edge; initiatives stopped for not fitting	Movement toward future-fit status
Epistemic resilience	Diversity of models and prompts, and rate of independent challenge in investment decisions	Incidence of correlated errors or missed non-consensus calls

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Advantaged

An agentic model could loosen the AUM-headcount link, on one 2026 projection. Constraint: compliance and data, with high-quality data getting costlier.

MID-SIZED

Exposed to technology- and data-driven acquisition if effort is spread thin, but the encoded-judgment moat does not need scale.

LARGE AND GLOBAL

Advantaged if converted

Scale funds proprietary data flywheels and institutional-memory infrastructure a follower cannot buy. Key risk: complacency about AUM and distribution.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST05

AMAT

P01-ST06

Responsible Innovation as a Strategic Principle

For: Senior Leadership | Risk and Compliance | Board and Executive

Treating responsible AI as a compliance matter for a risk team frames it as a constraint, and in an AI-native firm that is a strategic error. Trust is scarce: in PwC's surveys only about a third of CEOs report high trust in embedding AI into key processes. Cross-sector evidence, much of it modeled or self-reported, links responsible-AI practice to higher valuations, revenue and resilience. Responsible innovation is therefore an operating principle owned at board level, shaping which AI initiatives a firm funds and how they are built. The prescription is to embed it in the strategy, give it a single named owner, instrument the two best-evidenced practices and carry it into the investment process.

30%
fewer AI risks reported by firms with defined responsible-AI principles (2025 survey of 975 executives, self-reported)

4%
higher valuation, at most, modeled for robust responsible-AI firms versus compliance-only firms (PwC 2025, cross-sector)

WHERE FIRMS GO WRONG

The internal problem is location and ownership. Filed under compliance and handed to a second-line risk function, responsible AI is reframed as a constraint to satisfy rather than a capability to build, and governance arrives after the model exists. Capacity also lags deployment: in one survey of Dutch asset managers, half reported that governance was not keeping pace with their own AI adoption.

Reactive compliance vs proactive responsible innovation

Dimension	Reactive compliance	Responsible innovation
Ownership	Risk/compliance team	Board + named exec owner
Trigger	Regulation or an incident	Strategy and the value case
Scope	Point controls after build	Embedded across the lifecycle
Strategic value	A cost to minimize	Trust as an advantage
Outcome	Late, defensive, under-funded	Scales adoption; compounding trust

AMAT framework, drawing on McKinsey (2026) and EY (2025).

Takeaway. Treat responsible innovation as strategy: own it at board level, embed it in the portfolio, instrument it and extend it into the investment process. Those choices decide which AI uses the firm funds and how much client trust its AI earns.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST06

RECOMMENDED ACTIONS

1. **Put responsible AI in the AI strategy** and name one accountable executive owner, not the head of audit, this quarter.
2. **Set a prohibited risk tier** in policy, the firm's "won't build" list, with board approval and an annual review.
3. **Give an oversight committee a defined AI mandate**, building toward the full communicate-execute-govern operating model.
4. **Start with the two best-evidenced levers**: an oversight committee and model monitoring proportionate to risk tier, with every AI use inventoried and owned before it goes live.
5. **Require ethics-by-design for in-house models** and test bought models as configured in the workflow.
6. **Add responsible-AI expectations to one existing process**, such as a sustainable-investing questionnaire, building toward the whole investment and stewardship process.
7. **Explain the firm's AI governance consistently** to clients, employees and regulators, including in DDQ and RFP responses.

SEQUENCE AND DEPENDENCIES

1. **Name the owner, set the prohibited tier in a short policy** and start a basic inventory of live AI uses.
2. **Check data reliability** for each use case as it comes in.
3. **Stand up monitoring and the oversight committee** before broadening to the full program.
4. **Write the detailed policy** once ownership and the inventory are in place.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns ethics-by-design and risk-based scoping, maintains model documentation, and runs monitoring and incident escalation, including suspending a workflow.
Second Line	Owns responsible-AI policy and the monitoring standard by risk tier, runs the oversight committee, challenges first-line classification and evidence.
Third Line	Audits responsible-AI controls, provides independent challenge on fairness and monitoring, and tests that the "won't build" boundary is enforced.
Board and Executive	Owns responsible AI as strategy: sets risk appetite, approves the risk policy with its prohibited tier, names the accountable owner.
Cross-Functional	Moves trust, governance and workforce transformation together; sustainable value comes from combining innovation with all three, not technology alone.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST06

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Responsible AI as afterthought**	Make responsible AI a portfolio-intake filter through a prohibited risk tier and risk-based scoping.
HIGH **Compliance-only posture**	Invest beyond the minimum; run monitoring proportionate to risk tier and an incident-ready oversight committee.
MEDIUM **Weak trust stalls adoption**	Build and evidence trust through a consistent account of AI governance in DDQ and RFP responses and other client disclosure.
MEDIUM **Box-ticking**	Reframe as a value driver with board ownership and KPIs.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Strategic embedding	Responsible AI referenced in every funded AI business case; named executive owner in post	Share of live AI use cases meeting the outcome set in their business case
Portfolio scoping	Prohibited tier set in policy, board-approved and applied at intake	Realized AI incidents per year
Monitoring & oversight	Monitoring coverage by risk tier; oversight-committee cadence	Revenue and cost impact of monitored models
Trust	Consistent AI-governance answers maintained for DDQs and RFPs; client and employee trust surveyed	Client retention and trust-index trend

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Little room for a dedicated function. Name a part-time executive owner, keep a short inventory including vendor AI features, and spot-check outputs.

MID-SIZED

Governance usually trails adoption. Build the oversight committee and risk-tiered monitoring first, and formalize the prohibited tier in policy before scaling further.

LARGE AND GLOBAL

Formal program is unavoidable given multi-jurisdiction complexity. Extend it to portfolio oversight; the risk is bureaucracy, so keep it board-sponsored.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P01-ST06

AMAT

P01-ST07

From Experimentation to Enterprise Transformation

For: Senior Leadership; Technology

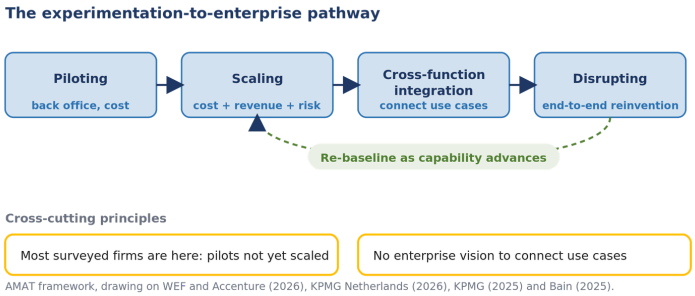
Adoption is no longer the question in asset management. In AIMA's 2025 survey of 150 alternative-investment fund managers, 95 percent reported using generative AI (self-reported), yet across all industries fewer than one in five enterprises had scaled it in any meaningful way by 2025, and these measures come from different populations. That gap between using AI and scaling it is where competitive advantage is now decided. Firms that cross it treat scaling as a business transformation, with executive ownership, a small number of concentrated bets, redesigned processes and a repeatable operating model. Those that stall keep running disconnected experiments on top of an unchanged operating model.

95%
of 150 surveyed alternative-investment managers use generative AI (self-reported, 2025)

58%
expect wider front-office AI integration in the next year, up from 20% two years earlier (same survey)

WHERE FIRMS STALL

The internal failure is rarely a shortage of pilots; it is the absence of the conditions that let a pilot become an enterprise capability. Three patterns explain most of the stall: a tactical, use-case-only focus with no enterprise vision, inertia in the operating model so each tool is bolted onto processes never designed for it, and treating AI as an IT deployment rather than a change in how the firm works. Without top-down direction and business-owned delivery, the teams who must change lack the permission to do so.



AMAT framework, drawing on WEF and Accenture (2026), KPMG Netherlands (2026), KPMG (2025) and Bain (2025).

Takeaway. Boards should act now but in the right order: classify before scaling, align before automating, and size the effort to the firm rather than copying a model built for a different balance sheet.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST07

RECOMMENDED ACTIONS

1. Name an accountable executive owner for AI transformation and document an enterprise AI vision that connects use cases end to end.
2. Concentrate on a few priority domains (Bain's leaders pick four to five; a boutique may start with one or two) and stop funding scattered pilots that do not fit them.
3. Start by redesigning the process from the ground up in one priority domain before scaling it; build toward redesign of every priority domain, rather than automating existing workflows as-is.
4. Start by installing a small central transformation team alongside business-owned delivery; build toward a codified, repeatable model for moving a capability from pilot to production.
5. Run an Enable-phase baseline: a shortlist of high-value use cases, an AI-literacy uplift, and proportionate guardrails.
6. Diagnose scaling readiness across governance, strategic focus, workforce and agentic AI preparedness before committing to scale.
7. Re-baseline AI readiness at least annually as capability advances; classify before you scale and align before you automate.

SEQUENCE AND DEPENDENCIES

1. **Diagnose readiness** across data quality, infrastructure, process readiness and people.
2. **Fix the binding constraint**, then redesign the priority process; scaling onto unready foundations produces poor outcomes and erodes confidence in AI investment.
3. **Scale** only after that, with a top-down mandate and business-owned delivery to keep the teams who must change how they work moving.
4. **Build in recalibration cycles** rather than a fixed multi-year roadmap, because each generation of AI capability demands a different governance and operating posture.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the transition in practice: decides which domains to scale, owns the pilot-to-production handover, and redesigns its own workflows.
Second Line	Builds proportionate governance for scaling, challenges classification and scaling decisions, and oversees the shadow adoption a use-case-only culture produces.
Third Line	Samples production workflows, not isolated pilots, testing handover acceptance, change approval, review records, incident handling and the stop route.
Board and Executive	Owns the transformation: the enterprise vision, the commitment to fewer, bigger bets and the timeline, as a standing agenda item.
Cross-Functional	The transformation team holds the repeatable operating model and the execution muscle that carries capability across domains.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST07

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Deployment, not transformation**	Executive-owned transformation with ground-up process redesign in priority domains.
HIGH **Compounding delay**	Act now on fewer, bigger bets and re-baseline as capability advances.
HIGH **Scaling on unready foundations**	Gate scaling on the four-dimension readiness diagnostic.
HIGH **No enterprise vision; use-case sprawl**	Document an enterprise AI vision and domain strategy and connect use cases end to end.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Scaling progression	Use cases in coordinated scaling or integration stage	Priority workflows in controlled production showing measured improvement against baseline
Pilot-to-production velocity	Median time from pilot approval to production	Pilots reaching production per year, and share still meeting quality and cost targets after handover
Ownership and run-state reliability	Named executive owner and funded team in place; share of live workflows with owner, monitoring and a tested stop route	Incidents, time to containment and repeat exceptions per live workflow, trended
Capability recalibration	AI readiness re-baselined against current benchmarks	Drift between assumed and actual capability

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Advantaged

Binding constraint: investment intensity against a small revenue base, thin technical capacity and amplified compliance cost (evidence from banks, by analogy).

MID-SIZED

Capacity squeeze

Copying the large-firm model risks cost-revenue mismatch, complexity beyond capacity and amplified risk. Leaders fund four to five domains properly.

LARGE AND GLOBAL

Advantaged if converted

Risks are teams without a top-down mandate to change how they work, and use-case sprawl without a connecting enterprise vision.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST07

AMAT

P01-ST08

How AI Changes the Economics of Asset Management

For: Senior Leadership; Finance; Board and Executive

Margin pressure in asset management is structural, not cyclical, and the technology sold as the cure can become part of the disease. AI changes the industry's unit economics by putting a large share of the cost base in play, by loosening the link between assets under management and headcount, and by accelerating fee compression. On a 2025 McKinsey estimate, AI, generative AI and agentic AI together could reach a quarter to two-fifths of an average manager's cost base, a modeled potential rather than a measured saving. Efficiency gains can leak to clients unless a firm reprices and repositions, and agentic AI carries a consumption-based cost that can rise rather than fall. The effect differs sharply by firm size.

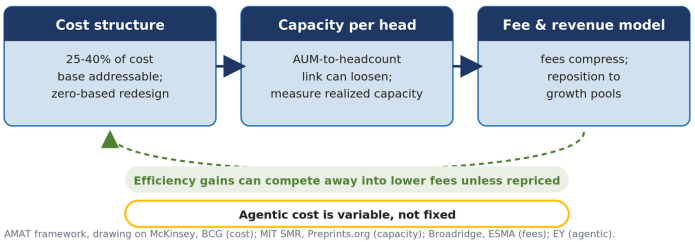
57%
of institutional investors say they are likely to replace a manager purely on cost

60 to 80%
of technology budget goes on running the business rather than changing it (McKinsey 2025 analysis)

WHERE FIRMS GO WRONG

The internal risk is treating AI as an efficiency add-on funded on top of the existing cost base rather than a reason to rebuild it. Technology is the fastest-growing cost center, at 8.9 percent a year against 4.5 percent for total costs, yet has the weakest-evidenced productivity return. Asset managers spend on average 60 to 80 percent of their technology budget running the business, so each AI initiative is layered onto an unchanged model and adds cost without restructuring it.

The AI economics model: three connected levers



AMAT framework, drawing on McKinsey, BCG (cost); MIT SMR, Preprints.org (capacity); Broadridge, ESMA (fees); EY (agentic).

Takeaway. Treat this as an economic decision, not a technology one: rebuild the cost base rather than adding to it, decide deliberately how to keep the gains, and size the response to the firm.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST08

RECOMMENDED ACTIONS

1. Run a zero-based rebuild of the highest-cost end-to-end workflow, using outsourcing, generative AI automation, and the elimination of dual-run costs: start this year with a cost baseline and a pilot run alongside the old process, and end dual-running only after a signed-off cutover.
2. Reframe technology from overhead to productivity multiplier. Start by holding the next round of AI investment decisions to a stated productivity return; build toward prioritizing, governing, and measuring all AI spend that way.
3. Start by picking the firm's cost archetype (alpha shop, solutions or distribution, beta factory) and testing the next budget round against it; build toward cost allocation aligned to it rather than to historical spend.
4. Budget agentic AI as a variable consumption cost with a named owner; do not treat an agent as a fixed-price license.
5. Set an explicit plan to keep, not surrender, efficiency gains: reprice deliberately and reposition capacity toward growth and higher-margin pools before fee pressure erodes the saving.
6. For sub-scale firms, do not replicate the large-firm platform; use vendor and shared platforms and concentrate AI on the workflows that differentiate.

SEQUENCE AND DEPENDENCIES

1. **Plan for a delayed, non-linear return**, and do not cut the program before the productivity J-curve turns.
2. **Rebuild the highest-cost workflow first**; the cost-structure lever is the largest and the enabler of the others.
3. **Use the freed capacity and savings to fund repositioning** of fees and revenue.
4. **Scale agentic use last**, where the consumption economics are understood; doing it without cost discipline tends to add cost faster than it removes it.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Rebuilds its cost base, redesigns highest-cost workflows, and manages agentic consumption as a variable operating cost, not a fixed bill.
Second Line	Reviews conduct implications of AI-driven pricing and advice before launch, and tracks the firm's exposure to market-level risks.
Third Line	Tests whether AI savings are real and net of consumption and oversight costs, not gross; does not certify the return.
Board and Executive	Owns the structural decision: zero-based cost rebuild, cost archetype, and, where it cannot fund the program, partnership or consolidation.
Cross-Functional	Finance and operations set the pre-AI cost and quality baseline per workflow and record realized savings and rework against it.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST08

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Margin illusion**	Reprice and reposition deliberately; track efficiency gains net of fee compression, not gross.
HIGH **Consumption-cost blowout**	Budget agents as variable operating cost with a named owner who can pause at a spending threshold; route routine steps to cheaper models.
HIGH **Capability concentration**	Board-level monitoring of the firm's relative AI position; deliberate build-buy-partner choices to avoid joining the lagging majority.
HIGH **Subscale cost trap**	Use shared and vendor platforms and focus on differentiating workflows; if the program is unaffordable, evaluate partnership or consolidation early.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Cost base restructuring	Cost per completed workflow against its pre-AI baseline, at recorded quality	Realized cost reduction net of implementation, consumption, review and transition costs, confirmed by Finance
Capacity per head	Automation coverage of high-volume workflows	AUM and revenue supported per FTE
Retained value	Net margin bridge: gross saving less fee pressure, agentic run cost and control cost	Operating margin trend
Agentic unit cost	Consumption cost per agent workflow versus prior baseline	Total agentic run-rate against budget

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE
Advantaged

A declining cost curve lets small teams run institutional-grade processes; a small revenue base limits capacity to absorb a variable AI bill.

MID-SIZED
Exposed

Carries enterprise complexity without boutique agility or global scale, leaving it exposed to consolidation if investment is spread thinly.

LARGE AND GLOBAL
Advantaged if converted

A fixed AI investment spreads across a large asset base; the advantage is conditional, as most technology budget runs the business.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P01-ST08

AMAT

Pillar 2 Executive Summary

For: Senior Leadership; Function Heads · Also relevant to: Investment, Risk, Compliance, Operations, Distribution

THE ARGUMENT

AI in asset management is not one capability but a portfolio of use cases across the front, middle, back and corporate functions. Pillar 2 maps where AI is deployed today, what each use case solves, and what must not be automated. The recurring lesson is that the value is real but conditional: in this guide's view, it accrues where each use case is governed as an investment input, an accountable human stays on the decision, and the firm tests whether its edge is real and plans for it to fade as the technique spreads.

~40%
of asset manager AI use cases in IOSCO's 2025 member survey are investment research, the most observed

56%
of firms in a 2026 survey name portfolio management and trading their top area for planned AI innovation

70%
shorter factor-signal half-life in the decade to 2024 than to 2014, in one 2026 arXiv preprint

WHY NOW

AI is now established in parts of the front office: among asset managers, investment research was the most-observed use case in IOSCO's 2025 member survey, and portfolio management and trading was the top area for planned AI innovation in a 2026 industry survey. Supervisors such as FINRA already run machine-scale surveillance, as the IMF reported in 2025, so weak controls are exposed quickly. And any edge may erode as adoption spreads: one 2026 preprint finds factor-signal half-lives about 70 percent shorter in the decade to 2024. The question has shifted from whether to deploy to how well, how fast, and how to protect an edge that may be perishable.

RISKS TO AVOID

Automating the decision, not just the analysis (ceding allocation, risk or fiduciary calls to models that should stay human-owned). **Ungoverned shadow use** (especially in corporate functions, where the footprint is broadest and least controlled). **Chasing a perishable edge** (over-investing in commoditizing techniques where the alpha is likely to fade with adoption).

Bottom line. Pillar 2 is the catalog of where AI earns its place in an asset manager, function by function. In this guide's view, what matters is not how much a firm automates but whether it picks the highest-value use cases, governs each as an investment input with an accountable human on the decision, and keeps moving as today's edge commoditizes. Deployed with that discipline, these use cases are where the strategy of Pillar 1 can become measurable value. The evidence is uneven: reported outcomes are largely the firms' own, no source gives a measured outcome for execution, corporate-function, distribution or social and governance AI at an asset manager, and beyond one hedge-fund study none shows AI alpha decaying across asset managers.

Explore Pillar 2 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P02-summary

PRIORITY MOVES FOR LEADERSHIP

1. **Prioritize use cases by value and readiness, not hype**, starting where use is already established (research, execution, client drafting) and governing each as an investment input.
2. **Keep accountable humans on every decision and external output**; automate the analysis and the drafting, not the allocation, risk or fiduciary call.
3. **Fix the operating model before buying capability**, especially in operations, compliance and corporate functions, and trace rules and data to source.
4. **Instrument for the edge and for its decay**: attribute performance to see whether AI inputs add return, monitor signal decay, and move spend up the frontier as techniques commoditize.
5. **Inventory the firm's full AI footprint**, including broad corporate-function usage, and put a named owner and review on each.
6. **Fix the distribution model before layering AI on it**: segmentation, coverage and data first, because ownership of the client relationship is becoming the defining advantage.

IN THIS PILLAR

P02-ST01	Investment Research and Idea Generation: Putting AI to Work in the Front Office	34
P02-ST02	Portfolio Construction and Decision Support	37
P02-ST03	Risk Management and Scenario Analysis	40
P02-ST04	Trading, Execution and Liquidity Analysis	43
P02-ST05	Operations, Reconciliations and Exception Management	46
P02-ST06	Client Reporting, RFPs and Investor Communications	49
P02-ST07	Compliance Surveillance and Regulatory Monitoring	52
P02-ST08	Finance, HR, Legal and Corporate Productivity	55
P02-ST09	Sales, Distribution and Relationship Management	58
P02-ST11	ESG, Stewardship and Sustainability Workflows	61

Explore Pillar 2 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P02-summary

AMAT

P02-ST01

Investment Research and Idea Generation: Putting AI to Work in the Front Office

For: Investment Teams

Research is where AI has moved furthest from pilot to production. IOSCO's 2025 survey of member regulators records investment research as the most frequently observed AI use case among asset managers, at around 40% of reported use cases, and one firm reported in 2026 that some research tasks that once took days now take under an hour. The evidence for a lasting edge is thinner. In a 2026 study of US AI hedge funds, early adopters earned roughly 6% a year above non-AI peers, but the advantage fell to statistically indistinguishable from zero after 2017, consistent with competitive exploitation. The question for a research leader is therefore how to build a capability whose advantage does not disappear when competitors copy the technique, which in this guide's view means proprietary data, differentiated models and disciplined human oversight.

40%

of use cases reported to IOSCO are investment research, the most frequently observed (2025 survey of member regulators)

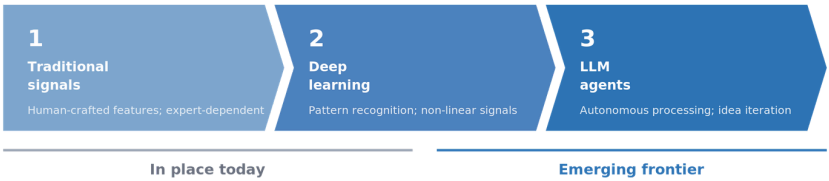
6%

a year of early-adopter outperformance by US AI hedge funds, indistinguishable from zero after 2017 (one 2026 study)

WHERE FIRMS GO WRONG

The gap between how much AI is used in research and how little of it is formal, governed or attributable is the most revealing in the industry. Firms use AI to do the same research faster rather than research they could not do before, few can say which AI components add value because they built no attribution, and cheaper short-term processing pulls analyst attention away from long-horizon work. The challenge is governing the tools as investment inputs and aiming them at durable advantage rather than speed alone.

The evolution of AI in investment research



Source: Cao et al. (2025).

Takeaway. The tools are mainstream, so the advantage has to be earned. Treat AI research as a governed investment input: attribute its signals, ground and verify its outputs, protect long-horizon judgment, and invest in proprietary data and models that competitors cannot copy.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P02-ST01

RECOMMENDED ACTIONS

1. Start by deploying grounded summarization of filings and transcripts for one coverage team; build toward AI research augmentation across the analyst team, with unstructured-data summarization and coverage expansion as the highest-return, lowest-risk entry points (RA 1).
2. Build AI-signal performance attribution from day one, so you can distinguish decaying signals from durable ones and retire the former (RA 2).
3. Start by naming one proprietary dataset or model your research could build that competitors cannot buy; build toward proprietary data and model differentiation rather than replicating published methods, which, in this guide's view, stop generating alpha once widely adopted (RA 3).
4. Point AI at alpha, not only at automation: prioritize research quality, novel signals, and coverage over pure efficiency (RA 4).
5. Require a human verification gate and source grounding on every AI research output that informs an investment decision (RA 5).
6. Size AI-driven strategies within liquidity limits and account for turnover and transaction costs in the business case (RA 6).

SEQUENCE AND DEPENDENCIES

1. Single-prompt extraction struggles with the cross-references, length, and verification needs of financial documents, which is why multi-agent designs that decompose work into specialist subtasks with verification loops now dominate serious.
2. Retrieval-augmented generation is the enabling pattern that grounds outputs in a firm's own documents and current market data, and it is the practical route to limiting hallucination.
3. The evidence points to a clear default: a hierarchical supervisor-worker architecture sits on the best part of the cost-accuracy frontier.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the use and the verification.
Second Line	Validate AI-derived signals and require interpretability where signals are opaque, since LLM-based text measures can limit auditability.
Third Line	Test whether attribution, verification gates, and documentation actually operate as described, and whether the firm can evidence how AI tools inform investment decisions.
Board and Executive	Set the strategic direction that AI research targets durable advantage, own the accountability that does not transfer to a tool.
Cross-Functional	Data and technology functions supply the retrieval infrastructure, data quality, and model architecture on which research capability depends.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST01

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **No demonstrated advantage**	Build signal attribution so the firm can see whether AI inputs add return; retire signals that stop paying and replace them.
MEDIUM **Short-termism drift**	Set research mandates and incentives that protect long-horizon coverage; monitor the horizon mix of AI-assisted output.
MEDIUM **Signal opacity**	Require an interpretability layer and documentation before a model-derived signal informs decisions.
MEDIUM **Hallucination and false precision**	Ground outputs in cited source documents and enforce a human verification gate before investment use.
MEDIUM **Memory and framing bias**	Test signal stability across configurations; benchmark against a neutral reference and control model inputs.

METRICS SCORECARD

AREA	LEADING	LAGGING
Research coverage	Names or asset classes under AI-assisted coverage	Analyst-adjusted coverage per head
Research speed	Time from question to grounded brief	Turnaround on complex research tasks
Signal quality	Model accuracy vs benchmark on a defined task	Realized risk-adjusted contribution
Extraction accuracy	Field-level extraction accuracy	Downstream error incidents
Adoption	Share of analysts using AI tools routinely	Analyst forecast timeliness and accuracy
Human oversight	Rate at which reviewers edit, override or reject AI output, trended	Post-decision error and rework rate

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST01

AMAT

P02-ST02

Portfolio Construction and Decision Support

For: Investment Teams

Portfolio construction has moved to the center of the industry's AI agenda: portfolio management and trading was the top area for planned AI innovation in a 2026 industry survey, cited by 56% of firms. The more consequential shift is what happens when many firms deploy similar tools, with one 2026 arXiv preprint finding factor signal half-lives compressing from around 60 months in the decade to 2014 to roughly 18 months in the decade to 2024. The evidence supports a division of labor in which AI finds patterns, screens ideas and stress-tests assumptions, while a transparent optimizer and accountable humans own the allocation decision.

56%
of firms cite portfolio management and trading as the top area for planned AI innovation (2026 industry survey)

18 months
factor signal half-life in the decade to 2024, down from around 60 months in the decade to 2014 (one 2026 arXiv preprint)

WHERE FIRMS GO WRONG

Many firms treat AI as the decision-maker rather than decision support, buying capability without redesigning the decision. Others run data-driven machine learning without investment expertise, or point generative models at return forecasting, and take on model risk because such models overfit and fail outside their training window.

The division of labor in AI-assisted construction

Dimension	AI leads	Optimizer + human decide
Pattern-finding & screening	Learns relationships; scans data at scale	Sets the opportunity set to search
Idea & security selection	Ranks and shortlists candidates	Approves the shortlist
Allocation decision	Proposes candidate weights	Owns weights via transparent optimizer and sign-off
Valuation	Drafts scenario and sensitivity bands	Justifies key drivers; signs off
Accountability	Logged, interpretable inputs	Named human owns the decision

AMAT framework, drawing on State Street (2026), Jadhav and Mirza (2025) and CFA Institute (2025).

Takeaway. Industrialize the analysis behind the allocation decision and keep the decision itself with a transparent optimizer and accountable humans. Build the unified data layer and validation discipline first, and make an explicit, sized call on the AI theme rather than drifting into it.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST02

RECOMMENDED ACTIONS

1. Split the workflow: use AI and language models for pattern-finding, screening and idea selection, but keep a transparent optimizer and human sign-off for the allocation decision.
2. Anchor every AI or machine-learning allocation model on investment logic, validate it against overfitting before production, and do not point generative AI directly at return forecasting.
3. Require scenario and sensitivity bands, not point estimates, for AI-assisted valuations, with named-analyst justification of the highest-leverage drivers before sign-off.
4. Build downside risk into any optimization or reinforcement-learning objective (for example a conditional value-at-risk term) rather than maximizing return alone.
5. Start by bringing positions and risk data for one multi-asset mandate into a single view; build toward a unified, real-time multi-asset data view so construction and pre-trade risk run on one source of truth.
6. Start by mapping the AI tools used in one strategy's investment process; build toward a map of AI across the five investment-process stages for every strategy, with decision rights proportional to model interpretability.
7. Make an explicit, sized call on exposure to the AI theme itself, with a pre-agreed checklist and timeline to course-correct.
8. Start by replacing static risk buckets with live goal and constraint data for one client segment; build toward personalized and automated portfolios driven by live client goals, constraints, tax and liquidity, with per-portfolio risk monitoring.

SEQUENCE AND DEPENDENCIES

1. Implementation succeeds or fails on two dependencies that precede any model choice: a unified data layer and a validation discipline.
2. On top of that layer, model governance should follow four foundational principles for investment modeling, namely foundational clarity, transparency, robustness and efficiency of process.
3. The single most cited front-office constraint is the inability to see multiple asset classes in one view, so a real-time.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the allocation decision and the exceptions.
Second Line	Validates AI allocation models for overfitting, robustness and transparency before production, applying the four foundational modeling principles; deeper model-risk governance is developed in Pillar 5.
Third Line	Tests whether AI allocation models were validated as claimed and whether human overrides are genuine rather than nominal, using the process map to trace accountability.
Board and Executive	Sets risk appetite for autonomous versus decision-support use and owns the explicit, sized call on AI-thematic concentration.
Cross-Functional	Delivers the unified, real-time multi-asset data layer that construction and pre-trade risk depend on.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST02

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Data-mining without investment logic**	Require investment-logic grounding and out-of-sample and cross-regime validation before production.
HIGH **Direct generative-AI forecasting**	Confine generative AI to analysis, summarization and idea support; keep forecasting in validated models.
HIGH **Signal crowding and alpha decay**	Track signal decay; diversify data and model architectures; re-budget risk toward layers still generating dispersion.
HIGH **Return-only reinforcement learning**	Embed a downside-risk term such as conditional value-at-risk in the objective.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Model validation	Share of AI allocation models with documented overfitting and cross-regime tests before production	Live-versus-backtest performance gap
Risk-adjusted performance	Downside-risk constraint embedded in the objective	Realized Sharpe and maximum drawdown vs a return-only baseline
Signal durability	Tracked signal half-life by strategy	Alpha decay rate
Human oversight quality	Rate at which reviewers edit, override or reject AI output, trended	Incidents traced to unchallenged AI output

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE
Advantaged

Agentic and vendor tools run institutional-grade, real-time personalized construction with small teams.

MID-SIZED
Exposed

Building in-house quant and ML alongside vendors; across firms surveyed in 2026, fewer than a quarter run one platform for public and private assets.

LARGE AND GLOBAL
Advantaged

Proprietary regime and factor models, run at one large manager since 2014, and the scale to build AI in-house.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST02

AMAT

P02-ST03

Risk Management and Scenario Analysis

For: Investment Teams | Risk and Compliance

Risk management sits at the center of the industry's AI agenda: in a 2026 EY survey, risk identification and emerging-risk horizon scanning was the most common AI application in risk, cited by 40% of firms using AI. Use tapers toward the harder governance tasks, falling to 20% for regulatory change management. Traditional stress testing fails when it is needed most, because linear and static assumptions break down in the nonlinear, time-varying crises it exists to anticipate. Machine-learning stress methods are designed to capture those dynamics, but the models doing the watching are fragile in their own right. The disciplined response is to widen what risk can see while keeping humans in the loop, so the risk function does not inherit the front office's model risk.

40%
of AI-using firms apply it to risk identification and horizon scanning, falling to 20% for regulatory change management (2026 EY survey)

3x
more high-risk cases missed under adversarial stress than in calm markets (one 2025 study)

WHERE FIRMS GO WRONG

Firms have applied AI quickly to forecasting and trading while leaving risk management under-developed, so the front office runs ahead of the oversight meant to constrain it. Many still calibrate stress tests and thresholds on calm-market data, which understates tail risk, and threshold alerts catch only known risks. AI is weakest in regime change, qualitative judgment and horizons beyond its training data, and treating a risk model as reliable there builds false confidence into the most consequential decisions.

AI adoption in WAM risk functions

40%
risk ID and horizon scanning
top AI/ML use

detection-led

20%
regulatory change mgmt
lowest AI use

Source: EY Wealth and Asset Management Risk Pulse Survey (2026).

Takeaway. Industrialize the detection and simulation behind risk decisions rather than automating the decisions themselves. Closing the gap between AI-driven alpha and AI-aware oversight turns risk from a gatekeeper into a strategic intelligence engine.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST03

RECOMMENDED ACTIONS

1. Start by running one machine-learning stress method alongside the existing linear model on one portfolio; build toward a machine-learning stress-testing suite (PCA, autoencoder and VAE) alongside traditional methods, rather than relying on a single linear model that fails in crises.
2. Make stress testing sector-aware and run models on rolling windows, so they reflect cross-sector dependencies and current conditions rather than a stale historical calibration.
3. Use distributional (Monte Carlo) stress scenarios to reveal tail risk, skewness and uncertainty that single-point stress tests miss.
4. Require documented explainability (for example SHAP or counterfactuals) for every AI model used in portfolio construction, allocation and risk, so humans can challenge outputs.
5. Calibrate decision thresholds and false-negative tolerances on stress-regime performance, not calm-market data, and stress-test models adversarially.
6. Architecturally bound risk and monitoring agents to escalate, not act, and keep a human decision on every material risk call.
7. Start by listing the AI models the front office uses and the risk control each has; build toward AI risk oversight that matches the pace of AI-driven alpha generation.
8. Track and diversify single-theme active-risk concentration, including hidden technology correlations across supposedly diversifying exposures.

SEQUENCE AND DEPENDENCIES

1. Implementation should start from the recognition that AI risk tooling is less mature than front-office tooling, and sequence accordingly.
2. The first is explainability method selection: SHAP values, heatmaps or counterfactual explanations should be chosen and documented up front.
3. Resist stacking multiple risk frameworks for their own sake.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns model challenge and the exceptions.
Second Line	Owns the ML stress-testing suite, explainability standards for attribution, adversarial and stress-regime validation, and cross-risk intelligence; deeper model-risk governance is developed in Pillar 5.
Third Line	Tests whether stress and risk models were validated as claimed, whether explainability is documented, and whether agent escalations are genuine human decisions rather than nominal ones.
Board and Executive	Sets risk appetite for AI-driven and single-theme concentration, treats reliance on a few AI providers as an enterprise risk, and owns systemic exposure.
Cross-Functional	Delivers data-quality and feedback-loop controls and pairs technology-driven risk tooling with strong governance as adoption scales.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST03

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Stress-test failure in crisis**	Deploy an ML stress-testing suite (PCA, AE, VAE), sector-aware and rolling-window, with multi-factor compound shocks.
HIGH **Adversarial fragility under stress**	Calibrate thresholds and false-negative tolerances on stress-regime performance; stress-test models adversarially.
HIGH **Unexplainable models block challenge**	Require documented explainability (SHAP, counterfactuals) for all production risk and allocation models.
HIGH **AI strategy fragility**	Apply tail-risk and cross-regime testing before relying on an AI strategy.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Stress-testing coverage	Share of portfolios under an ML stress-testing suite (PCA/AE/VAE)	Stressed VaR/ES/drawdown captured vs realized in actual drawdowns
Model explainability	Share of production risk and allocation models with documented explainability	Incidents traced to unchallenged model output
Adversarial and stress robustness	Thresholds calibrated on stress-regime data; adversarial tests run	False-negative rate under stress vs calm
Early-warning coverage	Risk factors under continuous AI monitoring; new holdings screened day one	Time to flag material conduct or risk signals

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE
Advantaged

Vendor and agentic monitoring give small teams continuous, multi-factor early-warning coverage that once needed a dedicated risk desk.

MID-SIZED
Exposed

Building in-house ML risk models alongside vendors, but exposed to overfitting and validation gaps while scaling.

LARGE AND GLOBAL
Advantaged

Proprietary ML stress testing, multimodal monitoring and integrated cross-risk architectures, with the data to run them.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST03

AMAT

P02-ST04

Trading, Execution and Liquidity Analysis

For: Investment Teams | Operations

Execution is the most AI-saturated part of the investment process, and algorithmic execution is now the default on electronically active buy-side desks. The algo wheel has matured from a tool for removing broker bias into the control layer for the whole execution lifecycle, from pre-trade intent through routing to post-trade transaction cost analysis. Expectations run ahead of deployment: in a 2025 study of buy-side equity traders, nearly 80% expected AI to significantly affect algorithm optimization, yet only about 15% had incorporated internal AI into their execution workflow. At institutional size price impact, not the bid-ask spread, is the dominant cost to manage. The disciplined approach is augmentation, with explainable and auditable execution logic and models validated against live-market frictions rather than clean backtests.

80%
of buy-side equity traders expect AI to significantly affect algorithm optimization (2025 study)

15%
have incorporated internal AI into their execution workflow, with a further quarter planning to within twelve months

WHERE FIRMS GO WRONG

Execution AI is only as good as the execution data behind it, and most desks lack enough of it, which bites hardest on the high-volume orders where execution matters most. Firms also treat automation as the source of value when gains come from how a wheel is designed and refined, and common mistakes include single-purpose wheels, too many brokers for the sample and mixed objectives. The deepest failure is specific to machine learning: market data are endogenous, so a backtest can mistake self-induced price moves for genuine predictive power.

The expectation-to-adoption gap in execution AI



Buy-side equity traders, 2025 study. Source: Crisil Coalition Greenwich (2025).

Takeaway. Treat execution AI as an engineering and governance problem: a well-designed wheel that generates clean data, a TCA loop that turns it into better routing, and models tuned within explicit guardrails, measured against price impact. Real-time optimization without those foundations meets a market that, unlike a backtest, pushes back.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST04

RECOMMENDED ACTIONS

1. Start by splitting one broad wheel into an objective-led wheel for a single strategy type; build toward narrow, objective-led algo wheels that group like strategies, rather than broad, static, catch-all setups.
2. Layer liquidity, spread, volatility, urgency and time of day into wheel logic and preserve trader discretion on large or complex orders.
3. Start by referencing pre-trade cost estimates in the routing of one wheel; build toward pre-trade TCA embedded as the wheel's control layer, so routing rests on dynamic, cost-aware classification rather than static buckets.
4. Use post-trade TCA as a decision tool: normalize for order difficulty and peer context, and iterate wheels on a regular (typically quarterly) cadence.
5. Set minimum per-broker data thresholds before drawing conclusions, and keep a small test allocation to avoid overfitting to incumbents.
6. Require execution logic to be explainable and auditable, and use structured, no-manipulation-constrained price-impact models so profit derives from forecasting, not from moving prices.
7. Validate any AI or LLM execution strategy against live-market frictions (slippage, transaction costs, market impact) before relying on backtested performance.
8. Keep the trader on the exceptions: bound execution AI to defined parameters with human strategy design, real-time monitoring and override.

SEQUENCE AND DEPENDENCIES

1. Implementation should start from the recognition that the industry is still early in building AI execution capability and expertise, developing and refining systems rather than running mature ones.
2. Three dependencies precede any move toward more automated routing.
3. As an illustration of the guardrail principle rather than a standard to adopt, execution guardrails can be encoded as an explicit, enforceable rule set.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns wheel design, routing logic and the exceptions.
Second Line	Owns explainability and auditability standards for execution logic, validation against live-market frictions, governance of multi-pool liquidity routing to avoid adverse selection.
Third Line	Tests whether execution models were validated as claimed, whether price-impact and no-manipulation constraints are documented.
Board and Executive	Remains accountable for decisions made by AI tools just as for those made by people.
Cross-Functional	Delivers the data quality, thresholds and infrastructure that execution AI depends on, and pairs dynamic routing with the guardrails that keep it explainable and controlled.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST04

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Illusory alpha from price impact**	Use structured, no-manipulation-constrained price-impact models; require forecasting to be separable from the model's own impact; independently validate before deployment.
HIGH **Price-impact mis-specification**	Calibrate impact models on adequate data; quantify expected cost before acting on an idea; treat price impact, not spread, as the cost to manage.
HIGH **Backtest overstatement of AI strategies**	Require net-of-cost, live-condition validation before reliance; treat backtested performance as unproven; benchmark against simple baselines.
HIGH **Data insufficiency and quality**	Set minimum per-broker data thresholds before drawing conclusions; keep a test allocation; do not over-interpret thin samples.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Execution cost (price impact)	Share of orders with pre-trade cost estimates validated against outcomes	Realized implementation shortfall vs pre-trade estimate
Wheel data quality	Share of brokers meeting minimum data thresholds before conclusions are drawn	Dispersion of like-for-like outcomes across comparable orders
Automation coverage	Share of standard execution flow routed through governed wheels	Share of equity / FX flow executed with AI assistance
Model trustworthiness	Share of ML execution models with documented explainability and no-manipulation constraints	Incidents traced to endogeneity or unexplained execution logic

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Advantaged

Vendor and broker algo wheels give small desks standardized execution and robust TCA without a build, since wheels are now standard across electronically active buy-side desks.

MID-SIZED

Exposed

Building in-house execution analytics without either the simplicity of a pure vendor stack or the data depth of the largest firms.

LARGE AND GLOBAL

Advantaged

Proprietary automation at scale, with one global manager reporting in 2026 roughly 75 percent of equity and 85 percent of FX trading automated.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P02-ST04

AMAT

P02-ST05

Operations, Reconciliations and Exception Management

For: Operations

Operations is where AI has worked longest and least visibly: in a 2026 survey of financial services firms, four of the top five AI use cases were back-office functions. The proven return is capacity, since the same team can support more assets and more complex products. Reconciliation, settlement, fund accounting and transfer agency move toward continuous, prediction-led exception management, with AI preparing exceptions and people certifying. Final NAV approval, valuations of hard-to-price assets and regulatory determinations stay human.

73%
of asset managers use AI to improve operational efficiency, the most common form of adoption (2026 Mercer survey)

1 in 6
firms had achieved complete straight-through processing from investment decision to execution (2026 survey)

WHERE FIRMS GO WRONG

Firms meet structural change with more people rather than a redesigned operating model: in a 2026 survey of buy-side firms, a large majority report meeting T+1 by adding headcount, which does not scale to T+0. Many also treat operations AI as a faster way to run the existing batch process.

The operating-model shift in middle and back office operations

Dimension	Batch and manual (current)	AI exception-based (target)
Processing mode	End-of-day batch cycles	Continuous intraday monitoring
Human role	Routine matching and entry	Review and certify exceptions
Data handling	Manual, multi-format entry	Extract and validate at source
Exceptions	Accumulate, then chased	Flagged and prioritized
Outcome	Delay and error risk	Capacity and reliability gain

AMAT framework, drawing on Alpha Financial Markets Consulting and PwC Singapore.

Takeaway. The prize is a larger operations team in effect: the same people supporting more assets and more complex products. Redesign the model around exceptions and certification, know what stays human, and fix data quality before scaling.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST05

RECOMMENDED ACTIONS

1. **Run probabilistic matching with an explicit confidence threshold** on one high-volume reconciliation, then extend to intraday monitoring that auto-resolves routine matches.
2. **Predict trade failures** for one market or counterparty set, then prioritize exception queues by exposure and regulatory deadline.
3. **Keep human certification of final NAV, Level 3 valuations and regulatory determinations**, with AI detecting variances and drafting commentary.
4. **In transfer agency, deploy document intelligence**, AI-driven AML and KYC screening and conversational copilots for investor servicing.
5. **Embed operational copilots** that surface exception histories, draft communications and give procedural guidance.
6. **Treat added headcount as a signal to modernize data flows**, starting by moving the highest-volume exception process from end-of-day batch to intraday monitoring.
7. **Fix data quality at the point of need**, with human-in-the-loop validation of AI-extracted data.

SEQUENCE AND DEPENDENCIES

1. Implementation should start from the recognition that the reconciliation and data foundation is the prerequisite for everything more advanced, and sequence from there.
2. The dependencies are practical and mostly about data and handoffs.
3. On frameworks, resist stacking parallel models that no one reconciles.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the operating-model shift: builds probabilistic reconciliation with confidence thresholds, runs intraday exception management, and retains certification of final NAV.
Second Line	Owns data-quality standards, human-in-the-loop validation requirements and tiered guardrails for higher-risk processes such as KYC, and governs where automation may and may not remove the human.
Third Line	Tests whether human certification was genuinely applied, whether AI-extracted data was validated as claimed, and whether exception handling and recordkeeping are complete and auditable.
Board and Executive	Remains accountable for operational outcomes whether produced by people or AI tools, with accountability personal where a regime such as SM&CR applies.
Cross-Functional	Delivers the data quality, integration and end-to-end process redesign that operations AI depends on, so gains are not lost in the handoffs between automated steps.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST05

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Automating away human certification**	Keep human certification mandatory for final NAV, hard-to-value assets and regulatory determinations; design the model as AI-prepares, human-certifies with a named owner of the sign-off.
HIGH **Premature automation of complex, illiquid instruments**	Sequence automation from liquid, standardized processes outward; require independent validation and explainability before automating complex instruments; keep humans on non-standard cases.
HIGH **AI agents acting without human validation**	Keep agent actions in operations bounded and pre-approved, require human validation before any action that changes a ledger entry, an instruction or a NAV input.
MEDIUM **Solving settlement compression with headcount, not automation**	Treat added headcount as a trigger to modernize; invest in straight-through processing and redesigned data flows rather than more manual capacity.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Capacity (assets per operations FTE)	Share of routine matching and data entry automated	AUM or transaction volume supported per operations FTE
Straight-through processing	Share of transactions auto-resolved above confidence threshold	Share of end-to-end flow with no manual intervention
Exception cycle time	Share of breaks identified intraday vs end-of-day	Mean time to resolve reconciliation and settlement exceptions
Operating-model efficiency	Duplicate reconciliations and manual touchpoints removed	FTE effort and rule-inventory reduction from redesign

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Advantaged

Vendor and service-provider automation delivers institutional-grade reconciliation, NAV and exception management without building a team; a single asset manager reached an overnight.

MID-SIZED

Exposed

Building in-house operational automation without either the simplicity of a pure vendor stack or the volume of the largest firms.

LARGE AND GLOBAL

Advantaged

Scale to spread a fixed automation investment across large volumes, and the balance sheet to run production platforms; one named manager reported in 2026 running production reconciliation and custodian-communication platforms.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P02-ST05

AMAT

P02-ST06

Client Reporting, RFPs and Investor Communications

For: Operations | Investment Teams

Client communication is where AI meets the investor, and customer experience and engagement was the most-cited AI use case in a 2026 survey of financial services firms. One large manager reported in 2025 generating fund reports and commentaries in two minutes at 98% accuracy, work that previously took weeks. The durable model is AI drafts at scale and humans review and own, because conduct and suitability obligations make human review before anything reaches a client non-negotiable. The risk that defines the sub-topic is trust lost, since a hallucinated return figure or a misstated fee is the highest-stakes error the technology can make, and overstating AI use in marketing and RFPs has become a US enforcement theme. Handled well, disclosure and human accountability are what let the firm earn client trust.

98%

accuracy for fund reports and commentaries generated in two minutes at one large manager, work that previously took weeks (2025)

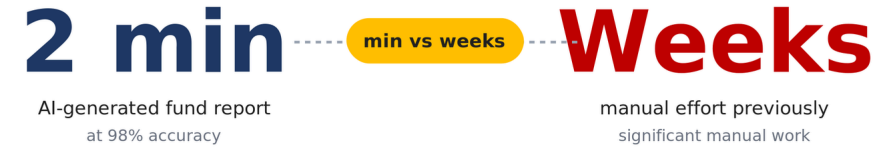
8%

of reported AI use cases were client-facing, 92% internal only (Luxembourg supervisory review, all entity types)

WHERE FIRMS GO WRONG

Firms mistake digitization for transformation: emailing PDFs built for print leaves the value of AI-native, interactive communication on the table. They also describe their AI use inconsistently across marketing, RFPs, DDQs and filings, and any overstatement becomes AI-washing, which US regulators have already acted on. A third failure is over-automating the relationship, since investors can often tell machine-produced communication and it can erode the credibility it was meant to scale.

Client reporting: minutes, not weeks



AI-generated fund reporting at a large manager. Source: ThoughtLab (Invesco), 2025.

Takeaway. The aim is the same people reaching far more clients with timely, personal output. Keep certification of every client-facing output, one consistent story about the firm's AI, and clear disclosure, and AI becomes a trust story rather than only an efficiency story.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P02-ST06

RECOMMENDED ACTIONS

1. **Certify every AI-drafted output before it reaches a client**, starting with commentary for one fund range or factsheet set.
2. **Draft answers to the most repeated RFP and DDQ questions** from an approved answer library with a validation step, then build an extract-validate-draft workflow.
3. **Tailor one recurring communication**, such as the quarterly letter, by client profile, then extend to plain-language translation of complex reporting.
4. **Stand up a single internal process** that reviews every description of the firm's AI use, in marketing, RFPs, DDQs, communications and filings, for accuracy and consistency.
5. **Disclose AI use to clients clearly, fairly and not misleadingly**, covering benefits and risks, and tell clients when they are interacting with an AI system.
6. **Give clients a route to a person** for any client-facing AI interaction, with plain-language explanations and human override.
7. **Size human-review controls to risk**: full review for client-facing and regulated outputs.

SEQUENCE AND DEPENDENCIES

1. Implementation should sequence from internal drafting to client-facing delivery, because customer-facing uses carry more direct reputational and compliance risk and therefore demand a higher degree of maturity and more comprehensive.
2. The dependencies are mostly about data, grounding and handoffs.
3. The realistic staging is: AI-assisted drafting of reports and commentary with full human review first, then RFP and DDQ automation with a validation step.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the client output: deploys AI drafting for reports, commentary and RFP/DDQ responses, personalizes by client profile.
Second Line	Owns disclosure standards and the single pre-publication review of all AI claims, pre-screens communications for unfair or unsubstantiated statements.
Third Line	Tests whether human certification was genuinely applied, whether client-facing AI output and AI-client interaction records are complete and retained.
Board and Executive	Remains accountable for client communications whether authored by people or machines, with accountability personal where a regime such as SM&CR applies.
Cross-Functional	Runs the single coordination process that keeps AI descriptions consistent across marketing, RFPs, DDQs, investor communications and filings.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST06

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **AI-washing and inconsistent AI claims**	A single internal coordination process that reviews all AI descriptions for accuracy and consistency before publication; disclose AI risks, not only benefits.
HIGH **Opaque robo-advice and suitability breach**	Keep generative advice human-assisted and explainable; block retail-facing advice patterns without a suitability profile; provide the rationale behind any recommendation.
MEDIUM **Authenticity and trust erosion**	Reserve human authorship for high-trust, relationship-critical communication; use AI to prepare, not to impersonate; disclose AI involvement.
MEDIUM **Digitizing paper instead of reimagining**	Redesign the communication and disclosure experience around AI-native, interactive formats rather than automating the existing document.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Report & commentary production	Share of reports and commentary AI-drafted	Time per client-reporting cycle
Content accuracy	Share of AI outputs passing human review unamended vs materially corrected	Client-facing errors and complaints
Personalization reach	Share of client communications tailored by profile	Client engagement and retention
Efficiency	Compliance-review effort automated	Cost per client-reporting cycle

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE
Advantaged

Vendor generative assistants, robo tools and chatbots deliver client-tailored summaries, commentaries and member-facing service without building a team.

MID-SIZED
Exposed

Building client-communication capability without the simplicity of a pure vendor stack or the volume of the largest firms, and most exposed to legacy data that slows personalization.

LARGE AND GLOBAL
Advantaged

Scale to run proprietary platforms that generate personalized reporting and commentary across large client volumes.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST06

AMAT

P02-ST07

Compliance Surveillance and Regulatory Monitoring

For: Risk and Compliance

Regulators moved first: as the IMF reported in 2025, FINRA uses machine learning to monitor all trading activity in stocks, options and bonds, so keyword-and-threshold monitoring inside the firm is no longer a defensible baseline. AI has also changed the threat, because spoofing, layering and momentum ignition can run at machine speed and across markets in ways traditional surveillance is likely to miss. The false positives that make compliance teams distrust their tools usually come from the operating model and the provenance of the rules, and accountability for compliance AI sits with senior leadership, not a vendor or a model.

£186m

in FCA fines in 2024/25 across all regulated sectors, against £42.6m in 2023/24 (cross-sector enforcement climate, not asset-manager cases)

WHERE FIRMS GO WRONG

Most firms have an operating-model problem that technology then amplifies. One global manager's order-management system generated false alerts across regions because a distributed monitoring model let staff without cross-business context create duplicate, incomplete or inaccurate rules, and fixing the governance, not the tooling, removed the noise. Treating compliance AI as point tools rather than a capability with owners and data lineage leads firms to over-trust it or abandon it after the first bad alert.

From keyword-rule monitoring to ML behavioral detection

Dimension	Keyword / rule-based	ML behavioral detection
Detection method	Fixed keyword lists, static thresholds	Models learn behavioral anomalies
Coverage	Sampled or rule-triggered alerts	Lineage evidences 100% reviewed
Alert quality	High volume of low-value alerts	Risk-ranked, fewer false positives
Scope	Single-channel, single-language	Voice, chat, email across languages

AMAT framework, drawing on Behavox vendor material (2025 and 2026).

Takeaway. The value comes from fixing the operating model first, tracing rules to source, pointing the technology at detection and synthesis, keeping human judgment on dispositions, and insisting every alert is explainable. Treated as a bolt-on, compliance AI only makes the noise faster.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P02-ST07

RECOMMENDED ACTIONS

1. Define one owner for compliance-monitoring rules and clarify jurisdictional responsibilities for global products. Start by consolidating duplicate rules in the region that generates the most false alerts; build toward one target operating model for compliance monitoring, with duplicate rules consolidated across regions.
2. Start by tracing the rules behind the noisiest alerts to the contractual or regulatory clause each enforces; build toward surveillance rules rebuilt from source documents with a governed, auditable source-to-rule mapping.
3. Require that every alert is explainable to an examiner. Start by running machine-learning behavioral detection alongside the existing lexicon on one surveillance channel; build toward machine-learning behavioral detection across trade and communications surveillance in place of keyword lexicons.
4. Keep a human in the loop where the technology is weakest: mandate human review and sign-off on high-stakes AML, KYC and market-abuse dispositions.
5. Configure third-party and vendor tools for supervision: reconfigure risky defaults, disable embedded chat features that escape capture, and ensure AI-assisted communications are recorded for review.
6. Put a generative-AI acceptable-use policy in place with active monitoring of employee AI use and a clear channel to report suspected misuse.

SEQUENCE AND DEPENDENCIES

1. The practical obstacle is almost never model accuracy in isolation; it is data provenance, sequencing and the division of labor between the tool and the team.
2. The second consideration is where to draw the human-machine line, and the evidence is consistent that it should be drawn at behavioral judgment.
3. Third, vendor and third-party governance is a first-class implementation task, not an afterthought.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Trading desks and client-facing teams own the primary market-abuse and conduct exposure, and bear the first-line duty to operate within market-abuse rules even where AI generated.
Second Line	Owns the surveillance operating model, rule provenance and the machine-learning detection stack, and is where the shift from keyword to behavioral detection is designed and governed.
Third Line	AI lets internal audit move from sampling to full-population testing of controls and policies, a material improvement in assurance completeness directly relevant to surveillance controls.
Board and Executive	Retains accountability for compliance and market-conduct failures, personal where a regime such as SM&CR applies; this cannot be delegated to a vendor or a model.
Cross-Functional	Vendor selection, data lineage and the generative-AI acceptable-use policy span the first and second lines and require joint ownership with technology and legal.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST07

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **AI-enabled market abuse**	Deploy machine-learning behavioral surveillance calibrated to the highest-risk pattern types; require explainable alerts and end-to-end lineage evidencing full coverage.
HIGH **GenAI-driven misconduct without intent**	Restrict data access by pipeline; log and review AI-generated investment-related outputs; mandate human sign-off before action on AI recommendations.
HIGH **Alert fatigue and false positives**	Redesign the monitoring target operating model; rebuild rules from source documents with governed source-to-rule mapping; risk-rank alerts.
MEDIUM **Model opacity and unexplainable alerts**	Make explainability a selection and acceptance criterion; retain human review where behavioral judgment dominates.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Detection coverage	Share of in-scope communications and trades under ML surveillance	Coverage evidenced to examiners via data lineage
Alert quality	False-positive rate on generated alerts	Confirmed cases per thousand alerts
Operating-model efficiency	Rule-inventory duplication rate	Headcount allocated to monitoring
Explainability	Share of alerts with a documented, examiner-ready rationale	Regulatory findings on unexplained dispositions

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE
Advantaged

(vendor-served).

MID-SIZED
Capacity squeeze

Key risk: region-by-region monitoring produces duplicate rules and false alerts before in-house capability is mature.

LARGE AND GLOBAL
Advantaged

Opportunity: centralize a global target operating model, with regional centers of excellence plus a shared rule-coding service.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST07

AMAT

P02-ST08

Finance, HR, Legal and Corporate Productivity

For: Finance; Legal; People and Change

Corporate functions are the largest category of AI use by volume in the available evidence, and the heaviest generative AI use sits in IT, communications, and finance and accounting (KPMG, 2025). In a 2026 survey of 107 private fund managers, more than half planned to use AI in HR, legal and compliance, and investor relations within a year. The value is real and quick to capture, but tools enter through individual productivity rather than a controlled program, so the firm often cannot say which functions use which tools on what data. The accountability for a finance disclosure, a contract review or an HR decision stays with the firm rather than the tool. Productivity gains that outrun assurance are a liability, not a win.

200,000

employees reached by one firm's enterprise generative AI copilot within a year (2026 report)

WHERE FIRMS GO WRONG

The challenge is not proving that AI helps but that adoption is running ahead of assurance. Tools enter the firm through individual productivity, so it often cannot say which functions use which tools on what data. Firms also treat corporate-function AI as free efficiency, yet in finance disclosures, legal review and HR decisions the accountability for a wrong output does not diminish because AI produced the first draft.

Where generative AI lands across corporate functions

Dimension	Manual today	GenAI-assisted
Finance	Manual disclosure drafting	AI drafts; human sign-off
Legal	Line-by-line contract review	Clause extraction, e-discovery
HR	Manual benefits/HR queries	GenAI HR/benefits assistant
Software/IT	Slow legacy-system change	Code completion, testing, migration

AMAT framework, drawing on PwC (2025), Norton Rose Fulbright (2025) and ABS/MAS (2025).

Takeaway. Inventory the tools, put a light but real governance layer over decentralized use, keep human sign-off on regulated outputs, and pursue the quick wins openly rather than in the shadows. Productivity captured with accountability intact is the win.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P02-ST08

RECOMMENDED ACTIONS

1. Start by inventorying which AI tools finance, legal and HR use, and on what data; build toward a maintained inventory of AI use across all corporate functions, kept current before any expansion, closing the gap between adoption and governance.
2. Put an approved-tools list and an acceptable-use standard in place, with mandatory human review before any AI-generated output is used in production or filed externally.
3. Keep a human in the loop on all regulated-function outputs: finance disclosures and estimates, legal review, and HR decisions require a named human owner and sign-off.
4. Start by deploying one approved drafting or transcription tool in the function with the heaviest administrative load and measuring the time recovered; build toward accessible drafting, summarization and transcription tools deployed deliberately wherever they remove administrative load.
5. Treat AI-assisted code as unassured until reviewed: require code review and testing of AI-generated code before it reaches production, especially for legacy modernization.

SEQUENCE AND DEPENDENCIES

1. Implementation in corporate functions is less about model selection and more about governing decentralized use without smothering the quick wins.
2. The one implementation area that needs its own discipline is software and legacy modernization, where AI-assisted coding is genuinely powerful for documenting and translating legacy systems but where unreviewed output.
3. The guardrail pattern that works is transversal tool governance (an approved-tools list and acceptable-use standard applied across the firm) combined with per-use-case restrictions and mandatory human review before production.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns adoption and the quick wins, and owns the first-line duty to review AI-assisted outputs before they are used or filed.
Second Line	Owns the approved-tools list, acceptable-use standard and the guardrail pattern (transversal plus per-use-case, with mandatory human review).
Third Line	Tests whether AI use is inventoried, governed and assured, and whether AI-generated outputs in regulated functions carry a documented human sign-off.
Board and Executive	Retains accountability for AI-assisted work product in regulated functions, personal where a regime such as SM&CR applies.
Cross-Functional	Tool governance, data controls and the acceptable-use standard span finance, HR, legal and technology and require joint ownership.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST08

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Confidential data exposure in external tools**	Approved-tools list and acceptable-use standard; restrict which tools may touch which data; inventory use across functions.
HIGH **Unassured AI-generated output quality**	Mandatory human review before production; code review and testing of AI-generated code, especially in legacy modernization.
MEDIUM **Over-reliance without human sign-off in regulated functions**	Human in the loop on all regulated-function outputs; named owner and sign-off.
MEDIUM **Ungoverned, decentralized adoption**	Inventory and govern AI use as a program, not a collection of tools; assign clear accountability for AI-assisted work.

METRICS SCORECARD

AREA	LEADING	LAGGING
Adoption depth	Share of employees with access to approved tools	Active use frequency
Developer productivity	AI-assisted code review throughput	Time-to-market for new applications
Assurance	Rate at which reviewers edit, override or reject AI output, trended.	Errors traced to unreviewed AI output
Governance coverage	Share of AI tools on the approved-tools list	Uninventoried tool use found in audit

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE
Advantaged
(quick wins).

MID-SIZED
Embedding.

LARGE AND GLOBAL
Advantaged
(scale).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST08

AMAT

P02-ST09

Sales, Distribution and Relationship Management

For: Senior Leadership | Operations

As product differentiation erodes, control of the client relationship is becoming the defining source of advantage, and BCG's 2026 report finds retail investors accounted for 61% of global AuM expansion between 2020 and 2025. BCG estimates that over three to five years AI can raise the clients each relationship manager covers three- to five-fold, unlock 35 to 55% of capacity, and add 0.5 to 1.0% of assets a year in net inflows. That only holds one way round. AI strengthens a sound distribution model and cannot substitute for one, so the discipline is to fix the commercial model first and treat segmentation as strategy. A recurring mistake is deploying AI before the commercial model is defined.

61%
of global AuM expansion between 2020 and 2025 came from retail investors (BCG, 2026)

35 to 55%
of relationship manager capacity could be unlocked by AI over three to five years (BCG estimate)

WHERE FIRMS GO WRONG

Most firms recognize the shift but have not reorganized around it: in a 2025 Broadridge and Money Management Institute survey of 142 member firms, 87% agree younger investors need different products or service models, but only 30% say their firm is actively reorienting growth strategies. The deeper failure is sequence. Without a complete, dynamic view of the client universe, coverage defaults to history and relationships, AI has nothing meaningful to optimize, and segmentation becomes the binding constraint.

AI as a distribution force multiplier

BCG estimates, over a three-to-five-year horizon, on a sound distribution model

3-5x
clients per RM
vs today

35-55% capacity

0.5-1.0%
net AuM inflows
per year

Source: BCG, 2026.

Takeaway. Treat segmentation as strategy, embed AI in the daily adviser workflow, and govern profiling and suitability as you scale. Buying tools ahead of the model spends more to reach the wrong clients faster.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST09

RECOMMENDED ACTIONS

1. **Segment the existing client book** on assets and fee potential from current CRM data, then build a complete, dynamic view of the client universe before deploying AI at scale.
2. **Map coverage logic and sales motion against the four pillars of the distribution model** (segmentation, coverage, sales motion, data) and fix all four before layering AI on top.
3. **Embed one AI tool in a single adviser workflow**, such as meeting preparation or call summaries, then extend across daily workflows.
4. **Keep the decision, the suitability judgment and the relationship with the adviser**, piloting next-best-action prompts for one client segment.
5. **Automate onboarding and suitability data capture** for one retail or next-generation product line to serve smaller-ticket clients profitably.
6. **Record the evidence behind every AI-driven recommendation** so advisers, clients and supervision can see why an action was suggested.
7. **Govern automated profiling**: define the lawful basis, run privacy impact assessments and keep human oversight of targeting.

SEQUENCE AND DEPENDENCIES

1. Implementation should follow the same sequence the recommendations imply: commercial model first, data foundation second, AI across the funnel third.
2. The first is data and CRM integration: the largest named deployments described here unify fragmented data, research and workflow tools into a single adviser surface.
3. The reason is practical rather than theoretical, because segmentation and the four pillars are what give AI something coherent to optimize.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns segmentation, coverage and the sales motion: builds the dynamic client view, uses AI for prospecting, next-best-action and proactive triggers, and retains the recommendation.
Second Line	Owns suitability governance, profiling and data-protection compliance, and marketing supervision: defines lawful basis for profiling, requires privacy impact assessments.
Third Line	Tests whether segmentation and coverage decisions are governed as claimed, whether profiling has a lawful basis and human oversight.
Board and Executive	Sets the strategy that treats ownership of the client relationship as the competitive priority, and holds accountability for distribution conduct and client outcomes.
Cross-Functional	Delivers the unified adviser data surface, CRM integration and clean external signals that AI distribution depends on, so capacity gains are not lost in fragmented handoffs.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST09

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Layering AI on a broken distribution model**	Fix the four pillars first (segmentation, coverage, sales motion, data); treat segmentation as strategy and make it the control layer before any AI deployment.
HIGH **Advice-like AI outputs and paid placement outside the perimeter**	Map every AI channel, the firm's own or a third party's, through which the firm's funds are presented.
MEDIUM **Automated profiling and privacy or conduct exposure**	Define lawful basis and run privacy impact assessments; keep human oversight of profiling-driven targeting; disclose and honor the right to object.
MEDIUM **Over-reliance on AI recommendations eroding adviser judgment and suitability**	Keep the adviser accountable for the recommendation and the suitability decision; treat AI outputs as prompts, not instructions; track the rate at which advisers edit.

METRICS SCORECARD

AREA	LEADING	LAGGING
Adviser coverage and capacity	Share of prospecting, prep and admin automated per adviser	Clients covered per relationship manager
Net flows from distribution	AI-sourced qualified leads and win-rate on AI-prioritized prospects	Incremental net AuM inflows attributable to AI-enhanced distribution
Client retention and depth	Proactive outreach triggered by AI life-event and outflow signals	Attrition rate and share-of-wallet
Adviser time reallocation	Preparation and admin time removed by AI tools	Share of adviser time spent in client-facing activity

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Advantaged

AI flattens the cost and information gap that once kept small managers out of retail, making it economic to serve smaller-ticket clients through robo-advice and low-cost personalized digital journeys.

MID-SIZED

Exposed

Building adviser-facing AI and a dynamic segmentation model without the data depth of the largest firms, and most exposed as asset and wealth management converge on the client relationship.

LARGE AND GLOBAL

Advantaged

Scale to run proprietary adviser platforms and own the client relationship: J.P.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST09

AMAT

P02-ST11

ESG, Stewardship and Sustainability Workflows

For: Investment Teams | Risk and Compliance

The problem in ESG is not a shortage of data but its state: selectively disclosed, self-reported and inconsistent across providers. AI does not solve that, but it makes imperfect data usable and timely. Controversy detection ran two to four weeks ahead of rating updates at one manager, coverage extends to names vendors do not rate, and stewardship analysis becomes systematic at portfolio scale. The same tools bring greenwashing that AI cannot yet reliably distinguish from genuine progress, model bias and opacity, so every AI signal needs review.

58%

of buy-side operations leaders name ESG investing the top area of opportunity for technological innovation, 81% in North America (2025 survey of 200)

2 to 4 weeks

lead of AI controversy detection over rating updates at one manager

WHERE FIRMS GO WRONG

Buying more vendor data misreads the problem, because vendor feeds carry the same coverage gaps and provider-specific biases. The technology is also less mature than vendor marketing implies: in one 2025 review, LLM-based ESG scoring is promising but early-stage, with limited validation against established scores. AI effort has concentrated almost entirely on the environmental pillar.

Matching AI techniques to ESG data limitations

Dimension	ESG data limitation	Mitigating AI technique
Coverage	Gaps (Scope 3, EM)	Geospatial AI, fusion
Credibility	Greenwashing	NLP + cross-source checks
Uncertainty	Unobservable futures	Scenario + stress tests
Dispersion	Scattered documents	Text mining / extraction
Fragmentation	Too many ratings	Feature selection
Consistency	Providers diverge	Ensembles + robustness

Source: Garrido-Merchan et al., AI-Powered Sustainable Finance, 2026.

Takeaway. Set the right expectation with the investment committee: AI makes imperfect ESG data usable, it does not fix it. Pair it with a three-tier data architecture, human review of every flag, real explainability and honesty about study-specific results, and push it into the social and governance pillars.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P02-ST11

RECOMMENDED ACTIONS

1. Add an AI controversy screen to the vendor ESG feed for one portfolio, with analyst review of every flag, building toward three tiers: vendor ratings, AI monitoring and analyst review.
2. Monitor news for the holdings with the thinnest vendor coverage, building toward streaming monitoring that updates within hours rather than quarterly.
3. Log the data-quality caveats of every estimated ESG signal, starting with unrated names already held and extending to small-cap, emerging-market and private names.
4. Embed ESG and climate-risk data as financially material model inputs and make the return-versus-ESG trade-off explicit where you optimize.
5. Treat every AI ESG output as a first-pass screen, with analyst or portfolio-manager review before it drives an investment, risk or voting decision.
6. Record which ESG factors drove each model-driven allocation, building toward SHAP and LIME explanations and stakeholder-facing visualizations.
7. Start with one social or governance use, such as screening for governance red flags, building toward the under-served social and governance pillars.
8. Govern the AI's own footprint with Green AI practices, and engage portfolio companies on the climate and water impact of their AI.

SEQUENCE AND DEPENDENCIES

1. In this guide's view, the first implementation step is the taxonomy: a controversy-monitoring system should start from a structured controversy taxonomy, by pillar and severity.
2. Use fine-tuned domain models rather than general-purpose ones, ingest news, filings and alternative data continuously so models update within hours rather than on the rating cycle.
3. The build-versus-buy question has a defensible middle path.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Run AI controversy monitoring and day-one screening, review every AI-flagged item before acting, and use AI-prepared briefs for engagement and proxy analysis.
Second Line	Validate greenwashing and model robustness, monitor cross-provider rating divergence and data quality, and own the ESG-AI model-risk and interpretability controls.
Third Line	Test ESG-AI model governance, bias across geography, sector and size, and data lineage from vendor feed to investment decision.
Board and Executive	Holds accountability for ESG-AI outcomes, personal where a regime such as SM&CR applies; sets ESG-AI oversight and explainability expectations, agree who owns AI stewardship.
Cross-Functional	A dedicated ESG data owner and cross-functional steering committee maintain the ESG taxonomy, controls and regulatory adaptability that every AI application depends on.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST11

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Greenwashing robustness**	Treat AI as first-pass screening only; require analyst confirmation of flagged items and check figures and claimed actions in AI-assisted ESG text against the source filing.
HIGH **ESG data quality and vendor dependency**	Build on diverse, multi-source data with a vendor-agnostic taxonomy; monitor rating divergence; assign central ESG-data ownership and controls.
MEDIUM **Model bias across geography, sector and size**	Test for bias across geography, sector and company size before deployment; monitor false-positive rates by segment; document tested populations.
MEDIUM **Over-reliance and loss of judgment**	Keep the analyst accountable for the decision; treat outputs as prompts; track the rate at which analysts edit.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Controversy detection speed	Share of controversies first flagged by AI vs vendor updates	Price impact avoided on flagged names
ESG data coverage	% of holdings with an AI-generated signal where vendor coverage is absent	Reduction in uncovered small-cap / EM names
Model quality	False-negative rate on a labeled controversy test set; bias tested by segment	Analyst-confirmed true positives vs AI flags
Stewardship reach	Share of meetings / holdings with AI-prepared engagement briefs	Engagement themes tracked over time

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE
Advantaged
Opportunity: close ESG coverage gaps in small-cap and emerging-market names where vendor data is thin.

MID-SIZED
Capacity squeeze
Exposure: rising ESG disclosure and stewardship expectations without a large data team.

LARGE AND GLOBAL
Advantaged
Opportunity: day-one, universe-wide AI screening and stewardship at scale.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P02-ST11

Pillar 3 Executive Summary

For: Senior Leadership; Data and Technology · Also relevant to: Investment, Risk, Compliance, Operations

THE ARGUMENT

AI in asset management is bounded by the firm's data foundations. Data infrastructure is among the most cited barriers to AI adoption in investment management, and the research, decisions and operations a firm wants AI to change all depend on its data, its knowledge and the record of its past decisions. Pillar 3 covers the data-side capabilities AI depends on: strategy, integration, quality, governance, privacy, and knowledge and decision intelligence. Its case is that scaling AI depends less on having the most sophisticated models than on treating data as the critical constraint and governing it as a proprietary asset.

80%+
of data available online is unstructured (State Street IM), and much of the signal sits there

58%
of financial-services tech and operations leaders agreed they have a clear data strategy (Broadridge, 2025)

81%
of private fund managers surveyed increased alternative-data use, 2024 to H1 2025 (late-2025 survey)

WHY NOW

Much of the signal sits in unstructured material that models can now read at scale, so integrating structured and unstructured data has become the gating capability for AI. Yet in Broadridge's 2025 survey of financial-services technology and operations leaders, only 58% agreed that they have a clear data strategy, and 81% of private fund managers in a late-2025 Lowenstein Sandler survey increased their alternative-data use even as a widely available source's edge can erode. Data is the point at which AI transformation either becomes possible or quietly stalls, which is why AMAT's position is to fund data work from the start and run it alongside model work, rather than treating data as a later problem.

RISKS TO AVOID

Buying AI first and treating data as a later problem or a reporting-era clean-up, rather than the critical constraint on AI (in AMAT's view, a common way programs stall). **Ungoverned access and privacy exposure** (the firm holds information whose exposure can be unlawful, and many AI tools retain what is put into them). **Over-paying for commoditized alternative data** whose edge may have eroded, instead of building proprietary signal.

Bottom line. Pillar 3 is the foundation: AI transformation is bounded by the firm's data. Scaling AI depends less on the most sophisticated models than on treating data, knowledge and decision memory as a governed proprietary asset: fund it from the start alongside model work, govern it as one capability, and build signal where others only buy it. Get the data foundations right and the rest of the guide becomes achievable; get them wrong and the program can quietly stall.

Chapter Leads, Pillar 3: Andrew Chin

Explore Pillar 3 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P03-summary

PRIORITY MOVES FOR LEADERSHIP

1. **Fund data work as a first-order AI priority from the start**, run it alongside model work, and build the governed foundation in stages from the data the first use cases need.
2. **Make unstructured-to-structured integration designed infrastructure**, with entity resolution and human-validated extraction served as reusable, governed data.
3. **Redefine data quality as owned, traceable and lineage-recorded**; name a golden source per domain and prepare readiness use case by use case.
4. **Treat governance and access as one capability**: classify and label data and enforce it through role-based, just-in-time access, not advisory permission.
5. **Build a knowledge and decision-intelligence layer** (thesis ledger, postmortems, retrieval) so the firm learns from its own decision outcomes, and test whether that builds an advantage that is hard to copy.
6. **Govern privacy and vendor licensing as live AI exposures**: non-retaining instances and information barriers for sensitive work; value datasets before buying.

IN THIS PILLAR

P03-ST01	Data Strategy for AI Transformation	66
P03-ST02	Structured and Unstructured Data Integration	69
P03-ST03	Knowledge Management, Institutional Memory and Decision Intelligence	72
P03-ST04	Data Quality, Lineage and Ownership	75
P03-ST07	Data Governance and Access Controls	78
P03-ST08	Data Privacy Controls and Information-Barrier Architecture	81
P03-ST09	Alternative Data and External Research Integration	84
P03-ST10	Data Marketplaces and Vendor Data Strategy	87

Explore Pillar 3 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P03-summary

P03-ST01

Data Strategy for AI Transformation

For: Senior Leadership; Technology; Data and Analytics

The competitive frontier has shifted from model sophistication toward proprietary data and operational excellence. In a 2024 survey of 106 asset managers and 94 asset owners, 71% named increased reliance on data and analytics as a top industry trend for the next three to five years. Yet data infrastructure is among the most cited barriers to AI adoption, and in a separate 2025 survey of 509 financial-services technology and operations leaders only 58% agreed they have a clear data strategy that lets them maximize the return on their technology. The strategic shift is threefold: treat data as the foundation AI depends on, architect it for AI and agent consumption rather than reporting alone, and treat proprietary data as competitive differentiation. Sequencing, ownership and control decisions decide whether an AI program scales or stalls.

71%

of 106 asset managers and 94 owners name reliance on data and analytics a top industry trend (2024 survey)

58%

of 509 technology and operations leaders agree they have a clear data strategy (separate 2025 survey)

WHERE FIRMS GO WRONG

A common failure, in AMAT's view, is inverted sequencing: firms buy AI tooling and treat data as a later problem, leaving use cases without the data they need. The problem is sharper in asset management because each asset class tended to acquire its own systems, leaving siloed environments with no front-to-back integration. Practitioners describe silos and bolt-on point solutions as the enemies of progress, and the barrier is not a shortage of AI ambition but a data foundation that cannot yet support it.

Two surveys, side by side

71%	58%
named data and analytics a top trend	agreed they have a clear data strategy
BNY Mellon (2024), 200 managers and owners	Broadridge (2025), 509 FS tech/ops leaders

Separate surveys, different populations and questions; not comparable. Sources: BNY Mellon, 2024; Broadridge, 2025.

Takeaway. Data strategy is where AI transformation becomes possible or quietly stalls. The immediate move for most firms is not to buy more AI but to fund the data work, start with the data the first use cases need, run model work alongside it, and build the governed foundation in stages.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST01

RECOMMENDED ACTIONS

1. **Set the data strategy as the central pillar of the AI strategy**, funded as a first-order priority.
2. **Bring the data your first AI use cases need onto a common layer**, with a named owner for each source, and build in stages toward a unified front-to-back platform layered over existing systems.
3. **Run data and model work in parallel**: bring priority data to a fit-for-purpose level, then keep cleaning alongside the model work.
4. **Model the underlying data for target use cases** instead of rewriting every legacy system.
5. **Implement data-quality controls before data enters production models**, and monitor live feeds continuously.
6. **Document the data behind your first use case as a dataset with a named business owner** before production, then build reusable data products under domain ownership.
7. **List the unstructured sources your first use cases depend on**, with an owner for each, and treat unstructured and synthetic data as a priority.

SEQUENCE AND DEPENDENCIES

1. The hardest part of executing a data strategy is resisting the two opposite failure modes: doing nothing until the estate is perfect, and doing everything at once.
2. Elevating data to a governed corporate asset means clear ownership that moves the organization from project-based pipelines to reusable data products and manages data for maximum value rather than rebuilding.
3. The practical path avoids both: build the priority data foundations for your first use cases rather than waiting for a clean estate.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns data requirements and data quality at source: investment, operations and data owners define what each use case needs and steward their domains' data.
Second Line	Treats data readiness as a governance prerequisite for safe AI deployment, not a technology project.
Third Line	Provides assurance over data lineage, ownership, and quality controls.
Board and Executive	Owns the link between data strategy and AI strategy, and holds accountability for it, personal where a regime such as SM&CR applies.
Cross-Functional	A single source of truth calls for ownership that spans multiple functions; the data-mesh model deliberately distributes domain ownership under centralized governance to make that work.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST01

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Structural data fragmentation**	Consolidate onto a unified, front-to-back data platform (lakehouse plus data mesh), built in stages from the data the first use cases need and layered over existing systems rather.
HIGH **Data treated as an afterthought**	Fund data work from day one and run it alongside model work, with a fit-for-purpose threshold for each use case.
HIGH **Poor data quality feeding models**	Engineer data-quality controls (noise reduction, outlier handling) before using data in production models, and monitor the data feeding live models continuously.
HIGH **Silos starve AI of context**	Model data for target use cases and build reusable, governed data products across domains.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Data strategy clarity	Share of priority AI use cases with a named data owner and mapped lineage	Board-agreed data strategy in force across the firm; share of priority AI workflows meeting their agreed decision-quality.
Unified data access	Share of front/middle/back-office data available through a common layer	Reduction in duplicate sources, and in manual reconciliation hours against a baseline taken before the first AI use.
Data quality at source	Share of priority datasets passing their agreed acceptance test before first production use.	Investment or model incidents traced to data defects; time to quarantine and clear a defective feed
Reusable data products	Count of governed, reusable data products versus one-off pipelines	Median time from an approved use-case request to fit-for-purpose data

BY FIRM SIZE

BOUTIQUE

Exposure: in LLM development and fine-tuning, data asymmetry favors larger incumbents with more data and resources, which can disadvantage smaller firms.

MID-SIZED

Binding constraint: fragmented, asset-class-specific systems create particular challenges here, as growth outpaces the coherence of the data estate.

LARGE AND GLOBAL

Opportunity: proprietary-data scale is the primary differentiation lever, and, on Strata FSC's view, the locus of alpha is shifting from model sophistication toward data.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P03-ST01

P03-ST02

Structured and Unstructured Data Integration

For: Technology; Data and Analytics

Earnings calls, filings, broker research, contracts, capital-call notices and private-asset reports do not live in a system of record, and over 80% of data available online is unstructured. In a self-reported 2026 survey of 131 asset managers, unstructured text was the most common AI input at 68%, narrowly ahead of traditional financial data at 66%. Models can now read that material at close to zero marginal cost and are learning to join it to structured data in one representation. The practical answer is designed integration: resolve entities once, transform documents to data with human-in-the-loop validation, and design for multiple data types. In a Grant Thornton survey, integrating data across departments was the data strategy AI leaders reported most often.

68%
of asset managers use unstructured text as an AI input, the most common (self-reported 2026 survey of 131)

53%
name sourcing and integrating unstructured data such as PDFs a major data challenge, 61% in EMEA (2026 survey of 200 buy-side operations leaders)

WHERE FIRMS GO WRONG

The barrier is rarely the model but the state of the data underneath it: most firms run on disparate systems never designed to talk to each other, and unstructured data remains largely ungoverned. Three failure modes recur: single-modality thinking, which hits a capability ceiling because financial decisions are multimodal; unmanaged heterogeneity in formats, timing and latency; and weak entity resolution, which every downstream workflow inherits.

Two ways to bring data together

Dimension	Point extraction	Designed integration
Data scope	One modality at a time	All modalities joined
Pipelines	Bespoke per use case	Reusable data products
Entities	Reconciled ad hoc	One shared entity layer
Trajectory	Hits a ceiling fast	Compounds across desks

AMAT framework, drawing on Bigdata.com (2025) and Jadhav and Mirza (2025).

Takeaway. Resolve entities, transform documents to data under human validation, fuse modalities, and serve the result as reusable, governed data products. Start with one high-value document type, get the entity layer and validation loop right, and let the capability compound.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST02

RECOMMENDED ACTIONS

1. **Map the companies, issuers and funds in one priority workflow** to a single maintained identifier list with a named owner, then build a shared entity-resolution layer.
2. **Build document-to-data pipelines for your highest-value unstructured sources**, with human-in-the-loop validation before output reaches a portfolio manager or client report; log vendor-embedded extraction in the AI inventory.
3. **Keep the source document, timestamp and entity identifier beside every extracted field**, so other data types can be joined later.
4. **Publish your first pipeline's output as a documented, owned dataset** when a second team needs it, then build reusable, governed data products.
5. **Write down the relationships your first workflow depends on**, such as issuer to parent, and move to schema-flexible representations once several workflows share them.
6. **Record the provenance, owner and usage rights of each unstructured source** before it feeds a decision, and govern it to the same standard as structured data.
7. **Use zero-shot techniques** to extend text-based analysis to new themes without new labeling.

SEQUENCE AND DEPENDENCIES

1. The build order is the main lever.
2. Document length is a dependency to test early: on long filings and lengthy earnings transcripts, model limits and retrieval quality can cause evidence to be missed.
3. Resolve entities before scaling extraction, start with the few document types that carry the most value rather than attempting everything at once.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Investment, research and operations teams build the extraction and fusion pipelines and validate outputs before they reach a portfolio manager or a report.
Second Line	Risk and compliance set standards for provenance, licensing and model use over integrated data and challenge the quality feeding AI.
Third Line	Internal audit tests data lineage and provenance end to end, from source document to the integrated record used in a decision; lineage mechanics are P03-ST04.
Board and Executive	The board funds integration as strategic infrastructure rather than a series of point tools, treats data as an asset managed for reuse.
Cross-Functional	Entity resolution and reusable data products are run as shared services across desks rather than rebuilt per team.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST02

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Entity mis-resolution**	Maintain an owned entity-resolution layer with monitored match and false-link rates, treated as standing infrastructure.
HIGH **Ungrounded extraction**	Require human-in-the-loop validation and source-grounding before extracted data reaches a portfolio manager or a client report; see quality assurance in P07.
HIGH **Ungoverned unstructured data**	Extend governance to unstructured data at creation and define provenance and usage rights per source.
HIGH **Point-in-time mismatch**	Store publication, retrieval and effective times with every extracted field; test as-of joins and sampled revisions before release; quarantine mismatches (data quality assurance is P07-ST08).

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Data-type coverage	Modalities and sources in scope	Share of decision-relevant sources available to AI
Extraction accuracy	Validation pass rate on extracted fields	Downstream error or restatement rate
Multimodal coverage	Modalities fused (text, tables, audio, images)	Signal lift versus single-modality baseline
Entity resolution	Entity match and dedup rate	Mis-linked-entity incidents

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The realistic move is to buy, not build: targeted extraction for the few document types that carry the most value, such as private-asset deal notes and offering documents.

MID-SIZED

The exposure is integration debt: pipelines built one use case at a time proliferate until no one can maintain them.

LARGE AND GLOBAL

Scale is an advantage and legacy is the constraint.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P03-ST02

P03-ST03

Knowledge Management, Institutional Memory and Decision Intelligence

For: Investment Teams; Senior Leadership; Data and Analytics

Generic models and the tools built on them diffuse, so on DiligenceVault's account the edge may no longer sit in the model. What can compound is the institutional knowledge and structured decision data wrapped around it. In NVIDIA's 2026 survey of 839 financial-services respondents, knowledge management and retrieval was the most-cited agentic workflow among those using or assessing AI agents, at 56%, ahead of internal process optimization at 52% (not an asset-management sample). A firm that captures the reasoning behind each decision and learns from it may build an asset that is hard to copy, though that has to be tested against its own outcomes.

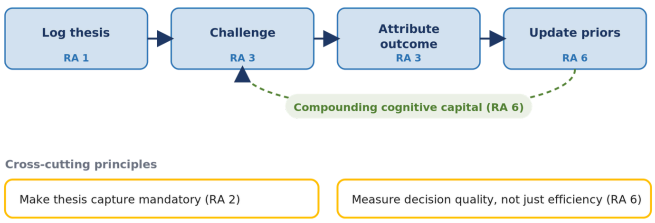
56%

of financial-services respondents using or assessing AI agents name knowledge management the top workflow, ahead of process optimization at 52% (NVIDIA 2026 survey of 839)

WHERE FIRMS GO WRONG

Wuebben argues that most firms are organized in ways that degrade the beliefs they produce: information is compressed at every handoff so a conditional thesis arrives as 'we like the name', committees reward conviction over calibration, and outcome attribution is rarely connected back to the originating hypothesis. Knowledge management is also treated as back-office housekeeping, which leaves duplicate research and knowledge that walks out of the door with experienced people. The pattern is documented for private equity; this guide extends it to other asset managers as a reasonable inference.

The closed learning loop



AMAT framework, drawing on AllianceBernstein (2026) and DiligenceVault (2026).

Takeaway. Start with the thesis ledger and make capture mandatory, build the postmortem-and-retrieval engine on top, give knowledge management a named owner, and close the loop so decision outcomes inform future judgment. The aim is to build the means to improve judgment over time, not only to manage documents.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST03

RECOMMENDED ACTIONS

1. **Log the theses behind new positions** in a simple structured template, then build a thesis ledger recording claim, mechanism, falsifiability conditions, horizon and calibrated probability.
2. **Make thesis capture mandatory for material decisions**, with a time-stamped rationale and named owner, and record and review every exception.
3. **Run a structured postmortem on closed positions with losses above a threshold**, plus a sample of other outcomes, then build an engine that retrieves analogous theses with base rates.
4. **Treat knowledge management as a governed capability**: assign a decision-record owner and a knowledge leader, build expert directories and deal memory, and assess readiness first.
5. **Encode judgment into firm-owned, tested prompt libraries** and a defined diligence architecture, refined after every cycle.
6. **Close the loop**: capture structured feedback from teams and reviews, record the correction and reviewer's reason, and measure decision quality, not just efficiency.
7. **Make the thesis ledger and postmortems searchable** by the team that writes them, then build a grounded, permission-filtered knowledge system over the firm's own corpus.

SEQUENCE AND DEPENDENCIES

1. The thesis ledger and mandatory capture come first because every later capability (postmortems, analogue retrieval, calibration metrics) depends on a complete record.
2. Where an existing research, CRM, order-management or document platform already offers an AI knowledge feature, name its business owner and record the feature in the firm's AI inventory before use.
3. The dominant obstacle is cultural: a culture that rewards conviction over calibration creates friction against documenting falsifiable claims and running honest postmortems.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Investment heads are accountable for their teams' original theses, investment decisions and postmortem approvals; investment teams log theses, run postmortems.
Second Line	Risk and compliance challenge how the decision record and AI tools are used, accessed and escalated against firm policy.
Third Line	Internal audit tests that the thesis ledger and learning loop are complete and actually used (documentation and postmortem-completion rates) and assesses the design and operation.
Board and Executive	The board sets direction and risk appetite for knowledge management and the learning-and-memory layer and holds accountability for AI-influenced decisions.
Cross-Functional	A named knowledge leader runs the shared KM ecosystem, directories and deal memory across teams; a run-state steward maintains access, taxonomy.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST03

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Broken learning loop**	Run structured postmortems on losses above the threshold and on a sample of other outcomes, each linked to the original thesis.
HIGH **Institutional-memory loss**	Capture procedural knowledge in the KM capability and in embedded copilots so it persists independently of individuals.
HIGH **Unchallenged AI outputs**	Embed automated challenge (base rates, historical analogues, crowding, counter-arguments) into the workflow; human-oversight obligations are P05-ST08 and P06.
HIGH **Self-confirming learning**	Preserve the original time-stamped thesis and contrary evidence; show the retrieved source record; require human validation of causal attribution.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Thesis documentation	Share of material decisions in scope with a time-stamped rationale	Undocumented decisions and unreviewed exceptions
Forecast calibration	Predicted vs realized probability, by layer	Brier score (reliability, resolution)
Postmortem completion	Losses above the threshold, and sampled other outcomes, with a structured retrospective	Repeated, uncategorized failure modes
Belief revision speed	Latency from disconfirming evidence to hypothesis update	Fat right tail of slow updates

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Less legacy fragmentation to overcome, so a lightweight thesis ledger plus copilots that capture what a handful of people hold is the pragmatic start.

MID-SIZED

This is where duplicate research, no cross-team transparency and siloed, unsearchable systems appear as AUM and complexity grow.

LARGE AND GLOBAL

Scale funds the full learning-and-memory layer and lets the firm compound cognitive capital; the constraint is bureaucratic knowledge management that no one uses.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P03-ST03

P03-ST04

Data Quality, Lineage and Ownership

For: Technology; Data and Analytics; Operations

In KPMG's survey of finance leaders across 13 sectors and a Precisely/LeBow survey of data leaders, data quality and readiness are among the most cited barriers to AI adoption, and data fit for reporting may not be fit to train or ground a model. In the 2026 Precisely/LeBow survey (not finance-specific), 43% of data leaders named data readiness, not model access, as the primary barrier to aligning AI with business goals. A defect a human analyst would catch in a spreadsheet becomes a systematic error once a model consumes the data thousands of times a day. Getting it right means treating data-management controls as a framework, readying data use case by use case, and naming an owner for every golden source.

43%
of data leaders name data readiness, not model access, the primary barrier to aligning AI with business goals (2026 Precisely/LeBow survey, not finance-specific)

50%
of 323 institutions surveyed by the Dutch regulator cite data quality as the biggest challenge (not only asset managers)

WHERE FIRMS GO WRONG

Data quality is often managed reactively and locally, so the same defects are fixed repeatedly and never at source. Golden sources are often undefined, and where lineage is not captured a firm cannot show where an AI output's input data came from. The failure is quiet: schema drift, unit mistakes and fiscal-period misalignment corrupt analysis without announcing themselves, and on AMAT's reading a silently wrong KPI is worse than a missing one.

Generic data hygiene vs AI-ready data

Dimension	Generic hygiene	AI-ready data
Goal	Clean records for reporting	Fit-for-purpose data for models
Scope	Structured systems of record	+ unstructured, per use case
Ownership	Diffuse / IT-held	Named golden source per domain
Lineage	Rarely tracked	Source-to-model, versioned
Validation	Periodic, manual	Layered controls + audit trail
Lifecycle	Retained by default	Controlled decommissioning

AMAT framework, drawing on MAS (2026), Hiniduma et al. (2024), Strata FSC (2026) and Ojewale et al. (2026).

Takeaway. Put the seven controls in place for the first use cases, harden the highest-risk domains first, and make lineage a habit rather than a scramble when an output is questioned. The AI-era data estate has to be fit for purpose continuously, not reconciled periodically.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST04

RECOMMENDED ACTIONS

1. **Apply the seven data-management controls** (fitness, representativeness, quality, classification, security, privacy, lineage) to the datasets behind your first AI use case, each with an owner and a test.
2. **Prepare and validate data readiness per use case** before model use, then monitor the data feeding live models continuously.
3. **Establish a golden source and named owner for each data domain**, wrapping the highest-risk domains, such as the IBOR-ABOR seam, in layered controls.
4. **Record the source, version and owner of each dataset** feeding your first production model, then build end-to-end lineage from source to model input.
5. **Engineer quality at source, not downstream**, with schema-constrained extraction and completeness and citation checks.
6. **Agree a shared schema and taxonomy** for one priority data domain, then build a semantic and metadata layer.
7. **Organize historical data, such as diligence history, now** so it compounds into reusable, governed data products.

SEQUENCE AND DEPENDENCIES

1. The sequencing trap is trying to make the whole estate AI-ready before doing anything.
2. Golden-source determination is an organizational decision as much as a technical one, and it needs a named owner (a chief data officer or data domain steward) empowered to arbitrate.
3. The workable path is to cleanse and ready the priority data that the first use cases actually depend on and deploy in parallel, rather than boiling the ocean.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns golden sources and quality at source; captures lineage, applies noise reduction and outlier handling before modeling, constrains extraction to a schema.
Second Line	Sets risk-based minimum expectations for the data-management control framework and challenges first-line control design and execution.
Third Line	Assures data quality, structure, and lineage; curated, clearly defined data improved audit AI performance in the CIIA case studies.
Board and Executive	Sets the mandate and funding for AI-ready data as a strategic capability and for clear golden-source ownership.
Cross-Functional	Data readiness is use-case-specific and the semantic/metadata layer spans functions; quality, lineage, and ownership require investment, technology, data, risk, and operations to act together.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST04

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Poor data quality caps AI**	Treat quality as a control framework with owners and tests; fix defects at source, not downstream.
HIGH **Mislabeled or unready training data**	Validate data readiness per use case (completeness, outliers, feature relevance, labels) before model use, and monitor it continuously once the model is live.
HIGH **Lineage and provenance opacity**	Record end-to-end lineage source to model, including third-party provenance and dataset versions.
HIGH **Golden-source failures at the IBOR-ABOR seam**	Define golden sources per domain and apply layered controls with audit-trailed overrides to the seam.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Data readiness	Share of priority datasets passing readiness checks (completeness, outliers, labels) before model use.	Model incidents traced to data defects
Quality controls	Share of higher-risk domains under layered controls with audit-trailed overrides	Reconciliation exceptions at the IBOR-ABOR seam
Lineage coverage	Share of AI datasets with recorded provenance and version lineage	Audit findings on undocumented or untraceable data
Schema and metadata	Task-aware schema/annotation quality score	Extraction errors: schema drift, unit or period mistakes

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Opportunity: a small, less-fragmented estate is inherently easier to keep AI-ready.

MID-SIZED

Binding constraint: golden sources and metadata fragment as asset classes and systems multiply, without a full data function to control them.

LARGE AND GLOBAL

Opportunity: scale to build reusable, governed data products and a semantic layer that compounds.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P03-ST04

P03-ST07

Data Governance and Access Controls

For: Technology; Risk and Compliance; Data and Analytics

Governance sets the rules for how data is classified, owned, shared and managed, and access control enforces them. Regulatory attention is shifting from model performance alone toward the governance of the data that trains AI, and a manager that cannot evidence how its AI data is governed is exposed however well the models perform. In a 2025 SimCorp report drawing on a WBR Insights survey of 200 investment-operations leaders, two-thirds named an integrated data management strategy and robust data governance as the solution to their data challenges (a respondent view, not a tested remedy). Consolidating data into a lake does not solve this, because an ungoverned lake is a liability with an AI system attached to it. Governance is the precondition for scaling AI, not a brake on it.

35%
of 323 institutions surveyed by the Dutch AFM had no technical or procedural controls on employees' use of AI tools (first half of 2025, not only asset managers)

75%
of 12 managers active in Italy report formal data governance policies, but only 33% have specific policies for algorithms (self-reported)

WHERE FIRMS GO WRONG

Three failure modes recur and compound. Data governance and data security run as parallel workstreams, a separation that was survivable when applications mediated access but not when an AI system can interpret information across the estate at once. Ownership is informal, so classification and protection lack a clear line of accountability. And governance is treated as documentation rather than enforcement, although documentation alone does not cover unstructured data or the breadth of data AI now touches.

A governance lifecycle for AI data



Microsoft Security, 2026 (CLPM lifecycle)

Takeaway. Treat governance and access control as one capability: name owners, classify and label data, enforce labels through role-based and just-in-time access, and attest periodically. Fix ownership and classification first and run the redesign alongside the AI program, not after it.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST07

RECOMMENDED ACTIONS

1. **Assign named data owners and stewards for each domain** before AI systems process that data, and include AI features already in existing platforms in the first inventory pass.
2. **Apply a simple classification schema** to the domains your first AI use cases touch, embedded in policy.
3. **Enforce classification labels through role-based access**, building to just-in-time access, and test with two differently entitled users before an AI search or agent goes live.
4. **Recertify owners of the data domains in AI scope**, then run periodic attestation as audit evidence.
5. **Map where policy does not cover unstructured data** and personal data in closed-source models, then redesign governance for AI.
6. **Define the firm's data framework in one place**: what data is used, who has access, how, and who oversees it.
7. **Check every AI use case, and every vendor-enabled AI feature before it is switched on, for data-use compatibility**, and confirm third-party licensing before use.

SEQUENCE AND DEPENDENCIES

1. The main lever is organizational rather than technical: fuse the functions that set data rules and the functions that enforce them into one accountable capability.
2. Moving toward domain ownership with self-service access requires a central governance layer that enforces quality and compliance standards across decentralized domains; without that central spine.
3. Where a CDO or data-governance function and a CISO or security function operate as parallel workstreams.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Data owners and stewards sit in the business and are accountable for classification accuracy and access decisions for their domains.
Second Line	Risk and compliance set or challenge the control standards for classification, access and attestation, oversee exceptions.
Third Line	Internal audit tests governance by examining attestation records, access grants and classification accuracy end to end, using the attestation program as its evidence base.
Board and Executive	The board and executive ensure a single accountable owner for data governance, examine reliance on data sources and licensing as an oversight matter.
Cross-Functional	The classification schema connects controls to business decisions across desks and functions, and a central governance layer enforces standards over decentralized domain ownership.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST07

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Weak governance foundations**	Redesign governance for AI end to end and define a single data framework (scope, access, use, oversight); make ownership a named business accountability.
HIGH **Governance-security split**	Operate governance and security as one capability with joint CDO/CISO ownership, so every classification has a matching enforced control.
HIGH **Ungoverned data access**	Enforce classification through RBAC and just-in-time access so labels are gates, not advisory permissions; embed enforcement in the architecture with audit trails.
HIGH **AI retrieval or agent permission bypass**	Require entitlement-matched retrieval, approve each AI feature before use, test restricted indexes and test with differently permissioned users before go-live; give agents controlled identities with prompt revocation (P04-ST07).

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Classification and ownership	Share of AI-in-scope data domains classified with a named owner	Unclassified or unowned data reached by AI
Access enforcement	Share of access grants issued via RBAC/JIT aligned to labels	Excess-privilege and orphaned-access findings; AI retrieval entitlement failures in a fixed monthly test set, by severity and workflow
Attestation	Attestation-cycle completion rate	Stale or invalid grants found at audit; median time from finding an unauthorized AI data grant to confirmed.
Governance-security integration	Joint CDO/CISO review cadence	Incidents traced to a governance-enforcement gap

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

With no dedicated data function, the realistic move is to inherit governance from the platforms already in use: adopt the classification schema and role-based.

MID-SIZED

The exposure is governance debt: data spread across silos with ownership and rules accreted use case by use case.

LARGE AND GLOBAL

The distinctive challenge is scale and reach: cross-business sharing rules, multi-jurisdiction enforcement and dependence on many data sources.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST07

P03-ST08

Data Privacy Controls and Information-Barrier Architecture

For: Risk and Compliance; Legal; Technology

Leakage of material non-public information through a generative AI system can lead to insider trading where the information is used in investment decisions, and failing to maintain adequate procedures to keep MNPI out of AI tools can itself create liability. In a Lowenstein Sandler survey of 107 private fund advisers (fielded late 2025, published 2026), data ownership and privacy was the most cited AI concern (42%) and among the top concerns for alternative data (39%). The controls that matter are a short, sequenced set: classify data first, run sensitive workloads through ringfenced instances that do not retain or train on firm data, enforce information barriers including for agents, and use privacy-enhancing technologies where data must be used without being exposed.

42%

of 107 private fund advisers cite data ownership and privacy as their top AI concern, 39% for alternative data (survey fielded late 2025)

18%

of the same advisers already cite insufficient rights to use alternative data with AI or other software tools

WHERE FIRMS GO WRONG

The most common failure is using public or consumer-grade AI tools for sensitive work, where prompts and uploaded documents may be retained and used for training, risking a confidentiality breach and, for legal material, waiver of privilege. Firms also deploy tools without knowing what they retain, and let data licenses fall out of date.

Public tools vs an approved enterprise or ringfenced deployment

Dimension	Public / consumer AI	Approved enterprise / ringfenced
Data retention	Retained; may train vendor models	No vendor retention or training, by contract or by design; verify prompts, outputs and logs
Confidentiality & privilege	Breach / waiver risk, hard to remedy	Risk reduced; privilege more likely to hold under counsel-directed use, assessed on the facts
Where data lives	Leaves the firm to a third party	Private instance; if vendor-hosted, location and permitted transfers verified
Control posture	None by default	Access controls, isolation, monitoring and audit

AMAT framework, drawing on Ropes & Gray (2025), White & Case (2026), ACA Group (2026) and Buczynski et al. (2025)

Takeaway. The immediate step is the least technical: classify the firm's data and run a retention assessment on every AI tool already in use, so the firm knows where its most sensitive information is actually sitting today. These controls are not optional refinements, because inadequate procedures can themselves create liability.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P03-ST08

RECOMMENDED ACTIONS

1. **Classify the data your first AI use cases will reach** as public, confidential or mission-critical before any LLM is connected, and confirm the legal basis to process each tier.
2. **Route sensitive workloads through a ringfenced or enterprise-grade instance** that contractually prohibits retention, training and disclosure of firm data, and keep AI use for legal work counsel-directed.
3. **Design information barriers and access controls explicitly**, and limit each autonomous agent, with a named owner, to the datasets, tools and network destinations its task requires.
4. **Confirm data licenses permit the specific planned AI use**, check model-training rights where training is planned, and use privacy-enhancing technologies where sensitive data must be used without being centralized.
5. **Run a data-retention assessment for every AI tool** in use or under evaluation, mapping what is stored, for how long, and whether MNPI could reside in it.

SEQUENCE AND DEPENDENCIES

1. Sequencing is the main execution decision, and it runs in the order of the four moves above: classification first, because instance choice.
2. Run-state ownership is set at first use, not added later.
3. Enterprise platforms reduce but do not eliminate unauthorized-access risk, so the choice of a reputable platform must be paired with access controls and monitoring rather than treated as the endpoint.

WHO DOES WHAT: LINES OF DEFENSE

First Line	The named owner of each deployed workflow approves the datasets and users it may reach.
Second Line	Risk, compliance and legal set and challenge the controls: they define the data-classification, MNPI, privilege, retention and licensing rules, review the data-retention and MNPI assessments.
Third Line	Audit tests whether classification, retention assessments, instance rules and agent isolation work in practice, including access barriers, sampled outputs and the response to vendor changes.
Board and Executive	Senior management sees that adequate MNPI and confidentiality procedures cover AI use, since a failure to maintain them can itself create liability, receives material incidents.
Cross-Functional	Information-barrier design spans business lines and is inherently cross-functional: PGIM's standards, for example.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST08

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Agent data leakage via tool use**	Isolate agents: a named owner, least-privilege access to datasets and tools, restricted network destinations and cross-barrier testing before go-live.
HIGH **MNPI leakage via GenAI**	Keep MNPI out of public and unapproved AI tools and document the MNPI control response before deployment.
HIGH **Privilege and confidentiality waiver**	Use enterprise/ringfenced instances; keep legal use counsel-directed.
HIGH **Inadequate MNPI procedures**	Run a data-retention assessment and document controls for every AI tool.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Data classification	Share of datasets classified before LLM connection	AI workloads running on unclassified data
Tool retention assessment	Share of AI tools with a completed retention/MNPI assessment	Sensitive data found residing in AI tools; unapproved retention findings or vendor-feature changes detected
MNPI and confidentiality	Pre-deployment MNPI control response documented; share of sampled AI outputs, retrieval results and tool calls passing permission tests	MNPI or confidential data found in public or unapproved AI tools; confirmed restricted-data disclosures or near misses
Instance coverage	Share of sensitive workloads on ringfenced/enterprise instances	Confidential work run on public/consumer tools

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

These firms depend on vendor platforms and have little leverage over contract terms, so the practical control is to be a careful buyer.

MID-SIZED

The move is to formalize: stand up data classification and run per-tool retention assessments so that AI use is governed by tier rather than by habit.

LARGE AND GLOBAL

The exposure is structural, because multi-strategy firms run many business lines whose data agents can reach across, and the feasible response, in the guide's reading.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P03-ST08

P03-ST09

Alternative Data and External Research Integration

For: Investment Teams; Technology

In a Lowenstein Sandler survey fielded in late 2025 and published in 2026, 81% of private fund managers increased their use of alternative data across 2024 and the first half of 2025, and only 5% cut back. The strategic problem is that the advantage can decay: in a study of 97 published stock-return predictors, returns were about 26% lower out of sample and about 58% lower after publication (an analogy, not a decay rate for any feed). Differentiation is more likely to come from proprietary and hard-to-replicate data, multimodal signals and data a firm generates itself than from the same feeds competitors buy. Much of the research process also rests on third-party data the firm does not own or control, which is a lineage, licensing and model-risk exposure as well as a sourcing one.

81%

of private fund managers increased alternative-data use across 2024 and the first half of 2025; only 5% cut back (late-2025 survey)

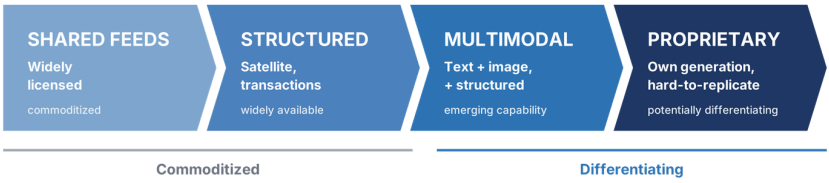
58%

lower returns after publication for 97 studied stock-return predictors, and about 26% lower out of sample (academic predictors, not alternative-data feeds)

WHERE FIRMS GO WRONG

Firms spend on alternative data as if buying a novel feed still bought an edge, but a widely licensed dataset is closer to a commodity, and a firm that keeps buying the same feeds as everyone else risks paying for parity rather than advantage. Part of the research process also rests on third-party data the firm does not control, since most private fund managers running custom AI train it on third-party proprietary data. Alternative data can also carry personal data, MNPI or restricted content, so buying through a vendor does not by itself give a lawful right to use it. The hard part has moved from finding data to judging it.

Where the alternative-data edge may sit



AMAT framework, drawing on Cao et al. (2025), Katona et al. (2025), BlackRock (2025) and T. Rowe Price (2026)

Takeaway. Spend differently rather than more: move from commoditized shared feeds toward proprietary, multimodal and self-generated data, with sourcing, provenance and validation discipline travelling with the data. Start by testing which of today's feeds still earn their cost.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P03-ST09

RECOMMENDED ACTIONS

1. **Assess licensing, usage rights and auditability before onboarding any source**, starting by applying a written vendor-evaluation and provenance process to the next source.
2. **Name the one or two areas where the firm has a genuine data edge** and the most commoditized feeds it pays for, then weight sourcing toward proprietary data, trialed against a baseline before scaling.
3. **Prioritize alternative data where a trial shows a signal a competitor cannot easily see**, not data competitors already hold.
4. **Start early-warning monitoring for one fast-moving signal**, then build streaming ingestion rather than relying on periodic updates.
5. **Govern AI-derived (synthetic) data as a first-class governed source**, recording when data is AI-generated, its inputs and lineage.
6. **Validate third-party data through documented lineage, sample-based source checks and human validation** of material AI-derived conclusions before they reach a portfolio manager, model or regulatory report.
7. **Test net value for the largest feeds at each renewal**, then build a net-value test for every source, with a named owner who authorizes retiring or renegotiating feeds that repeatedly fail.

SEQUENCE AND DEPENDENCIES

1. The main practical dependency is that alternative data is only as useful as the pipeline that ingests and validates it.
2. The second consideration is build-versus-buy at the data layer.
3. Sources arrive with unique schemas, delivery mechanisms and quality characteristics, so a firm needs the document-to-signal and multimodal extraction capability (P03-ST02) and the lineage and quality controls (P03-ST04) in place.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Investment and data-science teams source, evaluate and integrate alternative data, decide where it genuinely differentiates, validate extracted signals before they inform a decision.
Second Line	Risk, compliance and legal own licensing, lawful-sourcing and data-protection review of alternative data, and challenge third-party dependency; data-protection obligations are P05-ST07 and model risk is P05-ST04.
Third Line	Internal audit tests provenance, licensing and lineage of alternative-data sources end to end, from vendor contract to the signal used in a decision.
Board and Executive	The board and executive weigh alternative-data spend against demonstrable edge, given a market that, on Neudata's figures as reported by Lowenstein Sandler.
Cross-Functional	A data catalog and shared sourcing, provenance and validation process serve all desks rather than each team onboarding vendors independently.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST09

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Advantage decay**	Shift spend toward differentiating data, trialing each new source against the existing baseline before scaling, and test each feed's net value at renewal.
HIGH **Unlawful or restricted sourcing**	Before use, name a business owner; document origin, permitted purposes and rights for research, AI processing and onward sharing; require Compliance and Legal sign-off for heightened-risk feeds.
HIGH **Synthetic-data provenance and licensing**	Record AI-generated data and its lineage (source inputs, transformation, reuse restrictions, quality checks, owner); classification and access design in P03-ST07.
MEDIUM **Third-party data dependency**	Validate third-party data via documented lineage, sample-based source checks and human validation of material conclusions before use; formal vendor due diligence is P07-ST04, model risk is P05-ST04.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Signal value over time	Change in out-of-sample predictive contribution or decision coverage against an agreed baseline; refresh timeliness and error trend	Documented incremental investment or research value, net of licensed-data and integration costs, with uncertainty noted
Coverage and timeliness	Coverage/timeliness gained in fast-moving or thin-data segments	Decisions informed by signals not otherwise available in time
Sourcing-spend mix	Share of budget on commoditized vs proprietary/structured sources	Cost per differentiating source
Vendor AI-readiness	Share of vendors delivering AI-enabled / AI-ready data	Integration effort per new source

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Commoditization is an advantage here: the same alternative data that no longer differentiates large firms is now cheap and often AI-enabled by the vendor.

MID-SIZED

Capacity squeeze

The capacity squeeze is real: too large to compete on cost alone, not yet resourced for proprietary generation at scale.

LARGE AND GLOBAL

Scale is the advantage, and in the guide's reading the edge is likeliest to come from proprietary data.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST09

P03-ST10

Data Marketplaces and Vendor Data Strategy

For: Senior Leadership; Technology

Buying data is already the norm: in a 2026 Mercer survey of asset managers, 58% were using some vendor-provided data. Vendors are now embedding AI directly into their products, with 89% of private fund managers in a Lowenstein Sandler survey (late 2025, published 2026) reporting that vendors are fully or mostly doing so, and more than four in five have seen the cost of AI-enabled data products rise. Licensing terms are tightening in step, particularly on model training. Buyers can find it hard to establish what a dataset is worth before they have paid for it, while their own data sits locked in legacy systems. The discipline is to value a dataset before buying it, make the firm's own data visible before deciding build or buy, building an internal marketplace only where it is worth it, and treat licensing as a live exposure.

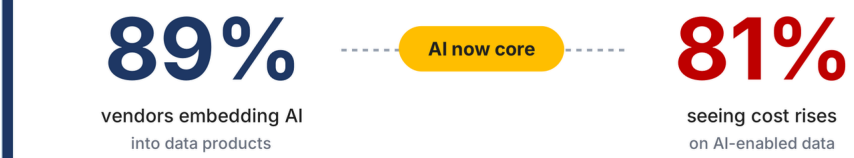
89%
of private fund managers report vendors fully or mostly embedding AI in their data products, with over four in five seeing costs rise (Lowenstein Sandler survey, late 2025)

58%
of asset managers use some vendor-provided data (2026 Mercer survey)

WHERE FIRMS GO WRONG

Buyers can find it hard to establish a dataset's value before paying. It depends on the context in which it is used, so a buyer may be unable to judge it without access while the provider is reluctant to reveal detail, and procurement defaults to marketing and trial and error. The firm's own data is often locked in legacy systems, so the build option is invisible, and licensing is treated as a settled contract when vendors are imposing tighter AI-use and training restrictions.

The vendor-data shift



Private fund managers surveyed November to December 2025 (107 respondents). Source: Lowenstein Sandler (2026)

Takeaway. The immediate step is the least glamorous: catalog what you already hold, and test the next material data purchase against a defined business need, a permitted trial and a measurable baseline. Treated with the discipline of any core system, the marketplace turns reactive buying into an evidence-based capability.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST10

RECOMMENDED ACTIONS

1. Start by cataloging the critical datasets, end users and use cases behind the next material data purchase; build toward an explicit build-versus-buy decision for each material data asset, choosing whether to buy externally, productize internally or, only where ownership and redistribution rights allow, monetize externally.
2. Start by testing the next material dataset against a named use case, using a permitted trial, a baseline and documented acceptance criteria; build toward repeatable value measurement, using machine-learning and explainable-AI methods where the data, tooling and expertise justify them.
3. Start by comparing the cost of the largest data renewal with the outcome it supports; build toward repeatable decision criteria by data category for every material purchase, estimating the performance achievable within a set budget, or the budget needed to reach a target level, where the data feeds a measurable model or signal.
4. Confirm your data licenses permit the AI use you depend on before deploying AI or training models on licensed data, and record vendor-enabled AI features in the AI inventory before use.
5. Design marketplace governance before the platform. Agree definitions, metadata standards, security protocols, quality benchmarks and entry criteria, then choose the platform via buy, build or partner with access and lineage controls built in.

SEQUENCE AND DEPENDENCIES

1. The governance layer, the agreed definitions, metadata standards, security protocols and quality benchmarks, must be designed before the platform is developed.
2. The value-assessment method carries its own trade-off.
3. On frameworks, the structured methodologies available for this sub-topic are the four-step marketplace build and the value-quantification and selection mechanism described above, which is still a research prototype.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Data and technology teams own the evaluation, value-quantification and procurement of external data and the build of any internal marketplace, including its access and lineage controls.
Second Line	Legal and compliance review data licensing for AI-use rights before deployment, and risk governs the data-sovereignty and IP exposure created when a model or API.
Cross-Functional	Marketplace governance, the incentive structures, entry criteria and quality benchmarks that give both providers and consumers confidence, is inherently cross-functional.
Board and Executive	Ownership of the build-versus-buy investment decision and of the data-spend trajectory sits at the executive level, given the material and rising cost of AI-enabled data.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST10

KEY RISKS AND CONTROLS

RISK	CONTROL
MEDIUM **Data-vendor dependency and concentration**	Maintain a register of critical vendor data, models and infrastructure (owners, contract end dates, downstream dependencies and the due-diligence record).
MEDIUM **Information asymmetry**	Test value before purchase against a defined use case and baseline, using ML and XAI methods where justified.
MEDIUM **Vendor cost escalation and restrictive licensing**	Value-based procurement plus a purchase-time licensing review of AI-use rights.

METRICS SCORECARD

AREA	LEADING	LAGGING
Value-before-purchase	Share of trials run with predefined acceptance criteria and a baseline	Incremental model or workflow result realized after purchase, net of dataset cost
Procurement sizing	Share of material purchases and renewals with a documented, use-case-specific value comparison	Spend avoided or reallocated, or analytical value retained, after those comparisons
Licensing coverage	Share of licensed datasets with confirmed AI-use rights	Licensing-related compliance incidents on AI use
Vendor cost governance	Cost trend of AI-enabled data products versus value delivered	Proportion of spend where cost growth outpaces value
Marketplace efficiency (if built)	Median time to discover, approve and provision a reusable data product	Hours of data preparation avoided and duplicate license spend eliminated



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P03-ST10

Pillar 4 Executive Summary

For: Senior Leadership; Technology · Also relevant to: Investment, Risk, Compliance, Operations, Finance

THE ARGUMENT

AI in asset management runs on a technology stack whose architecture decides whether AI scales safely or stalls expensively. Pillar 4 covers those engineering choices through an asset-management lens: the target architecture, the model layer, compute, integration, orchestration and retrieval, security, and the cross-cutting concerns of observability, controlled deployment, interoperability and cost. In this guide's view, AI scales where a firm makes deliberate, governed architecture choices and then instruments, contains and costs what it builds, more than where it runs the largest model. Strategy and use cases live elsewhere in the guide; Pillar 4 is how they are engineered to work in production.

92%
of Dutch asset managers in a 2026 KPMG survey call a mature architecture necessary

44%
of UK financial firms' third-party AI models came from the top three providers (Bank of England survey, 2024)

30x
growth in cost per interaction, 2023 to 2026, in one illustrative EY customer-service example

WHY NOW

Architecture and controlled deployment now draw supervisory attention: the UK FCA runs an AI sandbox and a live-testing scheme, the EU AI Act provides for member-state regulatory sandboxes, Singapore's MAS has issued final AI risk-management guidelines for all financial institutions, which it is phasing in, and the Financial Stability Board flags AI provider concentration as a systemic vulnerability. The supply side has concentrated: the top three providers supplied around 44% of UK financial firms' third-party AI models in 2024, up from roughly 18% in 2022, per a Bank of England survey. Gartner forecasts in 2026 that enterprises' overall inference costs will rise even as unit prices fall, because agentic workloads use far more tokens.

RISKS TO AVOID

Point-tool sprawl and legacy debt: accumulating tools instead of an architecture, so nothing scales and complexity taxes every change. **Unwatched, uncontained AI in production**: models that drift and fail silently, and agents deployed with no observability, no safe action space and no controlled rollout. **Concentration and cost blind spots**: single-vendor dependence with no tested exit, and a cost base that is expected to rise with usage while the firm watches the falling token price.

Bottom line. Pillar 4 is the engineering blueprint: AI scales safely or stalls expensively on the strength of these architecture choices. In this guide's view, what matters is less the model a firm runs than whether it designs a governed architecture, matches models and compute to the work, builds the connective layers once, and then instruments, contains and costs what it deploys. Make deliberate choices now, and do not ship material AI changes straight to production; the alternative is a sprawl of point tools that is expensive to run, hard to leave and hard to trust.

Explore Pillar 4 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P04-summary

PRIORITY MOVES FOR LEADERSHIP

1. **Design the target architecture deliberately**: put minimum controls first, clear legacy debt, choose an operating model, and use a governed agentic mesh as a reference point instead of accumulating point tools.

2. **Match models and compute to the work**: route across a governed model portfolio, right-size models, and place each workload across cloud and on-premises on its merits.

3. **Build the connective layers once and governed**: a governed integration layer and an API-first orchestration foundation the AI can safely act through, with RAG grounded in a governed corpus.

4. **Treat every agent as a non-human identity**: least-privilege access, isolation and a bounded safe action space, with security designed in at the architecture level.

5. **Instrument and contain before scaling**: an audit trail and continuous drift and quality monitoring for silent failure, and controlled deployment through sandbox, canary and phased rollout with a way back.

6. **Keep the exit open and the cost visible**: map provider concentration, design for portability with a rehearsed exit, and cost the full total cost of AI up front.

IN THIS PILLAR

P04-ST01	Target AI Architecture for Asset Managers	92
P04-ST02	Model Layer: LLMs, Specialized Models and Model Routing	95
P04-ST03	Cloud, Compute and Scalability for AI in Asset Management	98
P04-ST04	Integrating AI with Core Systems: PMS, OMS, EMS, CRM, Data Warehouses and Document Platforms	101
P04-ST05	APIs, Microservices and Workflow Orchestration: Building Composable, Orchestrated AI	104
P04-ST06	Retrieval-Augmented Generation and Enterprise Knowledge Bases	107
P04-ST07	Security Architecture and Identity/Access Management for AI	110
P04-ST08	Build, Buy and Partner Strategy for AI Capability	113
P04-ST09	Model Observability and Performance Monitoring: Seeing What AI Does in Production	116
P04-ST10	Sandboxes and Controlled Deployment: Experimenting with AI Without Betting the Firm	119
P04-ST11	Vendor Interoperability and Avoiding Lock-In: Keeping the Exit Open	122
P04-ST12	Cost Management Across Models, Compute and Data: What AI Really Costs at Scale	125

Explore Pillar 4 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P04-summary

P04-ST01

Target AI Architecture for Asset Managers

For: Technology; Senior Leadership

A target AI architecture is a small set of durable design choices rather than a shopping list of tools, and those choices decide whether every later technology decision compounds into one coherent system or fragments further. QuantumBlack proposed the agentic AI mesh in 2025, a vendor-neutral orchestration layer that coordinates specialized agents in place of point tools; it is a well-argued design direction rather than an industry standard. In a January 2026 survey of Dutch asset managers, 92% called a mature architecture necessary for their target operating model, but agreeing is not the same as having solved it. The obstacle is often less a missing tool than decades of legacy debt that AI cannot break on its own.

92%
of surveyed Dutch asset managers call a mature architecture necessary (self-reported)

75%
report moderate-to-very-high AI technology challenges

WHERE FIRMS GO WRONG

The usual failure is not a missing capability but legacy debt left unresolved. Firms that modernize often fail to decommission what they replaced, so licensing, maintenance and integration costs continue and erode the gains the new platform was meant to deliver. Accountability for platform delivery is also diffuse, and many firms default to a centralized-versus-decentralized IT debate instead of asking how strongly the enterprise drives modularity and reuse.

Point-tool infrastructure versus an illustrative agentic-mesh architecture

Dimension	Point tools	Agentic mesh
Structure	Disconnected AI point tools	Composable orchestration layer
Control	Governance decided per tool	Governance built into the mesh
Result	Fragmentation that compounds	Capability that compounds

AMAT framework, drawing on QuantumBlack / McKinsey (2025) and BlackRock (2025 Annual Report).

Takeaway. Treat architecture as a small number of durable choices rather than the sum of the tools bought: clear legacy debt, choose an operating model deliberately, and build a governed mesh instead of adding point tools. Fix legacy decommissioning and the operating-model choice first, because every later decision inherits from them.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST01

RECOMMENDED ACTIONS

1. **Make legacy decommissioning a required deliverable** of every modernization program, not an optional follow-on.
2. **Choose the IT operating model deliberately**, based on the innovation velocity needed (Adaptive Innovator for most asset managers, in this guide's view).
3. **Assess every new AI investment against the agentic-mesh principles**, building toward a governed, vendor-neutral orchestration layer.
4. **Give named delivery teams end-to-end accountability** for platform capabilities, and an owner able to stop each live AI workflow.
5. **Test each material AI workload** on accuracy, latency, confidentiality and cost, building toward model routing.
6. **Load-test document- and data-intensive AI pipelines** at target volume before production.
7. **Make registry, observability and evaluation coverage a go-live condition** for every agent, including vendor AI features.
8. **Right-size the architecture** to the firm's scale: managed services for boutiques, consolidation and legacy retirement at scale.

SEQUENCE AND DEPENDENCIES

1. Sequencing determines whether the architecture compounds or accumulates more debt.
2. The operating-model choice also has a practical implication for how centralized versus federated the build should be.
3. One recurring execution error is scaling AI pilots ahead of the enterprise infrastructure that is meant to carry them: without a robust foundation, McKinsey observes in financial services.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Delivery and business teams build on reusable enterprise platform components rather than one-off point tools, own each live AI workflow end to end.
Second Line	Risk and compliance set proportionate requirements and independently challenge access, monitoring, change and escalation arrangements.
Third Line	Internal audit tests whether decommissioning milestones were actually met, not only whether a modernization program delivered its new capability.
Board and Executive	The board and executive own the IT operating-model choice as a strategic decision, in this guide's view weighted toward Adaptive Innovator for most asset managers.
Cross-Functional	The agentic mesh includes compliance and risk-management capability as a core component.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST01

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Legacy maintenance vicious cycle**	Fund legacy exit as a standing budget line, not a one-off project, and treat decommissioning as a required deliverable of every modernization program.
HIGH **Failure to decommission legacy systems**	Write decommissioning into the program's formal deliverables and milestones, with program close-out gated on legacy retirement or an approved, time-limited coexistence plan, not only on new-system go-live.
HIGH **Data quality and bias**	Extend data governance to the data the firm controls, input, retrieval and fine-tuning data, assessing quality and provenance on an ongoing basis rather than only at deployment.
HIGH **Uncontrolled agent access and provider concentration**	Grant agents least-privilege tool and data permissions with owner approval, log material tool actions, test the stop and fallback path.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Live workflow quality	Error and exception rate on AI-assisted workflows, against the pre-AI baseline	Client- or business-impacting errors traced to AI-assisted work
Speed and cost to change	Time to deliver or modify a governed AI capability	Net run cost per completed, quality-accepted task
Incidents and resilience	Fallback and load tests passed at target volume before go-live	Incidents with business or client impact per 1,000 AI-assisted tasks, with severity and recovery time
Human review	Share of AI outputs reviewers edit, override or reject, trended (where the platform records reviews)	A sustained override rate at or near zero, investigated as a sign of rubber-stamping

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

With no dedicated architecture function, the realistic move, in this guide's view, is the minimum-viable architecture a 2026 preprint proposes for small US advisers.

MID-SIZED

Capacity squeeze

The exposure is a capacity squeeze: enough scale to have accumulated real technical debt, but often not yet a dedicated architecture function to resolve it deliberately.

LARGE AND GLOBAL

In this guide's reading the legacy complexity tax is likely to be heaviest here, built up over decades of enterprise systems.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST01

AMAT

P04-ST02

Model Layer: LLMs, Specialized Models and Model Routing

For: Technology

The model layer is a portfolio of engineering choices rather than a decision about which large language model to buy: whether an LLM is the right tool at all, whether to use a general-purpose or domain-adapted model, how to fine-tune it, and how to route work across several models. The starting discipline is a necessity test, since simpler and more transparent methods often beat an LLM on cost, interpretability and auditability for routine, deterministic tasks. General-purpose models used off the shelf frequently struggle with complex financial reasoning, and domain-adapted models can beat them on specific financial tasks. Singapore's regulator, MAS, has consulted on guidelines that would require firms to document why a model was chosen over simpler alternatives, which makes selection a decision the firm has to be able to defend.

60%

forecasting accuracy for a BlackRock domain-trained earnings-call model, against around 50% for general-purpose GPT models (one 2025 test)

90%

Gartner's 2026 prediction of how much cheaper trillion-parameter inference becomes for providers by 2030

WHERE FIRMS GO WRONG

The most common failure is defaulting to a general-purpose frontier model for everything and skipping the question of whether it is the right tool, which carries output-quality risk in the workflows where accuracy matters most. Firms with no explicit multi-model strategy make a default lock-in decision anyway, and without separating the orchestration layer from the model API, switching later becomes prohibitively expensive. Many also under-invest in evaluation, selecting on public leaderboards that measure general reasoning rather than financial-domain performance.

One BlackRock test: 40-day post-earnings return predictions, 500 Q1 2024 calls

~60% — ~10pt gap in this test — ~50%

accuracy: BlackRock proprietary earnings-call model

general-purpose GPT models, same 500-call test

Source: BlackRock Systematic, How AI Is Transforming Investing, 2025 (self-reported; Q1 2024, 500 calls).

Takeaway. A well-run model layer depends less on running the largest model than on asking whether an LLM is needed at all, adapting to the firm's domain, routing work across a governed portfolio and documenting every material choice. Start with the necessity test on the next model decision, and build the evaluation pipeline that lets the firm prove the choice was right.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST02

RECOMMENDED ACTIONS

1. **Apply a necessity test before every LLM adoption**, preferring simpler, more transparent methods for routine, deterministic tasks.
2. **Test a domain-adapted model against the general-purpose default** on the one task where accuracy matters most.
3. **Ringfence any model that handles sensitive data**, starting with one open-source model adapted through parameter-efficient fine-tuning.
4. **Put a model-agnostic abstraction layer in front of the current provider**, then build toward routing routine work to small or domain models.
5. **Build a financial-domain test set** for the highest-volume task and select models task by task on it, not on public benchmarks.
6. **Benchmark every fine-tuned model against the prior version** and re-verify alignment before acceptance and on each material upgrade.
7. **Document every material model choice**, including why it was selected over simpler alternatives.

SEQUENCE AND DEPENDENCIES

1. The dependency most firms underweight is evaluation infrastructure.
2. A second consideration is the accuracy-explainability trade-off, which is a design decision, not a technical accident.
3. Model routing, task-by-task selection and fine-tune validation all assume the firm can measure model performance on its own tasks.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Model selection, fine-tuning and prompt engineering are first-line build activities under this sub-topic's ownership of lifecycle Stage 2; first-line teams run the necessity test.
Second Line	Risk and compliance set model-risk policy, challenge selection assumptions and require alignment re-verification.
Third Line	Internal audit tests whether model choices passed the necessity and documentation gates, and whether fine-tuned models were benchmarked and alignment-re-checked before acceptance rather than after.
Board and Executive	The board and senior management own accountability for models the firm cannot explain or has not validated, personal where a regime such as SM&CR applies.
Cross-Functional	Bias and fairness testing (P07-ST09), explainability obligations (P05-ST09) and production observability (P04-ST09) are cross-cutting and are referenced here, not absorbed into the model layer.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST02

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Version-upgrade degradation**	Benchmark each model on the firm's own tasks before upgrading; do not assume a newer model is better.
HIGH **Unjustified model complexity**	Require model and feature selection to be justified and documented, in proportion to the use case's risk, with a bias toward simpler models unless complexity is genuinely warranted.
HIGH **Generative-AI risk amplification**	Apply GenAI-specific guardrails (input/output screening, tenant isolation, provider controls) proportionate to use-case sensitivity.
MEDIUM **Fine-tuning breaks alignment**	Re-verify alignment and benchmark the retrained model against the prior version before acceptance.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Decision usefulness	Share of material tasks with a representative test set and a pass threshold	Accepted, decision-useful answers per 100 representative tasks, against the previous model or manual baseline
Cost and effort	Baseline recorded for staff minutes and cost per task	Staff minutes and total compute or vendor cost per accepted task, trended
Output quality	Output-verification coverage on client-facing generation	Incidents, corrections and rework attributable to model output
Continuity	Documented fallback for each capability rated critical, where the firm controls the interface	Successful restoration during a model or provider unavailability test or event

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The right archetype is taker: use hosted, closed-source models for most work, test one task against the hosted baseline.

MID-SIZED

This tier carries the least distinct sourced evidence and is where model sprawl appears first, the capacity squeeze in practice: enough internal capability to adopt several models.

LARGE AND GLOBAL

Scale supports a governed model portfolio with routing and an in-house evaluation pipeline, and task-by-task selection against a financial-domain benchmark becomes realistic (as one manager describes it in an OpenAI.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST02

AMAT

P04-ST03

Cloud, Compute and Scalability for AI in Asset Management

For: Technology

AI runs on compute, and compute is where an asset manager's ambition meets a hard physical and financial constraint. Serving models at production volume needs elastic, GPU-backed capacity that most firms cannot provision on fixed on-premises hardware, which is why the industry has moved to the cloud. The decisive choice is a deliberate hybrid, matching each workload to the right tier and sizing for asset management's spiky load, such as the month-end performance run and the large risk simulation. The same shift creates concentration: three hyperscalers serve most of the sector, and that dependency belongs in the architecture decision and the firm's outsourcing framework.

62%

of surveyed EU securities-market firms run AI infrastructure only on commercial cloud (2026 regulator survey)

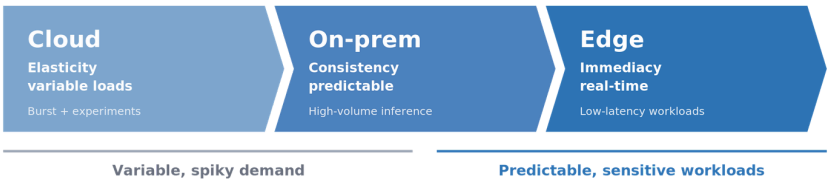
41%

depend on a single cloud provider (same survey)

WHERE FIRMS GO WRONG

The most common failure is treating compute as a fixed-capacity or lift-and-shift problem rather than a set of workloads that each belong on a different tier. Lifting systems into the cloud without re-architecting for elasticity delivers the cost of the cloud without its main benefit, while modern AI on aging on-premises infrastructure runs into latency, automation and legacy-compatibility limits. Defaulting every workload to public cloud ignores data-localization rules and, at high predictable volumes, economics that can invert.

Match each workload to the right compute tier



AMAT framework, drawing on Deloitte, Tech Trends 2026 (three-tier hybrid compute).

Takeaway. Stop treating compute as a fixed-capacity or lift-and-shift problem and treat it as a set of workloads, each placed deliberately on the tier whose economics fit it, with auto-scaling sized for spiky load. The same move creates a dependency on a few hyperscalers, which needs governing as an outsourcing and concentration risk.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST03

RECOMMENDED ACTIONS

1. **Move one spiky workload**, such as the month-end performance run, onto auto-scaling cloud capacity.
2. **Classify AI workloads** by demand pattern, cost and latency, and document a three-tier placement rule: cloud, on-premises or edge.
3. **Evaluate public cloud, private cloud and on-premises** against data-locality and regulatory constraints before placing regulated or sensitive data.
4. **Bring the cloud provider inside the outsourcing framework**, with an exit path and tested failover.
5. **Confirm the infrastructure is fit for purpose**, including GPU and accelerator capacity.
6. **Set the cloud-to-on-premises switching trigger deliberately**, for example when cloud cost reaches roughly 60-70% of the owned-equivalent cost.
7. **Invest in a modern, scalable cloud foundation first**, sequencing the platform ahead of the use case.

SEQUENCE AND DEPENDENCIES

1. Nothing in this article requires a platform migration before the first use case: workload tiering and capacity planning can begin on the firm's current platforms.
2. Four adjacent areas belong to neighboring articles and are referenced, not solved, here.
3. Where the manager hosts AI systems itself, test the existing compute against the use case's measured demand, including accelerator needs where applicable, and modernize where the test shows a shortfall.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Technology and infrastructure teams build and run the compute foundation: they manage vendor-hosted AI services against agreed performance, availability and data-handling standards.
Second Line	Risk and compliance own the cloud-outsourcing and operational-resilience controls: provider concentration and exit planning, data-residency compliance for regulated workloads, and resilience against a provider outage.
Third Line	Internal audit tests whether provider concentration is bounded and an exit path is real, whether resilience recovery objectives are met in outage testing.
Board and Executive	The board owns the cloud and data-center strategy and the concentration and outsourcing risk appetite.
Cross-Functional	Model selection (P04-ST02), security of the compute (P04-ST07), cost optimization (P04-ST12), vendor interoperability and lock-in (P04-ST11).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST03

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Hyperscaler concentration and third-party outage**	Govern the provider as a material outsourcing arrangement; run an outage exercise on the most important AI-supported process.
HIGH **Compute mis-provisioned for spiky demand**	Auto-scale compute up for spikes and back down afterward; tier workloads so burst load flexes in the cloud and steady load runs where it is cheapest.
MEDIUM **Existing compute falls short of a use case's production needs**	Test existing compute against each use case's measured demand before scaling, and modernize the compute layer as those requirements justify it; start from the use case.
MEDIUM **Unlawful data placement or transfer**	Assess permitted locations, provider controls and transfer arrangements before placing data; keep sensitive data on owned, dedicated or private-cloud infrastructure where that assessment requires it.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Workflow service outcome	Share of requests to each material AI-supported workflow that meet its agreed response-time and completion standard.	Business tasks not completed because of a provider or capacity disruption
Compute elasticity	Share of AI workloads on auto-scaling infrastructure that flexes with demand	Idle peak capacity paid for, or spikes (month-end, risk sims) that could not be served
Provider concentration	Share of AI workloads dependent on a single cloud provider; multi-region coverage	Business impact of a single-provider outage in resilience testing
Workload tiering	Share of workloads deliberately placed by tier (cloud / on-prem / edge) against a documented rule	Workloads defaulted to one environment regardless of profile

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The smallest firms will usually do best to buy their way out of the infrastructure problem.

MID-SIZED

On the same reading, this tier runs a real hybrid, balancing elastic cloud against cost and data-residency as private-infrastructure use rises with firm size.

LARGE AND GLOBAL

Scale makes private and dedicated infrastructure both feasible and common, used by up to around 64% of large enterprises in the same survey.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST03

AMAT

P04-ST04

Integrating AI with Core Systems: PMS, OMS, EMS, CRM, Data Warehouses and Document Platforms

For: Technology; Operations

An AI system is only as useful as its connection to the firm's core systems: a model that cannot read a live position, write to the order book, pull a client record from the CRM or retrieve a document is a demonstration, not a capability. The estate is fragmented by design, with many specialist systems, overlapping order and execution management definitions and, in many buy-side firms, several overlapping order-management systems. The decisive work is integration architecture: the choice between point-to-point links and a shared integration layer, and the real-time-versus-batch decision that sits underneath front-office AI.

82%
of surveyed firms use third-party systems for trade execution (68 senior leaders, 2025, self-reported)

79%
do so for portfolio management (same survey)

WHERE FIRMS GO WRONG

The most common failure is treating integration as an afterthought, bolted on once a model already works. A model is proven in isolation, and only then does the firm find that wiring it into order-management, CRM and accounting systems is most of the work, none of it scoped. The systems also hold incompatible data models, so the value an AI workflow promises leaks away at manual handoffs.

From point-to-point sprawl to an integration layer

Dimension	Fragmented point-to-point estate	Integration layer / operating platform
Connections	Bespoke link per system pair; grows n-squared	Publish once to a hub / message bus; read by approved subscribers
Data model	Incompatible per system; identifiers clash	Common, governed model across the estate
Change cost	Rework at every system change	Add once at the layer; downstream unaffected
AI readiness	Agents stall in manual handoffs	Agents reach core systems through governed contracts

AMAT framework, drawing on Strata FSC (2026), Citisoft (2025) and Broadridge (2025) (message bus; interoperability and integration challenges; systems fragmentation and incompatible data models).

Takeaway. Integration is the layer where AI value is realized or lost. Replace point-to-point links with a governed integration layer, tier real-time against batch deliberately, and integrate legacy, CRM and document systems rather than waiting to replace them. Design integration from the outset, map before connecting, and fix a process before automating it.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST04

RECOMMENDED ACTIONS

1. **Map the core-systems estate** and the data flowing between systems before integrating any AI, recording whether each flow needs real time or batch.
2. **Publish core events once to a shared integration layer**, starting with one stream such as trade executions, instead of building point-to-point links.
3. **Reconcile any real-time book against the official record continuously**: a real-time IBOR for intraday decisions, with the batch ABOR as the official record.
4. **Surface AI recommendations inside the CRM** and document systems staff already use, so they arrive where the work happens.
5. **Integrate legacy GRC and core systems with AI** where they carry the most value, rather than replacing them wholesale.
6. **Redesign a broken process before wiring AI to it**, and confirm the systems it touches expose modern APIs and real-time state.

SEQUENCE AND DEPENDENCIES

1. The dependency that gates everything else is readiness of the data and the interfaces.
2. Three adjacent areas belong to neighboring articles and are referenced, not solved, here.
3. An integration layer, real-time reconciliation and agent access assume the core systems can expose the state each use case needs through suitable interfaces; where they cannot.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Technology and operations build and run the integration: they map the estate, build the message bus and reconciliation engine, expose core-system state through modern interfaces.
Second Line	Risk and compliance set the control requirements and independently challenge whether first-line integration owners evidence permissions, lineage, reconciliation breaks, incidents and operational resilience.
Third Line	Internal audit tests whether the integration map is complete, whether systems genuinely connect through the governed layer rather than shadow point-to-point links.
Board and Executive	The board owns the operating-platform strategy and the legacy-modernization investment, and should see integration debt and coexistence plans, not only new AI features.
Cross-Functional	API, microservice and orchestration mechanics (P04-ST05), the target architecture (P04-ST01), data integration and quality as a discipline (P03-ST02, P03-ST04).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST04

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Integration complexity blocks scale**	Build a shared integration layer and map the estate before integrating; publish core events once to a bus where the use case fits rather than wiring bespoke links.
HIGH **Shared-layer outage, data duplication or stale-event propagation**	Monitor freshness and processing, control replay and idempotency, and test fallback and recovery.
HIGH **Batch latency creates intraday blind spots**	Run a real-time IBOR for decisions and keep the batch ABOR as the official record, reconciled bidirectionally and continuously on the data available.
HIGH **Aging, poorly integrated infrastructure**	Integrate the legacy stack with AI rather than waiting to replace it, and design the integration lifecycle from the outset of any new platform.

5 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Integration coverage	Share of core systems (PMS/OMS/EMS/CRM/warehouse/doc) reachable through the shared integration layer	New point-to-point links created outside the layer
Real-time data availability	Share of front-office AI decisions served by real-time (IBOR-grade) data	Intraday decisions delayed or taken on stale end-of-day data
Reconciliation integrity	Share of ABOR deliveries fully reconciled within the agreed window; systematic-break auto-match rate; intraday data freshness	Break backlog and time-to-resolve at the IBOR/ABOR seam
Handoff automation	Share of cross-system AI workflows running without manual export or re-keying	Manual handoffs and re-key steps found in production workflows

4 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The advantage available to the smallest firms is that much of their integration can be bought rather than built.

MID-SIZED

Exposed

This tier is the exposed middle: it typically has enough systems to matter and enough fragmentation to hurt.

LARGE AND GLOBAL

BlackRock reports in its 2025 annual report that Aladdin began as a single-workstation risk tool and is today a centralized operating platform serving the end-to-end investment process across public.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST04

AMAT

P04-ST05

APIs, Microservices and Workflow Orchestration: Building Composable, Orchestrated AI

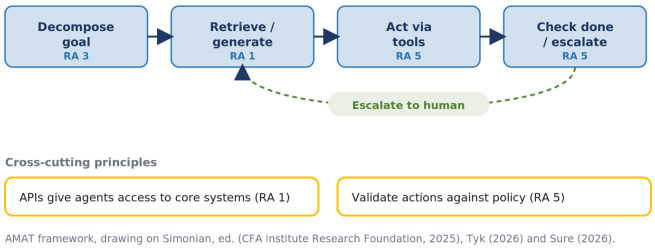
For: Technology

The model is no longer the hard part of an AI system; the architecture that connects and coordinates many models, tools and services is. Single-prompt copilots are increasingly joined by agents that carry a task end to end, and an agent is only as capable as the systems its APIs can reach. The decisive work is composition and orchestration: exposing capabilities through well-designed APIs, packaging AI functions as reusable microservices, and orchestrating them into reliable multi-step workflows. In this guide's view, composability and orchestration, not raw model quality, now determine whether agentic AI delivers in production.

WHERE FIRMS GO WRONG

The most common failure is embedding AI logic across the estate instead of designing composable services, so the same function is rebuilt repeatedly and a change to a model or prompt forces rework across systems. The second is deploying agents on an estate with no API foundation, so they cannot act because the governed interfaces they would call do not exist. The third is expecting coordination to emerge on its own, when bolting single-purpose tools onto different functions produces siloed outputs.

How an agentic workflow actually runs



AMAT framework, drawing on Simonian, ed. (CFA Institute Research Foundation, 2025), Tyk (2026) and Sure (2026).

Takeaway. Build a governed API foundation first, package AI capabilities as reusable microservices with clear boundaries, and orchestrate them into reliable workflows connected by open standards. Constrain autonomy deliberately and engineer workflows for failure and cost, so that a fault degrades gracefully.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST05

RECOMMENDED ACTIONS

1. **Deploy no agent against a system it cannot reach through a governed API**; put standardized, authenticated, rate-limited interfaces in front of the systems the first workflow needs.
2. **Package shared AI capabilities as reusable services** with clear contracts, such as document extraction, kept separate from business logic.
3. **Run high-value multi-step workflows through an orchestration layer**, building toward coordinated agents that share context.
4. **Use open standards such as MCP for new tool connections**, building toward open agent-communication standards in place of proprietary integrations.
5. **Constrain what agents can do**: force structured outputs, gate tool calls and validate high-impact actions against policy, with approval gates and transaction limits.
6. **Give each production workflow retries, a fallback model and a per-run budget**, building toward circuit breakers and asynchronous processing for heavy inference.
7. **Register agents built by many teams through a central platform** with central quality control and access management.

SEQUENCE AND DEPENDENCIES

1. The dependency that gates everything else is the API foundation, and it is worth restating as a sequencing rule rather than a checklist item: the interfaces.
2. Cost is an implementation constraint as much as a design one.
3. The second sequencing reality is that the tooling landscape is immature and moving fast.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Technology builds and runs the composable stack: the API foundation, the AI microservices, the orchestration layer, and the failure-handling and cost controls.
Second Line	Risk and compliance set and challenge the autonomy control requirements, test their adequacy and monitor exceptions, covering whether tool calls are gated.
Third Line	Internal audit tests whether the API foundation and orchestration controls work as designed: whether agents can act only through governed contracts, whether tool gating.
Board and Executive	The board oversees the firm's approach to orchestration and its material risks; executive management approves the platform strategy.
Cross-Functional	Connectivity to core systems (P04-ST04), security of the orchestration surface and agent identities (P04-ST07), sandboxes for testing multi-agent behavior (P04-ST10).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST05

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Uncontrolled agentic execution**	Validate high-impact actions against authorization and policy before they execute, with tool gating, approval gates and transaction limits; insert the checks in the orchestration logic.
HIGH **Cascading orchestration failure**	Engineer for failure with a per-run budget, timeouts, bounded retries for idempotent calls, circuit breakers and tested fallback models.
HIGH **Duplicate or irreversible actions after retries**	Record an action identifier and completion state, make retries idempotent where possible, require approval for irreversible actions and route uncertain outcomes to a named operator before replay.
HIGH **Agents cannot reach systems (no API foundation)**	Put standardized, authenticated, rate-limited and governed interfaces in front of the systems each workflow needs before agents are deployed against them, and extend coverage workflow by workflow.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
API foundation coverage	Share of systems agents need that are reachable through a governed, authenticated API	Agent actions blocked or hand-worked for want of an API
Workflow reliability	Automated recovery coverage (retries, circuit breakers, fallbacks) across orchestrated workflows	Cascading failures, unrecovered faults and duplicated actions in production
Cost per workflow	Cost profiled per workflow type against the single-model baseline	Unbudgeted inference spend on multi-agent runs
Workflow value realized	Share of candidate workflows with an agreed baseline for time, quality and cost	Change in median end-to-end completion time, reviewer corrections and all-in cost per successfully completed workflow versus its baseline

4 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The advantage available to the smallest firms is that they do not have to build orchestration from scratch.

MID-SIZED

This tier often has enough candidate workflows to justify agents but not yet the foundation to run them.

LARGE AND GLOBAL

Scale makes a federated orchestration model feasible, and the BlackRock case in ZenML's summary shows one in use.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST05

P04-ST06

Retrieval-Augmented Generation and Enterprise Knowledge Bases

For: Technology; Investment Teams

Retrieval-augmented generation connects a general language model to the firm's own data at the moment it answers, so outputs are grounded in current, firm-specific documents rather than the model's fixed training knowledge. For an asset manager, it lets a model answer from the research archive, filings, policies and market data, and cite the source it used. In a 2025 systematic review of enterprise RAG studies, compliance and governance and financial reporting were among the most-studied domains, and hallucination was the most-cited challenge, so retrieval reduces the problem rather than removing it.

12.8%

of financial-NLP task assignments in 2025 were retrieval-enhanced question answering, up from 5.5% in 2022

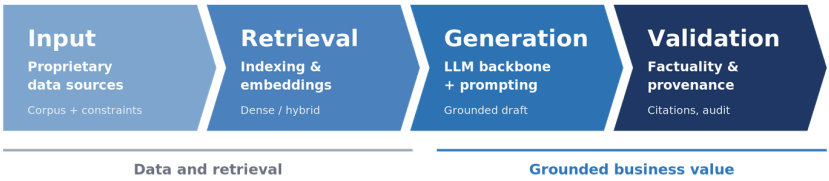
nearly half

of enterprise RAG studies name hallucination as a challenge, the most cited (2025 systematic review)

WHERE FIRMS GO WRONG

The most common mistake is pointing a language model at a large, unstructured pile of documents and expecting good answers, when broad access to a sprawling repository increases both hallucination and incomplete answers. A second is treating retrieval as solved once a vector database is in place, although pure similarity search returns noisy candidates and degrades on multi-part questions. The third is building for the demo rather than production: fewer than 15% of RAG studies address real-time integration, and more than half rely on public data that risks domain shift.

The enterprise RAG value chain: from proprietary data to grounded output



Source: Karakurt and Akbulut, 2026 (RAG-Enterprise Value Chain).

Takeaway. Curate and scope the knowledge base, engineer and test retrieval deliberately, prefer hybrid or structured designs over raw vector search, ground every answer in a cited source, and lock down who can write to the corpus and who can retrieve from it. Grounding reduces hallucination but does not remove it, so the residual risk has to be engineered against.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST06

RECOMMENDED ACTIONS

1. **Constrain what the retriever can search**, starting with one curated repository and building toward a centralized knowledge base.
2. **Restrict corpus write-access to authenticated, trusted sources** and log every ingestion, as a privileged security control.
3. **Test retrieval precision and recall on real queries before go-live**, recording the chunking, embedding and retrieval choices.
4. **Add hybrid or structured retrieval**, starting where proprietary knowledge matters most, rather than defaulting to pure vector search.
5. **Configure strict RAG**: require a retrieved source for every answer, with an "I don't know" below a retrieval-confidence threshold.
6. **Require traceable citations and freshness indicators** on every output, and test that responses derive strictly from retrieved content.
7. **Enforce user entitlements at retrieval**, and make the retrieved context for each answer inspectable before generation.

SEQUENCE AND DEPENDENCIES

1. The dependency most teams underweight is the offline work that has to happen before the first query.
2. A concrete technical ceiling shapes the whole design: real investment documents routinely exceed what a model can hold in a single context window.
3. Structure-first retrieval, curated repositories and task-aware schemas all move effort to an upfront structuring phase, and the payoff, more accurate retrieval in the one preprint that measured it.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Investment and technology teams build and use the retrieval layer and operate the access, corpus and output controls: they scope the corpus, engineer chunking.
Second Line	Risk and compliance set policy for the controls that make grounding safe (corpus write-access, entitlement enforcement at retrieval.
Third Line	Internal audit tests whether corpus write-access is authenticated and logged, whether retrieval enforces entitlements, and whether outputs are genuinely source-cited rather than plausibly worded.
Board and Executive	Senior leadership is accountable, personally where a regime such as SM&CR applies, for a grounded system that breaches an information barrier or cites a fabricated source.
Cross-Functional	Knowledge-management strategy (P03-ST03), data quality and tagging (P03-ST04), security architecture (P04-ST07) and model risk (P05-ST04) are cross-cutting inputs to a retrieval system and are referenced here.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST06

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Silent retrieval failure and hallucination**	Run strict RAG with a no-source fallback and, once validated, a confidence threshold; require traceable citations; test the whole pipeline, not just the output text.
HIGH **Unauthorized retrieval / entitlement bypass**	Enforce user entitlements at the retrieval step, not only the application layer; filter the candidate set by access rights before generation.
HIGH **Retrieval-dependent persistence**	Sanitize and filter retrieved content at the pipeline boundary before context integration; scan the corpus for injected content, not just live queries.
HIGH **Vendor or index change disrupts grounding**	Require notice of material changes and keep a tested fallback; re-run the retrieval test set after any embedding, index or vendor change.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Retrieval quality	Retrieval precision and recall measured against a real investment-query test set	Answer errors traced to a retrieval miss or wrong document
Grounding and citation	Share of answers carrying a traceable citation to a retrieved source	Client-facing or compliance outputs published with no source chain
Abstention discipline	Rate at which the system returns 'I don't know' when no source is retrieved or, once validated.	Confident answers later found to be fabricated (silent-retrieval-failure incidents)
Latency	Retrieval-plus-generation response time on production queries	User abandonment or fallback to manual search

4 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The right archetype is to buy rather than build.

MID-SIZED

Capacity squeeze

This tier faces a capacity squeeze and hits corpus sprawl first: enough internal capability and enough document sources to stand up retrieval across the firm.

LARGE AND GLOBAL

Scale supports institutional-grade retrieval over large proprietary corpora with modular, auditable pipelines; enterprise search is in use at a public institution.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST06

P04-ST07

Security Architecture and Identity/Access Management for AI

For: Technology; Risk and Compliance

When AI is in the loop, security stops being a control added later and becomes a property designed in. An agent holds credentials, reads sensitive data, calls tools and takes actions, and can be steered by whatever text it ingests, which breaks the assumption behind identity and access management built for human users. Australia's prudential regulator notes that identity and access management has not yet adjusted to non-human actors such as AI agents. In a 2026 survey of 496 organizations, security and risk concerns were the most-cited barrier to fully scaling agentic AI.

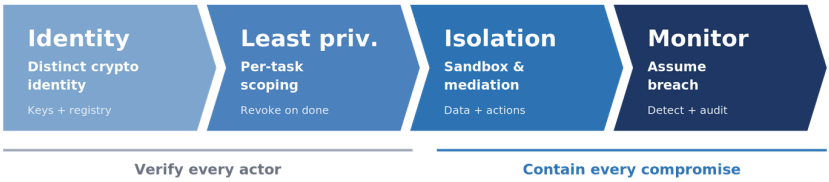
two-thirds
of 496 surveyed organizations, nearly, name security and risk as the top barrier to scaling agentic AI (2026)

about half
of respondents at fund managers under $1 billion report no restrictions on generative AI tools (2025)

WHERE FIRMS GO WRONG

The most common failure is treating AI security as a control layer added after the system works. Firms that skip designing it in end up retrofitting controls onto agents that already have broad access, and where no identity model exists for agents, permissions default to broad. What a compromised AI application can do is set by the permissions and tools it holds, so permission scoping is a security primitive rather than an operational convenience. Shadow AI compounds the problem when staff use public consumer tools on unmanaged channels.

Extend Zero Trust to AI agents as non-human identities



AMAT framework, drawing on Microsoft Security (2026), ACSC/CISA/NSA et al. (2026) and Brodt et al. (2026).

Takeaway. Treat every agent as a non-human identity to be verified and constrained: a distinct identity in a reconciled registry, least privilege scoped to each task, isolation of the model from data and actions it does not need, and firm data kept out of untrusted systems. Closing the identity gap before agents are scaled, rather than after, is the defining security task.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST07

RECOMMENDED ACTIONS

1. **Deny any agent or key not in the trusted registry**, registering every production agent.
2. **Allow-list each agent's tools and remove standing broad access**, with least privilege per task and rights revoked on completion.
3. **Vault agent credentials in dedicated stores**; never embed keys or secrets inside agents.
4. **Require human confirmation for irreversible operations**, and sandbox the highest-impact agent actions.
5. **Classify AI workloads by data sensitivity**, placing MNPI and proprietary-strategy workloads on isolated infrastructure.
6. **Ban public chatbots for firm data** and route staff to approved enterprise platforms.
7. **Constrain agent tool access to honor information barriers**, and alert on agent credential access and unusual outbound traffic.

SEQUENCE AND DEPENDENCIES

1. The dependency that gates everything else is the agent identity model.
2. The second consideration is that these controls carry real usability and effort costs, and the right level is a deliberate trade-off rather than a maximum.
3. Least privilege, credential vaulting, audit and revocation all assume each agent is an identifiable principal.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Technology and business teams build the architecture: they issue agent identities, scope least-privilege access, vault credentials, and isolate models and high-impact actions.
Second Line	The CISO sets the Zero-Trust-for-agents policy and security standards for non-human actors and provides expert challenge; Risk and Compliance independently review how information barriers.
Third Line	Internal audit tests whether the agent registry is complete and reconciled, whether access is genuinely least-privilege and revoked, whether credentials are vaulted.
Board and Executive	Senior leadership is accountable for AI security failures; whether that accountability is also personal depends on the regime (see Risks and Controls).
Cross-Functional	Adversarial and red-team testing (P07-ST07), AI policy and acceptable use (P05-ST03), data-protection obligations (P05-ST07).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST07

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Prompt injection / agent hijack at the architecture level**	Isolate the model with a mediation layer and sandbox high-impact actions; require human confirmation for irreversible operations; assume a filter alone will not hold.
HIGH **Data leakage and the lethal trifecta**	Break the trifecta: isolate sensitive data behind a mediation layer, restrict external egress, and keep confidential work on approved enterprise platforms with enforceable commitments on retention and training.
HIGH **Unmanaged shadow AI**	Ban public chatbots for firm data and route staff to approved enterprise platforms; block unauthorized platforms at the network and apply the P07-ST04 vendor terms.
HIGH **IAM not adapted to non-human actors**	Give every agent a distinct cryptographic identity in a reconciled trusted registry; bind identities to roles and deny anything not in the registry.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Agent identity coverage	Share of production AI agents with a distinct identity registered in the trusted registry	Access events or incidents traced to an unregistered or shared agent identity
Least-privilege enforcement	Share of agents whose data and tool access is scoped to task and revoked on completion	Over-permissioned agents found in review; blast radius of any incident
Isolation of high-impact actions	Share of irreversible or high-impact agent actions behind a sandbox and human-confirmation gate	Incidents where a compromised agent reached data or actions outside its remit
Shadow-AI closure	Share of staff routed to approved enterprise AI platforms; unauthorized platforms blocked	Firm data found in unapproved or public AI tools

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The exposure is the shadow-AI gap: in the 2025 AIMA survey, around half of respondents at fund managers under $1bn in AUM reported no restrictions on Gen AI use.

MID-SIZED

This tier typically has enough agents deployed to matter but no identity model for them (an AMAT firm-size reading).

LARGE AND GLOBAL

Scale makes an enterprise Zero Trust program extended to non-human identities both feasible and necessary (an AMAT firm-size reading).

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST07

AMAT

P04-ST08

Build, Buy and Partner Strategy for AI Capability

For: Senior Leadership; Technology

Build, buy or partner is not one decision made once but a sequence made for every AI capability, and a firm that gets it wrong is often answering a vendor's question, which product to buy, instead of its own: whether the capability needs to be differentiating or is table stakes. A structured six-decision framework gives firms a repeatable way to answer the question capability by capability. This article sets out that framework, the evidence on when each path fits, and the governance that keeps a hybrid stack coherent.

21%

of surveyed asset management professionals reported building AI in-house, while 63% used third-party platforms (industry survey; overlapping responses)

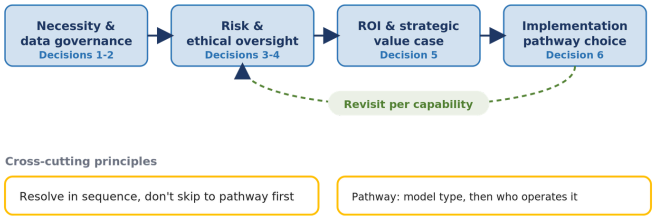
one in ten

is the pure-outsourcing share in a separate 12-firm survey of firms in Italy (self-reported)

WHERE FIRMS GO WRONG

The most common failure is treating build, buy or partner as a single, one-time decision rather than a capability-by-capability discipline, so the firm makes a default lock-in choice without noticing until it tries to change course. A second is assuming generic, off-the-shelf tools are adequate for institutional investment management, when they often cannot handle structured and unstructured data together or meet compliance standards. The third is organizational: the business side is reluctant to give up costly customizations on a vendor platform, and vendors have every incentive to add more.

The six-decision sequence for AI capability sourcing



AMAT framework, drawing on Tavasoli, Sharbaf and Madani (2025; six-decision framework).

Takeaway. Getting build, buy and partner right is not about insourcing the most or the least. Run every material AI capability through the same deliberate sequence, decide what differentiates the firm, and govern the vendor relationships it keeps instead of letting customization accumulate by default.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST08

RECOMMENDED ACTIONS

1. **Run every material AI capability through the six-decision sequence** before choosing build, buy or partner.
2. **Reserve in-house build for differentiating capability**, defaulting to vendor or hybrid for the rest.
3. **Decide per use case**, not once for the whole firm, using the taker, shaper, maker lens.
4. **Require enterprise-grade, tenant-isolated AI tools with audit logging** for sensitive workflows; no public consumer tools.
5. **Route each use case by data sensitivity** to a commercial API or an internally hosted model.
6. **Govern vendor customization** with a fit-to-standard default, deviating only for documented, differentiating cases.
7. **Add flexibility and interoperability to selection criteria** at the next core-platform review, and record what is built and bought.
8. **Partner with specialist vendors and fintechs** where in-house build is not yet affordable, reassessing as the firm scales.

SEQUENCE AND DEPENDENCIES

1. Sequencing matters here as much as in any technical build.
2. The build-vs-buy decision also has an organizational dependency that is easy to underweight: fit-to-standard governance and the six-decision framework both require someone with the authority to say no.
3. The six-decision framework only works if it is applied before a vendor relationship exists, not retrofitted afterwards; in this guide's view.

WHO DOES WHAT: LINES OF DEFENSE

First Line	First-line teams propose build/buy/partner choices and own the six-decision-framework paperwork for their own capability requests, including the documented business case required for any vendor customization.
Second Line	Risk and compliance should require that the risk-management and ethical-oversight decisions in the six-decision framework are resolved and evidenced before procurement proceeds.
Third Line	Internal audit tests whether capabilities were actually routed through the six-decision framework, not only whether a build/buy/partner decision was documented after the fact.
Board and Executive	The board receives portfolio reporting on AI investment across data, talent, infrastructure and sourcing.
Cross-Functional	Vendor ecosystem management spans first, second and third line by design; cross-reference P03-ST10 for data marketplace vendors specifically and P07-ST04 for due diligence as a risk.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST08

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Vendor and provider concentration**	Map provider concentration across the AI stack explicitly and, for each critical capability, record the upstream dependencies it concentrates on and an achievable continuity option.
HIGH **Default lock-in without a defined strategy**	Require the six-decision sequence before any vendor commitment, so the lock-in decision is made deliberately rather than by default.
MEDIUM **Off-the-shelf tools may not meet institutional workloads**	Test any off-the-shelf tool against the firm's actual compliance and orchestration requirements before committing, not against a vendor demo.
MEDIUM **Ungoverned platform customization**	Apply fit-to-standard governance with a documented business case required for any deviation.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Workflow value	Comparable cycle time before and after deployment, with volumes and review quality recorded	Sustained cycle-time improvement without more errors or rework
Net run-rate cost	Total license, integration, infrastructure, personnel and review cost per completed workflow	Observed cost change at steady production volume
Provider resilience	Share of material workflows with a tested fallback or exit, and measured provider outages	Minutes of service unavailability and the workflow losses they cause
Human oversight	Rate at which reviewers edit, override or reject AI output, trended (if reviews are logged)	A sustained drop to or near zero despite meaningful volume

5 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Building proprietary investment AI is a scale luxury, and specialist vendors are emerging to serve smaller firms in narrower domains such as corporate credit research.

MID-SIZED

Capacity squeeze

This tier faces a capacity squeeze between the boutique partnering pattern and the large-firm insourcing pattern: growing internal capability makes selective build newly affordable.

LARGE AND GLOBAL

Scale supports the core-data-layer and reusable-recipe patterns described above, and, in this guide's view, boards at this scale should receive reporting on AI investment across data, talent.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST08

P04-ST09

Model Observability and Performance Monitoring: Seeing What AI Does in Production

For: Technology; Risk and Compliance

A model that passed its pre-deployment checks does not stay correct: in live markets it drifts as conditions move away from its training data, its edge decays, and it can fail silently without throwing an error. Model observability lets a firm see what its AI is doing in production, through telemetry, drift detection, output-quality evaluation and a tamper-evident audit trail. For asset managers degradation is an investment problem as well as a quality problem, because AI systems that trade on their own outputs can erode the signal they were built to capture.

a quarter

of 102 surveyed financial-services IT decision makers already use agentic AI in observability (2025, vendor-sponsored)

37%

plan to do so within the next two years (same survey)

WHERE FIRMS GO WRONG

A common failure is assuming the monitoring tools the firm already owns will do the job. Application monitoring shows that the service is up while leaving the firm blind to whether its answers are still any good. That produces silent failures, because a model upgrade, a prompt refinement or a data update can introduce new errors or reintroduce fixed ones with no change to the application code.

Traditional monitoring cannot see inside AI

Dimension	Traditional application performance monitoring (APM)	AI observability
Watches	Uptime, latency, error rates	Drift, quality, tool use, reasoning
Cadence	Thresholds on known metrics	Continuous eval + drift detection
Failure	A crash or exception	Silent degradation, no crash
Audit	Logs the request	Tamper-evident decision trail

AMAT framework, drawing on Uslu (2026), Kapoor (2026) and Ojewale, Suresh and Venkatasubramanian (2026).

Takeaway. Pre-deployment sign-off is a snapshot, not a control. Instrument models for continuous observability, with a tamper-evident audit trail as the shared foundation, drift detection and behavioral monitoring on top, output-quality evaluation on every release, and a feedback loop that turns human review into systematic improvement.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST09

RECOMMENDED ACTIONS

1. **Log the model version, configuration and decisions** of production models, building toward a tamper-evident, append-only audit trail.
2. **Route drift alerts to the second line** from the first model in production, and treat back-tests as insufficient evidence of live health.
3. **Run a regression suite of known higher-risk prompts before each release**, building toward output-quality evaluation in the deployment pipeline.
4. **Add output accuracy and hallucination rate to existing dashboards**, building toward AI-specific metrics beyond uptime.
5. **Log every action and tool call of any production agent**, adding a fairness layer where a model influences a client outcome.
6. **Record every human review and override with the reviewer's reasoning**, and feed it into versioned evaluation prompts and retraining.
7. **Name an owner for each production model's alerts**, with authority to roll back or retire the model.

SEQUENCE AND DEPENDENCIES

1. The dependency that gates the rest is instrumentation: drift detection, quality metrics and audit querying all assume the events they need are being captured in the first place.
2. Several adjacent areas belong to neighboring articles and are referenced, not solved, here.
3. Continuous inference and fine-tuning across interacting models generate volumes of telemetry that only automated summarization, querying and anomaly-surfacing can make usable.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Technology builds and runs the observability stack: the audit trail, drift detection, evaluation pipelines and behavioral logging.
Second Line	Risk and compliance set the monitoring standard and the materiality at which a monitored deviation escalates, challenge material escalations and exceptions.
Third Line	Internal audit tests whether the observability layer works as designed: whether the audit trail is complete and tamper-evident, whether drift and quality monitoring actually fire.
Board and Executive	The board owns the observability investment and should see model-incident, drift and monitoring-coverage metrics, not only the count of models deployed.
Cross-Functional	Ongoing monitoring as a control discipline (P07-ST02), independent model validation (P07-ST05), pre-deployment testing (P07-ST01), outcome feedback loops (P03-ST03).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST09

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Modeled signal erosion**	Monitor live signal performance and drift continuously, not with back-tests; watch the signal-to-noise ratio decay and retire or recalibrate a signal before it inverts.
HIGH **Competence illusion**	Measure output quality against independent, out-of-sample reality, not self-consistent metrics; treat suspiciously stable accuracy as a monitoring signal in its own right.
HIGH **Silent degradation goes unnoticed**	Deploy continuous drift detection with alerts to the first-line owner and escalation of material exceptions to the second line.
HIGH **Hallucination regression with no code change**	Run regression suites of higher-risk prompts on every release inside the deployment pipeline; track error and hallucination rates across versions.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Model drift	Population-stability / distribution shift on inputs and outputs, alerting the owning first line.	Live performance divergence detected late or via client impact
Output accuracy	Share of a risk-based sample of outputs failing a defined material-error test	Downstream errors traced to unreviewed low-quality outputs
Hallucination rate	Regression-suite hallucination rate run on every release	Rising hallucination trend detected only in production
Agent behavior	Tool-selection accuracy, context adherence and step-wise latency for agentic workflows	Unexplained agent actions with no telemetry to reconstruct them

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The smallest firms can lean on the monitoring built into the vendor platforms they already use rather than an internal build, so the bar to start is low.

MID-SIZED

This tier often has enough models in production to drift but not yet the monitoring to see it, which is where silent regression, behavior changing with no code change.

LARGE AND GLOBAL

At scale, observability becomes indispensable: beyond latency, throughput and error rates, the firm has to watch model-specific signals across every model it runs.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST09

P04-ST10

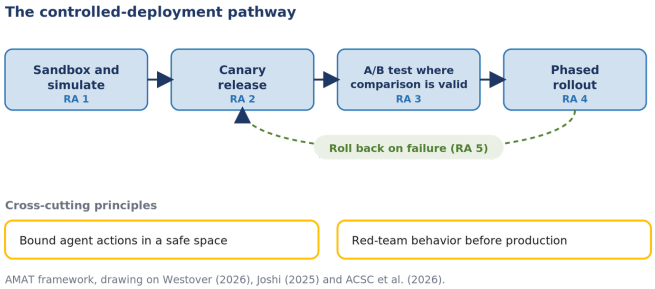
Sandboxes and Controlled Deployment: Experimenting with AI Without Betting the Firm

For: Technology; Investment Teams

The gap between an AI capability that works in a demonstration and one that is safe in production is where firms get hurt, and a sandbox is how they cross it without betting the firm. A production-like sandbox, a bounded safe action space for agents and a staged path to live use through canary, A/B and phased rollout let a firm find failures on a small slice before they reach a client. Supervisors now support this: the UK FCA runs an AI sandbox and a live-testing scheme, and the EU AI Act, as amended, requires each member state to establish a regulatory sandbox, though these grant no relaxation of the rules. A firm that never ships an AI change straight to production can experiment boldly because it can fail safely.

WHERE FIRMS GO WRONG

The common failure is deploying AI straight into production because it worked in a demonstration, with no controlled path in between. A firm that skips canary, A/B and phased rollout finds its failures in production, where the only remedy left is to roll back after the damage is done. The failure is sharper for agentic AI, because an agent takes actions as well as producing outputs, so experimentation has to happen inside a bounded safe action space that contains what the agent can actually do.



AMAT framework, drawing on Westover (2026), Joshi (2025) and ACSC et al. (2026).

Takeaway. A sandbox moves the point of failure from the client to a controlled environment the firm can afford to break. Build a production-like sandbox with a bounded action space for agents, release through canary and A/B stages with exit criteria, roll out in phases with a way back, and red-team agent behavior before it goes live. A regulatory sandbox relaxes no rules.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST10

RECOMMENDED ACTIONS

1. **Bound what any agent can do before it touches live systems**, using a sandbox for AI changes that can act or reach clients.
2. **Release each material AI change to a small group first**, with canary deployment to a defined slice of traffic as the standard route.
3. **A/B test material AI changes against the current process**, with success and exit criteria set in advance.
4. **Keep a way back for every AI release**, with phased rollout as the default route to live.
5. **Red-team agents that can take actions in live systems** in the sandbox before every production release.
6. **Engage a supervisory AI sandbox or live-testing scheme** for novel, higher-risk or infrastructure-light cases.
7. **Treat the sandbox as an enabler, not a compliance pass**: regulatory sandboxes grant no rule relaxation and accountability stays with the firm.

SEQUENCE AND DEPENDENCIES

1. The consideration that determines whether any of this works is sandbox fidelity: a sandbox that does not resemble production gives false confidence.
2. Several adjacent areas belong to neighboring articles and are referenced, not solved, here.
3. Fidelity does not mean copying confidential client or portfolio data into a lower-control test environment.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Technology builds and runs the sandbox, the safe action space, the canary and phased-rollout pipeline and the incident controls.
Second Line	Risk and compliance set which AI changes must go through a controlled rollout and the minimum exit and rollback thresholds.
Third Line	Internal audit tests whether AI changes genuinely went through the controlled path rather than shipping straight to production, samples release evidence.
Board and Executive	The board owns the appetite for experimentation speed versus deployment control and should expect that no material AI change, and especially no agentic one.
Cross-Functional	Pre-deployment testing as a control discipline (P07-ST01), cybersecurity and adversarial red-teaming (P07-ST07).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST10

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Big-bang deployment**	Release through a canary deployment to a slice, A/B test against the baseline where a valid comparison is possible.
HIGH **Uncontained agent actions**	Run agents inside a bounded safe action space in a production-like sandbox before any live use.
HIGH **Test-to-production data leakage**	Use approved masked or synthetic data, separate credentials, restricted tool allow-lists and tested environment boundaries before any run.
MEDIUM **Untested behavior reaches production**	Red-team agent behavior and resilience in the sandbox before production (methodology in P07-ST07).

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Escaped changes	Share of material AI releases and expansions completed with no live incident and errors within the agreed baseline	Incidents or user- or client-impacting errors per material release or expansion, with severity and time to containment.
Stop and unwind	Share of material AI services whose stop has been exercised and restored the prior workflow within the firm's.	Actual restore time when a stop is invoked in live use
Controlled-rollout coverage	Share of AI changes released via canary or phased rollout rather than big-bang	Incidents traced to a change shipped straight to all users
Agent containment	Share of agentic capabilities tested in a bounded safe action space before live use	An agent taking an unintended real action during experimentation

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The smallest firms can treat the sandbox as something to borrow rather than build: in the UK.

MID-SIZED

This tier is where AI adoption is most likely to outrun controlled-deployment discipline, and where changes start reaching production without the staged path set out above.

LARGE AND GLOBAL

The largest firms have the scale to run an enterprise sandbox across the estate, and can take their most novel systems to supervisory live testing where it is offered.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST10

P04-ST11

Vendor Interoperability and Avoiding Lock-In: Keeping the Exit Open

For: Technology; Senior Leadership

Every AI capability a firm adopts creates a dependency, and the market supplying those capabilities is consolidating into a handful of foundation-model, cloud and hardware providers. The question is no longer whether an asset manager depends on vendors but whether it can still move if one fails, reprices or changes its terms. AI lock-in compounds across the model, the orchestration layer, the runtime and the working patterns of the teams that build on it, so switching costs harden quietly.

44%
of third-party AI models used by UK financial institutions came from the top three providers in 2024, up from roughly 18% in 2022

9%
of 131 surveyed asset managers named vendor concentration among their largest regulatory blind spots, while 63% use vendor or off-the-shelf AI tools (Mercer, 2026)

WHERE FIRMS GO WRONG

A common failure is to integrate first and think about exit later, then find the exit plan is a document rather than a tested capability. Australia's prudential regulator found several regulated entities heavily dependent on a single AI provider with no tested substitution or exit strategy. AI lock-in also compounds across the model, orchestration, runtime and developer patterns at once.

Captured versus portable AI architecture

Dimension	Locked in	Built to move
Providers	One vendor, many use cases	Multiple vendors, model families
Switching cost	Compounds across the stack	Contained by abstraction layers
Exit plan	Documented, never tested	Rehearsed fallback, tested
Negotiating position	Price-taker, terms dictated	Credible exit is leverage

AMAT framework, drawing on MinterEllison (2026) and Waehner (2026).

Takeaway. Vendor dependence is unavoidable, but lock-in is a choice made by default when a firm integrates for speed and defers exit. Make concentration visible through a vendor register, prove exits by rehearsing them, and contract for portability the firm can actually invoke.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST11

RECOMMENDED ACTIONS

1. **Maintain an AI vendor register**, including AI embedded inside SaaS tools, that maps every critical dependency and where it is used.
2. **Assess concentration risk for each critical provider**, modeling the impact of a sudden loss of service or a material change in model behavior.
3. **Rehearse the exit for every critical provider** with a continuity fallback, tested rather than documented.
4. **Write exercisable termination and portability provisions** into every new critical AI contract, and add them at each renewal.
5. **Make exit feasibility and interoperability a procurement criterion**, favoring open standards and an abstraction layer at the model interface.
6. **Name a second, structurally different model family** for the strategic decisions that rely most on AI.
7. **Review provider concentration at each vendor renewal**, moving to continuous monitoring within third-party oversight.

SEQUENCE AND DEPENDENCIES

1. The dependency that gates everything else is visibility: concentration assessment, exit testing and portability all assume the firm knows which AI it relies.
2. A bespoke architecture is hard to reverse once embedded in production, so the interoperability and portability decision belongs at the design stage, not after the system is live.
3. The second reality is that exit rights are only worth what they are worth in practice; a rehearsed fallback that maintains continuity is the difference between a resilient firm.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Technology builds and runs the abstraction layers, portability and multi-vendor architecture and owns the vendor register.
Second Line	Risk sets and monitors provider concentration limits and owns the evidence that exits are tested.
Third Line	Internal audit tests whether exit plans are genuinely executable rather than merely documented, whether concentration is within appetite.
Board and Executive	The board owns the firm's concentration appetite and the trade-off between the speed of deep vendor integration and the flexibility of a portable architecture.
Cross-Functional	The build, buy or partner decision (P04-ST08), vendor due diligence (P07-ST04), data-vendor strategy (P03-ST10) and the open orchestration and agent-communication standards (P04-ST05) are cross-cutting.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST11

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Untested single-provider concentration**	Run a concentration risk assessment per critical provider and keep a rehearsed, executable exit with a continuity fallback.
HIGH **Compounding agentic lock-in**	Design for portability from the start: an abstraction layer at the model interface and open interoperability standards (P04-ST05).
HIGH **Architecture lock-in**	Assess architecture reversibility and cost early, and avoid committing to a bespoke multi-agent design before requirements are understood.
HIGH **Systemic provider concentration**	Diversify across vendors and model families and set firm-defined concentration limits on any single critical dependency.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Provider concentration	Share of critical AI workloads on the single largest provider	Loss of service or repricing that the firm cannot absorb
Exit readiness	Share of critical providers with a rehearsed, executable exit and continuity fallback	Time and service continuity achieved in the most recent controlled failover or withdrawal test.
Portability	Share of models and data with a documented, exercised migration path	Migration blocked by proprietary formats at the point of need
Model-family diversity	Strategic decisions that rely most on AI with a second structurally different model family tested.	A single withdrawn model halting a strategic workflow

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Exposed

Unusually for this pillar, the smallest firms are the most exposed: in this guide's view they are the most reliant on vendor solutions and therefore on repricing.

MID-SIZED

This is the tier where lock-in is most often engineered in unnoticed, because it starts building bespoke architecture as it scales.

LARGE AND GLOBAL

The largest firms have the scale to build abstraction layers, run genuinely multi-vendor architectures on open interoperability standards and monitor concentration across the estate.



Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P04-ST11

P04-ST12

Cost Management Across Models, Compute and Data: What AI Really Costs at Scale

For: Finance; Technology; Senior Leadership

The unit price of AI is falling while the total cost climbs, and a firm that watches only the first number is the one surprised by the second. As token prices fall, agentic workloads consume far more tokens per task and the bill rises with usage. The costs that break budgets are usually those left out of the business case: governance and compliance overhead, organizational change, failure and recovery, and the recurring run cost of models that can cost more to operate than to build. This article covers the true total cost of an AI capability, the levers that control it, such as routing and tiering workloads, right-sizing models and cost-aware architecture, and how to measure and govern spend.

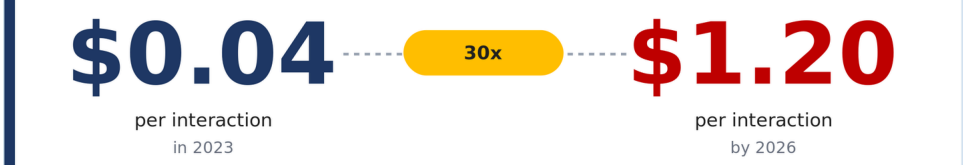
$1.20
per AI-assisted customer-service interaction, from around four cents two years earlier, as workflows expanded (2026 EY analysis)

10 to 20%
of build cost is the typical annual run cost of traditional IT; gen AI at scale can rival or exceed its build cost

WHERE FIRMS GO WRONG

A common failure is to cost only the visible part of AI. Most business cases include tokens, licenses and platform infrastructure, and leave out governance and compliance burden, organizational change, expected failure and recovery, and potential AI-specific levies. A second is treating AI systems as finished builds, when gen AI at scale can incur recurring costs that rival or exceed the original build. The third is masking an inefficient architecture with cheap tokens, which defers the reckoning to the point of scale.

AI unit costs fall, AI bills rise



One customer-service AI assistant as its workflow expanded; illustrative, cross-industry. Source: EY (2026).

Takeaway. The falling price of AI is real, and it is the wrong number to manage. Cost the whole thing up front, match models to work by routing and right-sizing, engineer efficiency into the architecture, budget the recurring run cost, and monitor and attribute consumption as it moves with usage.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST12

RECOMMENDED ACTIONS

1. **Cost all seven categories of the total cost of agents** in every AI business case.
2. **Route and tier workloads across a portfolio of models**, starting with one routine, high-volume workload.
3. **Make right-sizing the default**, requiring a justification whenever the top-tier model is chosen.
4. **Build cost-aware architecture**: semantic caching, selective invocation, fallback models and tiered service levels.
5. **Budget the recurring run cost as an ongoing line**, estimated up front.
6. **Decide deliberately where costly explainability is worth paying for**, rather than assuming interpretability is free.
7. **Attribute AI spend by team, model and workload** and monitor it as usage scales.

SEQUENCE AND DEPENDENCIES

1. Total-cost visibility is the target state, built alongside the first cost controls rather than before them.
2. Several adjacent areas belong to neighboring articles and are referenced, not solved, here.
3. Start with the largest invoice items and a transparent estimate of the governance, change and recovery costs that business cases usually leave out.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Technology implements routing, right-sizing, caching and usage instrumentation.
Second Line	Risk and compliance ensure cost decisions do not quietly drop required controls (for example skipping explainability to save compute) and that provider-pricing dependency is within appetite.
Third Line	Internal audit tests whether business cases carried the full total cost of agents, whether run costs were budgeted and realized.
Board and Executive	The board owns the firm's AI cost appetite and the trade-off between speed of adoption and cost discipline.
Cross-Functional	The investment case and funding model (P08-ST05), ROI measurement (P08-ST03), cloud and compute economics (P04-ST03).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST12

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Consumption cost spiral**	Route and tier workloads across a model portfolio and engineer cost-aware architecture rather than relying on cheap tokens.
HIGH **Hidden costs omitted from the case**	Cost the six current categories of the total cost of agents up front, carry potential AI levies as a scenario, and re-forecast as usage scales.
MEDIUM **Unbudgeted run cost**	Estimate and budget the recurring run cost up front as an ongoing line item.
MEDIUM **Stranded by rising prices**	Forecast for consumption-based and rising prices; keep model choices substitutable so the firm is not locked to one provider's economics (interoperability in P04-ST11).

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Total-cost coverage	Share of AI business cases costing the six current TCA categories.	Programs running materially over an approved budget
Cost per unit of work	Cost per completed, quality-checked task for the largest AI-enabled workflows.	Cost per accepted task and share of tasks needing rework, where logs allow
Workload tiering	Share of eligible tasks routed to approved lower-cost models	Net savings after retries, human review and failures
Run-cost budgeting	Share of deployed capabilities with a funded recurring run-cost line	Run costs competing unplanned with new projects for funding

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

The smallest firms benefit from consumption pricing, which removes the large upfront build and can keep AI infrastructure a small absolute cost.

MID-SIZED

A mid-sized manager can face a particular challenge when usage expands across desks before a consistent cost-allocation process is in place.

LARGE AND GLOBAL

The largest firms have the volume to make model-portfolio routing worthwhile and to justify on-premises economics for steady, high-volume inference, the compute detail of which sits in P04-ST03.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P04-ST12

AMAT

Pillar 5 Executive Summary

For: Board and Executive; Risk and Compliance; Senior Leadership · Also relevant to: Legal, Technology, Investment Teams

THE ARGUMENT

AI in regulated asset management is governed whether or not a firm chooses to govern it. Pillar 5 is the floor of the guide: the controls and obligations that keep an AI-using firm safe inside the regulated environment, regardless of strategic ambition. It spans the governance model and board accountability, policy, model risk, conduct, data protection, human oversight, explainability and the AI inventory, then incident response, cross-border compliance and systemic risk. The firms that get this right treat governance less as a brake on AI than as a feature that AI leaders tend to have, and as how they show a supervisor or client that their AI is under control.

1 in 6
of Dutch asset managers (one 2026 survey) report adequate AI-risk structures and controls

99%
of large firms surveyed by EY in 2025, across sectors, report AI-related losses (avg US$4.4m)

EUR 35m
or 7% of turnover: the EU AI Act's top penalty, for prohibited practices

WHY NOW

Existing technology-neutral rules apply to AI now, and AI-specific obligations are being added. The EU AI Act's prohibitions already apply; its high-risk duties apply from December 2027, under the Act as amended, and most asset-management uses fall outside them. Authorized EU fund managers already carry DORA's incident and ICT third-party duties, AI-related losses are widespread, and the Bank of England has assessed correlated, concentrated AI as a systemic risk, so governance is supervised and cross-border, not optional.

RISKS TO AVOID

Governance as paperwork: a policy library and org chart no one can trace to an actual AI decision or an accountable owner. **Unmanaged, unrecorded AI**: models in production without validation, oversight, or a record complete enough to explain a decision to a regulator or client. **Blind to the outside**: missing extraterritorial and cross-border obligations, and treating correlated, concentrated AI as someone else's systemic problem.

Bottom line. Pillar 5 is the floor: the governance, risk, regulatory and control obligations a regulated asset manager must meet to use AI at all. The firms that lay it before scaling, rather than bolt it on afterwards, are better placed to show any supervisor or client that their AI is under control.

Chapter Leads, Pillar 5: Declan Sheehy

Explore Pillar 5 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P05-summary

PRIORITY MOVES FOR LEADERSHIP

1. **Lay the governance floor first**: name owners and a decision-rights model, including who can stop a live system, and put board accountability on the AI decisions that cannot be delegated.
2. **Build the AI inventory and classify by risk**: everything else operates on the register, so keep it complete, current and risk-tiered, under a policy calibrated to real use.
3. **Extend model risk management to AI**: validate for hallucination, drift and sensitivity to irrelevant context, with oversight calibrated to risk, clear escalation and an unconditional failsafe on irreversible decisions.
4. **Make every AI decision evidenceable**: explainability, transparency and records built into the pipeline, starting from the books and records already kept, to the standard the regulator and client require.
5. **Run one high-bar, cross-border program**: map each system to the regimes it touches, prepare in-scope systems for the EU AI Act's high-risk duties (from December 2027, as amended), and assess each system for the personal data it processes.
6. **Design for failure and for the system**: AI-specific incident runbooks and kill switches, plus model, data and signal diversity and macroprudential engagement against systemic monoculture.

IN THIS PILLAR

P05-ST01	AI Governance and Decision Rights: Turning Ownership of AI Into a Working System	130
P05-ST02	Board and Senior Management Accountability: Owning the AI Decision You Did Not Make	133
P05-ST03	AI Policies, Standards and Acceptable Use: From a Document on the Shelf to a Control That Holds	136
P05-ST04	Model Risk Management for AI: Validating Systems That Do Not Stay Correct	139
P05-ST05	AI and the Regulators: Meeting Supervisory Expectations Without Waiting for AI-Specific Rules	142
P05-ST06	Conduct Risk, Fiduciary Duty and Client Outcomes: When the Algorithm Advises, the Firm Still Answers	145
P05-ST07	Data Protection and Privacy: Governing the Personal Data Behind the Model	148
P05-ST08	Human Oversight and Escalation Protocols: Keeping a Person Able to Overturn the Machine	151
P05-ST09	Explainability, Transparency and Recordkeeping: Making an AI Decision You Can Reconstruct	154
P05-ST10	AI Inventory and Use-Case Classification: You Cannot Govern What You Have Not Counted	157
P05-ST11	Cross-Border AI Regulation: Governing One AI Estate Across Many Rulebooks	160
P05-ST12	Incident Management and AI Failure Response: Detecting, Containing and Learning When AI Goes Wrong	163
P05-ST13	Systemic AI Risk and Market Stability: When Everyone's AI Makes the Same Bet	166

Explore Pillar 5 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P05-summary

P05-ST01

AI Governance and Decision Rights: Turning Ownership of AI Into a Working System

For: Senior Leadership; Risk and Compliance; Board and Executive

Every other control in this pillar assumes a governance model that says who owns each AI decision and how that ownership is enforced. AI governance is a permanent operating state rather than a project with an end date. In a 2026 survey of Dutch asset managers, only around one in six said their structures and controls were adequate to manage AI risk while still enabling innovation.

one in six

of surveyed Dutch asset managers say their structures and controls are adequate to manage AI risk (2026)

1.5 times

as likely to report a cross-functional AI governance board among AI leaders in a self-reported cross-sector survey (2026); 1.7 times for a responsible-AI framework

WHERE FIRMS GO WRONG

The most common mistake is the wrong shape of governance rather than its absence. A separate AI track alongside the standard development lifecycle can duplicate controls and cause confusion and delay, as one financial-services example suggests, and housing oversight in a single silo keeps it from being embedded where the business holds the risk. Where ownership is unclear, risk accumulates in the gaps, and AI that a vendor switches on inside an existing platform can reach production without an owner, an approval or an inventory entry.

Embed AI governance in the three lines, do not build a parallel one

Dimension	Parallel AI org	Embedded in 3 lines
First line	Separate AI team owns build	Business + IT own it
Second line	Standalone AI committee	Governed in existing frame
Third line	Ad hoc AI reviews	Audit assures within mandate
Result	Confusion, delay, duplication	One spine; scales with autonomy

AMAT framework, drawing on Weinert and Gürsching (2026) and PwC (2025).

Takeaway. Governance is the floor of this pillar and has to be laid first. Embed AI oversight in the three lines the firm already runs, give it a cross-functional committee anchored to a recognized framework, make decision rights explicit including the right to stop, scale oversight to what the AI does, and bring vendor-enabled AI inside the same model.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST01

RECOMMENDED ACTIONS

1. **Give AI governance one accountable home** inside the existing three lines of defense.
2. **Establish a C-suite-appointed, cross-functional AI governance committee** with authority to block launch or require conditions on higher-risk use cases.
3. **Publish a decision-rights matrix** naming who proposes, approves, vetoes, suspends and is notified.
4. **Tier governance intensity** by what the AI does and whose outcome it affects: assistive, advisory or autonomous.
5. **Route new AI use cases through the existing product lifecycle** with AI templates, with no separate AI track.
6. **Fund AI governance and name owners now**, starting by recording AI oversight spend in the next budget cycle.
7. **Review and update the framework on a set cadence**, re-baselining maturity as AI capability moves.

SEQUENCE AND DEPENDENCIES

1. Sequence matters more than speed, with one exception: ownership cannot wait.
2. MindForge's handbook says responsibility for AI governance is generally not a standalone capability and is often integrated with existing governance functions.
3. The dependency chain runs home, then forum, then decision rights, then proportionality, but a firm should name an interim business owner and a stop authority for every live use case.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the responsible build and operation of AI and its live use cases, each with a named person able to stop it.
Second Line	Govern AI within the existing frame: set the AI policy and control standards, propose risk appetite for board approval and report breaches of it.
Third Line	Independently assure the governance model and its decision rights, testing that oversight is operating in practice, not merely documented.
Board and Executive	Hold accountability for the governance model, personal where a regime such as SM&CR applies; name the accountable senior executives, approve risk appetite and delegation limits.
Cross-Functional	The C-suite committee spans IT, legal, compliance, and the business, reviews higher-risk use cases and material changes, and sets the criteria for delegated approvals.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST01

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Accountability gaps ('AI between the lines')**	Map decision rights explicitly (propose / approve / veto / reclassify / suspend / retire / notify) for every significant AI decision.
HIGH **Parallel-track governance**	Collapse AI requirements into the existing lifecycle so there is one governance spine, not two.
HIGH **Siloed ownership**	Assign shared ownership across the three lines through a cross-functional committee.
HIGH **Governance lags adoption**	Adopt a proportionate, AI-specific framework anchored to a recognized standard and re-baseline maturity.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Governance coverage	% of AI use cases logged in the inventory and risk-tiered, including AI enabled inside vendor platforms.	% of production models with no assigned owner
Governance maturity	self-assessed maturity by governance dimension	independent audit rating of the model
Decision-rights clarity	% of significant AI decisions with a named approver, veto holder and person able to suspend	incidents traced to unclear ownership
Oversight quality	Rate at which reviewers edit, override or reject AI output, trended.	Material AI outputs accepted without a substantive check, and errors found after use

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The advantage is simplicity: a proportionate model with a named owner, a quarterly review, and a change log is lightweight but real.

MID-SIZED

Capacity squeeze

The capacity squeeze is where the governance gap bites hardest, which is why the priority there is the committee and the three-lines embedding rather than more use cases.

LARGE AND GLOBAL

The move is to run a genuine multi-layer oversight model and then spend its effort on consistency.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST01

P05-ST02

Board and Senior Management Accountability: Owning the AI Decision You Did Not Make

For: Board and Executive; Senior Leadership

Accountability for AI is the one governance obligation a board cannot delegate, automate or outsource. In the UK and EU, supervisors take the position that responsibility for AI-driven outcomes stays with senior management and the management body, whether decisions are taken by people or by AI-based tools. In US corporate law, two legal scholars have proposed that failing to use and oversee AI may come to be treated as a breach of the duty of care, a scholarly argument and not settled law. Naming an accountable person is worthless without the competence, authority and time to exercise the role. A firm that assigns accountability on paper has documented the gap rather than discharged the duty.

12%

of surveyed C-suite executives at large organizations, across sectors, could match the right controls to five common AI risks; chief risk officers scored 11% (2025)

nearly three quarters

of boards were perceived to have only moderate or limited AI expertise (2026 survey)

WHERE FIRMS GO WRONG

The internal failure is to treat accountability as an assignment rather than a capability. Firms name an accountable executive to satisfy the regime and leave that person unable to exercise the role: in EY's 2025 survey of C-suite executives at large organizations, a cross-sector sample, only 12% could match the right controls to five common AI risks, and chief risk officers scored 11%. The person also needs the authority to stop or change an AI-driven outcome and the time to do it, and nearly three quarters of boards were perceived in a 2026 survey to have only moderate or limited AI expertise.

Two separate indicators of board AI oversight

Different populations and measures; no combined gap calculated

~8%

disclose board-level AI oversight

of more than 3,000 US-listed companies

ISS Governance analysis, as reported by PRI, 2026

nearly

3 in 4

boards perceived to have only

moderate or limited AI expertise

KPMG survey, as reported by KPMG and INSEAD, 2026

Sources: Abrams, PRI (2026), reporting ISS Governance; KPMG International and INSEAD (2026).

Takeaway. Board and senior-management accountability cannot be delegated, automated or reduced to a signature. It is discharged only when the accountable executive has the competence, authority and time to oversee the model, the board can challenge management, and the oversight is documented rather than assumed.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST02

RECOMMENDED ACTIONS

1. **Map primary AI oversight explicitly** across the board, audit committee and any board risk committee, writing down who owns what.
2. **Map each material AI use case to a senior executive** and a named control owner, with a named person authorized to suspend the workflow from first use; delegating to a committee or vendor moves nothing.
3. **Equip the accountable body to exercise the role**: an approved AI strategy, a structured oversight question set, and the authority and time to act.
4. **Weigh inaction risk alongside implementation risk**, adding AI outcomes to executive review and later to compensation.
5. **Document board-level risk acceptance**, sized to materiality, for material or opaque AI, including third-party models, and disclose board oversight of AI.
6. **Refresh board and senior-manager AI briefings** and re-map oversight whenever the AI estate or the regulation materially changes.
7. **Make AI a board lens** wherever it is material, within the board's strategy, talent and risk discussions, rather than a separate AI report at every meeting.

SEQUENCE AND DEPENDENCIES

1. The sequence is to locate accountability, then map each use case to an accountable executive, then build the competence to exercise it, then evidence and disclose it.
2. The governance model itself, the design of decision rights and the three lines of defense, is P05-ST01, and this article assumes it.
3. The binding constraint is competence rather than structure: a firm can redraw its committees in an afternoon.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the AI they deploy, escalate issues promptly, and give the accountable executive the information needed to oversee it.
Second Line	Operate the accountability regime: map each material AI use case to the accountable executive's remit and keep the responsibilities map current, as AMAT recommends.
Third Line	Independently assess whether accountability is real, reporting to the board or audit committee on whether named owners can and do exercise oversight.
Board and Executive	Hold primary accountability, personal where a regime such as SM&CR applies; map oversight across the board and its committees, build the competence to exercise it.
Cross-Functional	Tie executive review and compensation to AI outcomes so accountability is shared and felt across the leadership team, not isolated in one role.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST02

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Accountability without competence**	Require sufficient knowledge before assigning accountability; equip the accountable body with a management-body-approved strategy, a structured question set.
HIGH **Delegation illusion**	Map each material use case to an accountable executive and a named control owner; keep the AI strategy owned by the management body and keep in-house competence.
HIGH **Oversight assumed, not evidenced**	Document board risk acceptance and disclose oversight; treat the record as evidence that oversight happened.
HIGH **Board AI-expertise gap**	Build board competence, refresh briefings and review the record of challenge, and re-map oversight regularly; route education through P06-ST10.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Accountability coverage	% of material AI use cases mapped to an accountable executive and a named control owner	AI incidents with no clear accountable owner
Oversight competence	% of board members and accountable persons completing a structured AI briefing in the past year	material AI items where board minutes show no substantive challenge.
Oversight mapping	AI oversight responsibilities mapped across board and committees	material AI risk that fell through a gap between bodies
Evidenced oversight	% of board decisions on material AI with documented risk acceptance resting on an impact assessment	material AI deployed without documented board risk acceptance

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Accountability is structurally easy to locate because it concentrates in a few named leaders; the exposure is that those same leaders must be competent to oversee mostly-vendor AI.

MID-SIZED

Capacity squeeze

The AI estate grows faster than its oversight discipline, ownership fragments, and accountability drifts onto paper.

LARGE AND GLOBAL

Advantaged

The large firm is advantaged in capability, with multi-committee oversight and the resource to build genuine competence, but it carries the heaviest disclosure and duty-of-care surface.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST02

P05-ST03

AI Policies, Standards and Acceptable Use: From a Document on the Shelf to a Control That Holds

For: Risk and Compliance; Senior Leadership

An AI policy is worth only what it prevents and what it safely enables. In a 2026 survey of 107 private equity, hedge fund and venture capital respondents, 86% reported a formal AI policy. That is self-reported adoption, not evidence that the controls work, but on it a firm with no written policy is in a small minority. A written policy is routinely mistaken for the control itself: a generic document filed to satisfy a due-diligence questionnaire governs nothing.

86%
of surveyed private fund respondents reported a formal AI policy covering research, optimization and trading (2026, self-reported)

roughly half
of surveyed respondents below $1 billion in assets reported no restrictions on generative-AI tool use (2025)

WHERE FIRMS GO WRONG

The most common failure is to treat the policy as the control. A generic or AI-drafted policy that does not reference the firm's own AI environment governs behavior no one actually performs. Shadow AI is the live exposure: in a 2025 AIMA survey roughly half of respondents below $1 billion reported no restrictions on generative-AI tools, and in a cross-industry 2025 EY survey only about a third of companies prohibited employee-built agents outright. Where policies exist they often rely on reminders and after-the-fact review rather than controls that prevent the prohibited action.

A policy on the shelf vs a policy that governs

Dimension	On the shelf	That governs
Origin	Generic, imported or AI-drafted	Built from a real inventory of AI use
Tools	No approved / prohibited list	Approved-tools list; prohibited uses
Enforcement	Reminders and after-the-fact review	Enforced technically at runtime
Currency	Reviewed rarely, if ever	Reviewed on a set cadence

AMAT framework, drawing on CBIZ (2026), FS-ISAC (2023) and APRA (2026).

Takeaway. An AI policy earns its place only when it governs what people actually do. Build it from a real inventory, calibrate acceptable use to risk, publish what is approved and prohibited, enforce it technically rather than by reminder, and keep it current as the tools move.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST03

RECOMMENDED ACTIONS

1. **Build the policy from a real inventory of how AI is used**, not a generic imported policy.
2. **Define acceptable use as a risk-calibrated, two-speed framework**: permissive for low-risk uses, stringent for sensitive data and higher-risk functions.
3. **Publish an approved-tools list** based on a vendor-compliance assessment, including AI features inside platforms the firm already runs, with access controls.
4. **Cover the eight policy elements**, from tool inventory and approval to recordkeeping, training and scheduled review.
5. **Address AI-specific legal risks**: Gen AI IP and attribution limits, AI-generated code licensing and restricted-use content categories.
6. **Enforce technically, not by reminder**, starting by preventing unapproved tools on the highest-risk uses and offering an approved alternative.
7. **Close the shadow-AI and citizen-developer gap with real visibility**, and review the policy at least annually.

SEQUENCE AND DEPENDENCIES

1. The sequence is to inventory real use, define acceptable use against it, publish the approved-tools list, enforce technically, and review continuously; in AMAT's assessment the enforcement and review steps.
2. The governance model that owns the policy, decision rights and the three lines of defense, is P05-ST01.
3. The binding constraint is that a policy is only as good as its enforcement and its currency.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Use only approved tools within the acceptable-use policy, verify AI outputs, and exercise judgment on the edge cases the policy cannot anticipate.
Second Line	Own the policy, its minimum standards and the approved-tools list; challenge risk classification and exceptions; set the AI requirements that vendor due diligence applies (P07-ST04).
Third Line	Test the policy against actual AI use, checking that approved-only is real and that citizen-developer and public-tool use is visible and governed.
Board and Executive	Approve the policy and the risk appetite behind its acceptable-use tiers, hold accountability for it, personal where a regime such as SM&CR applies.
Cross-Functional	Run the access, device, logging and blocking controls that apply the policy and, where the platforms support it.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST03

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Shadow AI**	Govern public-tool and citizen-developer use with real visibility and technical blocks, not a stated ban alone.
HIGH **Policy not enforced**	Enforce technically: block unapproved tools and separate policy from code so rules are applied at runtime.
HIGH **Generic or imported policy**	Build the policy from a real inventory of the firm's AI use; define firm-specific approved and prohibited uses.
HIGH **Unapproved data or output exposure**	Restrict permitted inputs by data classification and tenancy; validate external-facing output before release, log exceptions and use the incident escalation procedure (P05-ST12).

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Data exposure prevention	share of in-scope AI workflows with tested input-data restrictions	confirmed confidential-data exposures through AI per quarter
Unapproved-use detection	share of managed devices and enterprise AI tenants covered by usable detection logs	unapproved tools and ungoverned employee-built agents found, and time from discovery to containment
Output validation	share of client- and regulator-facing AI workflows with a designated reviewer empowered to stop release	errors caught before versus after client or regulatory release
Policy exception effectiveness	share of active exceptions with an accountable approver, an expiry and a compensating control	expired or breached exceptions, and repeat incidents

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

In AMAT's reading, the boutique is the most exposed tier despite the simplest estate, because it has the least capacity to enforce a policy.

MID-SIZED

Capacity squeeze

In AMAT's framing, the mid-sized firm's AI estate and its citizen developers outgrow a policy that was written once and never enforced.

LARGE AND GLOBAL

The large firm is advantaged in resource, with the capacity to run the two-speed permissive-and-stringent policy and centralized enforcement that the acceptable-use framework describes.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST03

P05-ST04

Model Risk Management for AI: Validating Systems That Do Not Stay Correct

For: Risk and Compliance; Technology

A validated AI model does not stay valid. Established model risk management frameworks are technology-neutral and expected to extend to AI, though most were written for banks and are benchmarks rather than rules for most asset managers. AI strains them: models are opaque, can hallucinate and drift, and when autonomous can change behavior between review cycles, so validation has to become continuous. Firms that extend their existing model risk management to the failure modes classical models never had, and scale validation intensity to each model's risk, will deploy AI safely. Firms that treat AI as exempt, or as identical to a classical model, will not.

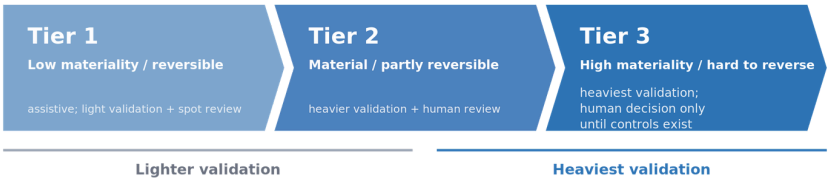
74%

of respondents in a cross-industry 2026 survey of about 500 organizations cite inaccuracy as a highly relevant AI risk, yet active mitigation lags awareness across nearly every risk category

WHERE FIRMS GO WRONG

The failure model risk management is meant to prevent is a gap between knowing and doing: in a cross-industry 2026 survey of about 500 organizations, 74% identified inaccuracy as a highly relevant AI risk, yet mitigation lagged awareness across nearly every risk category. Two opposite mistakes widen the gap. One treats AI as exempt from model risk, so a model can reach production with no inventory entry, owner or risk assessment. The other applies classical model risk management unchanged, although generative models are nondeterministic and a deployed model is a living asset that can degrade silently.

Match validation depth to model risk: materiality tiers



AMAT framework, drawing on SR 26-2 (materiality-scaled validation) and Wells Fargo GenAI MRM (2025).

Takeaway. AI does not escape model risk management; it strains it. Do not exempt AI or apply the classical template unchanged: risk-rate and tier every model, validate generative models for their specific failure modes, and revalidate at a cadence set by autonomy and drift.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST04

RECOMMENDED ACTIONS

1. **Bring AI models and AI-enabled vendor functions inside the existing risk-governance process**; run no parallel AI-model process and let no model reach production ungoverned.
2. **Risk-rate and tier each model before deployment** by business reliance, impact, complexity and autonomy, separating decisioning models from assistive tools.
3. **Extend validation for generative models**, starting with outcome analysis and hallucination measurement on the decision-relevant ones.
4. **Match validation depth to materiality**: light for assistive and reversible uses, heavier with human review for material ones, heaviest for high-materiality, hard-to-reverse uses, with human decision only until controls exist.
5. **Log every model, prompt, data or vendor-version change**, revalidate material changes in proportion, and set monitoring frequency by autonomy, reversibility and data non-stationarity.
6. **Test for the failure modes classical validation misses**: hallucination, sensitivity to irrelevant context, drift and use beyond intended purpose.
7. **Maintain a model inventory with a named owner and risk tier** for every model; none in production without an owner and a documented risk assessment.

SEQUENCE AND DEPENDENCIES

1. The sequence is tier, then validate, then monitor, and it depends on an honest model inventory.
2. Independent challenge and validation as a distinct second-line function is P07-ST05; pre-deployment testing is P07-ST01.
3. In practice the constraint is often validation capacity and skill as much as policy: GenAI validation requires people who can design conceptual-soundness and outcome-analysis tests and calibrate hallucination metrics.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own model risk-rating, documentation and first-line testing; state each model's intended use and escalate when it is used beyond it.
Second Line	Independently validate higher-risk AI models with conceptual-soundness and outcome-analysis testing, and set or challenge the AI MRM policy and tiering criteria under the firm's governance arrangements.
Third Line	Assure that the AI MRM framework actually operates; in the AMAT three-lines model, this means testing that validation, tiering and the inventory work in practice.
Board and Executive	Approve model-risk appetite and tiering thresholds and hold accountability for them, personal where a regime such as SM&CR applies; ensure known model vulnerabilities are cured.
Cross-Functional	Classify decisioning versus assistive models and route work by risk tier.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST04

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Classical MRM blind to AI failure modes**	Extend MRM for AI-specific complexity with conceptual-soundness and outcome-analysis testing, not a classical validation template.
HIGH **Undetected drift**	Set validation frequency by autonomy, reversibility and data non-stationarity; log and triage every model, prompt, or data change, and revalidate material changes, including cumulative recalibrations.
HIGH **Sensitivity to irrelevant context**	Test representative investment prompts with irrelevant contextual variations and check whether decisions change without an economic reason.
HIGH **Model used beyond purpose / shadow AI**	Constrain models to their risk-assessed scope and inventory every model with an owner; block deployment that bypasses validation.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Model coverage	% of AI models risk-rated and in the inventory	% of production models with no completed validation
GenAI validation depth	conceptual-soundness and outcome-analysis tests run per model	measured hallucination rate in production
Validation cadence	monitoring frequency set by autonomy and data non-stationarity	drift incidents caught before client impact
Bias and robustness	disparate-impact and input-perturbation tests run	biased or unstable outputs reaching decisions

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The boutique advantage is proportionality: applying the same risk-based principle, a small firm running a handful of assistive models can meet the bar with risk-rating, vendor validation evidence.

MID-SIZED

Capacity squeeze

In AMAT's reading, the capacity squeeze applies the awareness-mitigation gap McKinsey's cross-industry survey describes to firms that run consequential models without yet having a validation function independent of the people.

LARGE AND GLOBAL

The large-firm move is to run the full GenAI-adapted validation lifecycle through an independent function and then spend its effort on the agentic problem.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST04

P05-ST05

AI and the Regulators: Meeting Supervisory Expectations Without Waiting for AI-Specific Rules

For: Risk and Compliance; Legal; Senior Leadership

Asset managers do not get to decide whether their use of AI is regulated. Supervisors in major jurisdictions treat existing, technology-neutral rules as applying to AI, the EU has added an AI-specific statute that can reach firms outside the EU, and the SEC has already brought AI-related enforcement actions. This article sets out how the EU AI Act maps onto frameworks the firm already runs, how the US, UK, EU and Asian regimes differ in instrument and posture, and what the live enforcement and supervisory-priority signals are. The winning posture is not to wait for bespoke AI rules but to map today's obligations onto existing governance and be ready to explain the firm's AI to a supervisor.

7%

of global annual turnover, or 35 million euros, is the maximum EU AI Act penalty for the most serious breaches (prohibited practices)

four

SEC enforcement actions against registrants for misrepresenting their AI use ("AI washing") within a single year to early 2025

WHERE FIRMS GO WRONG

The most common failure is a category error: waiting for AI-specific rules that, in most of the world, are not arriving in that form, since existing agencies engage first through current rulemaking and enforcement and the EU is the outlier. The second is misjudging the firm's own status, because under the EU AI Act a firm that substantially modifies or fine-tunes a general-purpose model may cross into provider obligations. The third is a gap rather than a rule: MiFID II contains no provisions written for AI.

Mapping EU AI Act obligations onto rules you already run (EBA)



Source: AMAT adaptation of the European Banking Authority banking-and-payments mapping (2025).

Takeaway. An asset manager does not choose whether its AI is regulated, only whether it is ready. Map AI to the obligations already owed, classify the firm's status under the regimes that apply, build the controls that need building, and engage supervisors so that engagement becomes assurance rather than exposure.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST05

RECOMMENDED ACTIONS

1. **Map every AI use to the existing obligations it already triggers**, such as conduct, suitability and recordkeeping, rather than wait for an AI rulebook.
2. **Classify each system's EU AI Act status** as provider or deployer, and its risk tier, before deployment, reassessing on any fine-tuning.
3. **Map AI Act obligations against existing frameworks** such as DORA, MiFID and AIFMD.
4. **Treat supervisory scrutiny as active now**, preparing for named examinations with documented AI governance.
5. **Substantiate every external AI claim** in marketing, RFPs, fund documents and filings to eliminate AI-washing exposure.
6. **Build a regulator-communication capability**, starting with a named owner and a log of every supervisor contact on AI.
7. **Build governance to durable principles** and track jurisdictional divergence for cross-border operations.

SEQUENCE AND DEPENDENCIES

1. The sequence is inventory, then classify, then map, then build the gaps, then engage: a firm cannot classify its AI Act status or right-size compliance without first knowing which AI.
2. The EU AI Act was amended before it was fully in force: the Digital Omnibus on AI deferred the high-risk dates and softened the literacy duty.
3. A practical challenge is bridging the two regimes' different logics: the AI Act is process-centric and human-centric while MiFID II is outcome-driven and technology-neutral.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the AI inventory and the first-line mapping of each use to the obligations it triggers.
Second Line	Own the AI Act mapping, the provider/deployer classification and risk-tiering, and the regulator-engagement process; integrate AI Act obligations into existing governance rather than a parallel regime.
Third Line	Assure that the obligations map, the status classification and the regulator-communication process actually operate, and horizon-scan for ESMA and legislative developments on AI in investment services.
Board and Executive	Hold accountability for the firm's AI regulatory compliance, personal where a regime such as SM&CR applies.
Cross-Functional	Run the regulator-communication capability: a named owner, supervisor-contact log and engagement decision rule, AI Office code engagement, voluntary sandbox or live-testing routes where offered.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST05

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **AI-washing enforcement**	Substantiate every external AI claim with an evidence file; review all AI language in client-facing and regulatory materials.
HIGH **Provider/deployer misclassification**	Classify provider/deployer status and risk tier before deployment; re-run on any new build, material modification, rebranding or repurposing.
HIGH **Unpreparedness for active supervision**	Treat examination as active now: maintain documented AI governance mapped to regulated activities, ahead of any new rule.
MEDIUM **Overlapping oversight and compliance burden**	Integrate AI Act obligations into existing governance via the derogation/integration/standalone map rather than a parallel regime.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Obligations mapping	% of AI systems mapped to the existing obligations they trigger and to their AI Act risk tier	% of production AI with no completed obligations map
Status classification	% of systems with a documented provider/deployer determination	provider obligations triggered without prior classification
AI-claim substantiation	% of external AI claims with an evidence file	AI-washing findings, complaints or enforcement contact
Supervisory engagement	sandbox / live-test participations and engagements logged	adverse examination or supervisory findings on AI

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The boutique advantage is footprint: a small firm that mostly deploys vendor AI carries the lighter deployer obligations under the EU AI Act, not none.

MID-SIZED

Capacity squeeze

The capacity squeeze is the awareness-obligation gap in its regulatory form: the firm has begun building AI in-house, which risks pulling it into the heavier provider obligations.

LARGE AND GLOBAL

The large-firm position is advantaged but not easy: it has the dedicated legal and compliance capacity to engage supervisors directly, through formal channels and, where offered, sandboxes.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST05

AMAT

P05-ST06

Conduct Risk, Fiduciary Duty and Client Outcomes: When the Algorithm Advises, the Firm Still Answers

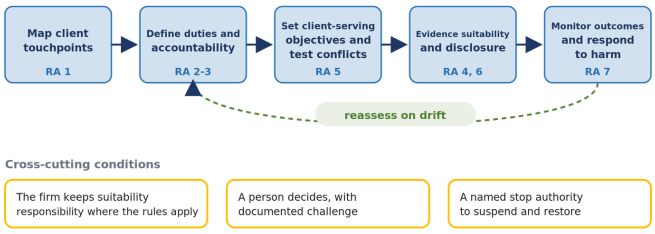
For: Risk and Compliance; Legal; Investment Teams

Conduct obligations and fiduciary duty do not change when a model does the analysis; they become harder to discharge, and which duty applies depends on the activity. Existing conduct rules apply to AI as they apply to any other tool, and supervisors are checking whether firms can show they meet them. Suitability and best-interests obligations were written for human advisers, so when a model supports a recommendation the firm owes the same duty, but traceability and a demonstrable basis for the advice are exactly what black-box systems struggle to provide. AI turns conduct into a design constraint: a firm that cannot explain, evidence and stand behind an AI-shaped recommendation is unlikely to meet its duty, however good the output looks.

WHERE FIRMS GO WRONG

The first failure is letting the model drift from decision-support into decision-maker: the duty attaches to the decision, not the tool, and an override that is available but never exercised leaves the reviewer a rubber stamp. The second is a conflict of interest hidden inside model complexity, for example a model that could favor the firm's own funds, a higher-fee share class or a product the firm is seeding. A wider form is asymmetry, where AI-driven pricing or advice can work against clients without any single decision looking conflicted.

Designing conduct in: the AI client-outcome control sequence



AMAT framework, drawing on Ghetti et al. (2025) and Tracz (2026).

Takeaway. Conduct and fiduciary duty are the reason the other controls in this pillar exist, and AI raises the bar for evidencing them. Keep AI as decision-support, make suitability reconstructable, set client-serving objectives, disclose accurately and monitor real outcomes, because when the algorithm advises, the firm still answers.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST06

RECOMMENDED ACTIONS

1. **Map where AI touches client outcomes**, starting with the client-facing workflows already in use, including AI features embedded in existing platforms, and building toward every system that profiles clients, shapes suitability or communicates with them.
2. **Apply the three fiduciary duties to every AI-assisted mandate**: awareness that AI is in use, a defined scope and terms of delegation, and review in periodic performance reviews.
3. **Keep AI as decision-support**, with a named person owning each AI-shaped recommendation and a documented, reconstructable basis.
4. **Preserve suitability and traceability**, treating the firm, not the vendor, as remaining responsible for suitability where the rules apply.
5. **Set client-serving objectives and surface conflicts**, with conduct review in the business case for any AI-driven pricing or advice change.
6. **Disclose AI use accurately and avoid AI-washing**: clear disclosure of AI-related risks and marketing claims that match actual use.
7. **Monitor outcomes for fair value and vulnerable clients**, starting with the AI-enabled client journeys that could affect suitability, fees, access to advice or client understanding, and act on the findings.

SEQUENCE AND DEPENDENCIES

1. The sequence is to map client-outcome touchpoints first, then define duties and accountability, then set client-serving objectives and test conflicts, then evidence suitability and disclosure.
2. Bias as an outcome is a conduct concern and appears here as fair treatment and discrimination risk.
3. The binding constraint is explainability: the duties assume a recommendation whose basis can be shown from a decision record.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the AI-shaped recommendation and its basis, keep AI as decision-support, and confirm the firm retains suitability responsibility.
Second Line	Set and test client-serving objectives, run outcome and fair-value monitoring, and reconcile AI marketing claims to actual use.
Third Line	Independently test whether AI-assisted recommendations are reconstructable and whether conduct controls operate as designed, including conflicts and disclosure.
Board and Executive	Confirm that fiduciary duty and the best-interests test are being met in AI-enabled services, and hold accountability for conduct outcomes.
Cross-Functional	Align the client-outcomes map, disclosure, and vendor oversight so responsibility for AI-shaped advice is clear end to end.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST06

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Over-reliance on AI**	Keep AI as decision-support with a named owner and a documented basis for every recommendation.
HIGH **Suitability breach**	Require reconstructable suitability rationale; constrain high-autonomy uses that cannot be explained.
HIGH **Adviser conflicts of interest**	Set client-serving objectives and test for proprietary-product steering and engagement optimization.
HIGH **Liability-attribution gap**	Treat suitability responsibility as non-delegable; keep a clear line of responsibility to the firm and its named owners, and route client complaints through the firm's redress process.

7 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Suitability defensibility	% of AI-assisted recommendations with a reconstructable, documented basis	unsuitable or corrected AI-assisted recommendations per 1,000, compared with the pre-AI process.
Human ownership and challenge	% of client-facing AI outputs with a named accountable owner; rate at which reviewers edit.	AI outputs reaching clients with no owner
Conflict and objective integrity	% of client-facing models with a documented client-serving objective and conflict test, including masked-product comparisons	instances of proprietary-product steering or engagement optimization found by conflict testing, and the share remediated
Disclosure accuracy	% of AI marketing and mandate claims reconciled to actual AI use	AI-washing findings, and disclosures corrected after client or second-line challenge

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

The client base is small enough that conduct can be genuinely personal and well documented, so the discipline is straightforward.

MID-SIZED

Capacity squeeze

AI-enabled advice and distribution scale faster than the firm's conduct controls, and advice-like outputs and embedded conflicts can spread before monitoring catches them.

LARGE AND GLOBAL

The large firm has the resources to evidence suitability, disclosure and fair value at scale but faces multi-jurisdiction conduct rules and high AI-washing visibility.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST06

P05-ST07

Data Protection and Privacy: Governing the Personal Data Behind the Model

For: Legal; Risk and Compliance; Technology

AI does not come with a privacy exemption. The European Data Protection Board takes the position that the GDPR regularly applies to AI models trained on personal data. Separately, a firm should assess whether each AI system processes personal data at every stage, including prompts, retrieved records, outputs and logs, whether or not it trains a model. It must be able to show a lawful basis, evidence any anonymity claim and document its accountability. Where training was unlawful, a supervisory authority can in some cases order erasure of the data or the model itself.

74%
of industry respondents rate data privacy among the top AI risks, as do 65% of vendors and 80% of regulators (2026 global survey)

46%
of Luxembourg supervised entities cite data protection as their second most cited AI challenge (2024)

WHERE FIRMS GO WRONG

The first failure is assuming the model is anonymous and treating privacy as an architecture detail to handle later; a firm cannot make an anonymity claim for a model it has not assessed, and for a purchased tool the entry point is mapping the personal data in inputs, retrieved records, outputs and logs. The second is data walking out through prompts, since employees can enter customer personal information and proprietary data into generative AI tools, with the greatest exposure in public and consumer-grade tools. The most acute version is confidentiality and legal privilege, which a US court has held can be forfeited by using a public chatbot.

EDPB model classification: explicit vs implicit

Dimension	Explicit data-providing	Implicit data-embedding
Design intent	makes personal data available	embeds data in parameters
GDPR scope	always in scope	in scope; assess case by case
Anonymity	cannot be anonymous	anonymous only if evidenced
What to do	treat as personal data	run the anonymity test

Source: Apostle and Schröder, Orrick (2025), summarizing EDPB Opinion 28/2024.

Takeaway. Assess scope, test anonymity, establish a lawful basis, document accountability, and govern the data going into and coming out of AI. The firms that get this right can show their work for each model, because a firm that cannot evidence lawful processing cannot defend it.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST07

RECOMMENDED ACTIONS

1. **Assess data-protection scope for every AI model that touches personal data**, including AI features inside existing platforms, and do not assume a model is out of scope.
2. **Classify each model trained on personal data as explicit data-providing or implicit data-embedding**, run the EDPB three-element anonymity test before treating any output as anonymous, and map the data flows for purchased tools.
3. **Establish and document a lawful basis**, completing the three-step legitimate-interest assessment where it is relied on.
4. **Maintain accountability documentation**: processing records, decisions and DPIAs before high-risk production use or any pilot on live personal data.
5. **Apply mitigating measures beyond the minimum**, such as pseudonymization and masking, and enforce purpose-based data minimization on queries.
6. **Govern third-party AI data**, confirming whether vendors train on firm data, where they store it and who else sees it.
7. **Prohibit confidential or privileged data in public consumer AI tools**.

SEQUENCE AND DEPENDENCIES

1. The sequence is to assess scope, then classify and test anonymity, then establish and document a lawful basis, then run the accountability documentation continuously.
2. Decide early how the firm will get any special-category data it needs to test AI for bias.
3. This article is the regulatory and obligation frame; the data-architecture controls that implement it, information-barrier design, encryption, and privacy-enhancing technologies such as differential privacy and federated learning, are P03-ST08.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Keep personal and confidential data out of public tools, and out of approved tools unless the purpose needs it.
Second Line	Set the data-protection assessment standard and the AI-specific vendor due-diligence approach; advise on and challenge scope, anonymity and lawful-basis assessments, DPIAs and accountability documentation.
Third Line	Independently test whether anonymity and lawful-basis claims are evidenced and whether shadow AI and query controls operate as designed.
Board and Executive	Confirm the firm can evidence lawful, anonymized and accountable AI processing, and hold accountability for data-protection risk, personal where a regime such as SM&CR applies.
Cross-Functional	Align the data-protection obligation with the P03-ST08 architecture controls, the P05-ST11 cross-border rules and the regional regimes in P05-ST05 so obligation and mechanism meet.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST07

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Model inversion / inference**	Run the anonymity test and evidence it; apply privacy-enhancing controls (architecture in P03-ST08).
HIGH **Unlawful training data**	Establish a lawful basis and keep accountability documentation before training.
HIGH **PII leakage via prompts**	Enforce purpose-based data minimization on queries and bound acceptable use (P05-ST03).
HIGH **Query retention by external AI**	Restrict to approved tools and minimize data before submission; confirm retention and training terms.

5 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Scope and lawful basis	% of AI models with a completed data-protection scope assessment and documented lawful basis	models processing personal data with no assessed basis
Anonymity evidence	% of models claiming anonymity with a completed three-element test on file	anonymity claims unsupported by documentation
Query and tool control	% of AI use through approved tools with data minimization enforced	PII-in-prompt or shadow-AI incidents detected
Third-party data governance	% of AI vendors handling personal data with verified data-use terms (training.	vendors using firm data to train without confirmation

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

A boutique runs on vendor AI it cannot fully audit, so its priority is contractual: confirm in writing whether vendors train on the firm's or clients' data.

MID-SIZED

Capacity squeeze

The problem is scale of use, where growing generative-AI adoption spreads PII-in-prompt leakage and shadow AI faster than controls follow.

LARGE AND GLOBAL

The large firm has the resources for DPIAs and evidenced anonymity but faces the sharpest version of the training risk.

P05-ST08

Human Oversight and Escalation Protocols: Keeping a Person Able to Overturn the Machine

For: Risk and Compliance; Investment Teams; Operations

Human oversight is the control that keeps a person answerable for what an AI system does, and it fails in two opposite ways. Too little leaves high-stakes or irreversible decisions delegated to a model. Too much of the wrong kind leaves a human formally in the loop who adds no scrutiny, which is no better than none. Oversight has become urgent because AI has started to act, not just answer: in a self-reported, cross-industry 2026 survey of about 500 organizations, 44% flagged autonomous or unintended system actions as a relevant risk while only 29% were actively mitigating it. The design goal is oversight calibrated to risk, backed by a real override, an explicit escalation path and an audit trail that proves a human could and did intervene.

44%
of surveyed organizations across industries flag autonomous or unintended system actions as a relevant AI risk (2026, self-reported)

29%
are actively mitigating that risk, a 15-percentage-point gap

WHERE FIRMS GO WRONG

The first failure is automation bias: rising workloads can cut review time, confident-looking output can breed undue trust and manual skill can degrade, leaving a human formally in the loop but functionally absent. These are risk factors, not proven independent causes. The second is treating the presence of a human as the control, when the control is the human's ability and willingness to overturn the machine. The third is applying one oversight setting everywhere, which makes review so heavy on routine tasks that reviewers stop engaging while high-stakes decisions are no better protected.

Human oversight modes, calibrated to materiality



AMAT framework, drawing on IOSCO (2025) and Gaurav et al. (2025).

Takeaway. Human oversight only counts when a person can and does overturn the machine. Calibrate oversight to what is at stake, set explicit boundaries and escalation thresholds, keep an unconditional failsafe on decisions that cannot be undone, and test that a person can halt, reject or reverse an action, investigating a challenge rate that stays near zero.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST08

RECOMMENDED ACTIONS

1. **Calibrate oversight to the materiality and reversibility of each use case**, with human approval for high-stakes or irreversible decisions.
2. **Define explicit autonomy boundaries and escalation thresholds** for every AI system at deployment, with a named responder, a response deadline and a safe default.
3. **Maintain an unconditional human failsafe** for the highest cost-of-error decisions, such as large trading positions, final NAV approval and Level 3 valuations.
4. **Design against automation bias**: give reviewers the time, authority and ability to overturn outputs.
5. **Add the agentic control set for systems that act**, starting with a hard human-approval gate before irreversible or high-impact actions.
6. **Monitor the oversight itself** through a trended challenge rate, investigated whenever it falls to or near zero.
7. **Record and evidence the oversight**, with a named owner for each AI-assisted workflow and the actual approver of each gated material decision.

SEQUENCE AND DEPENDENCIES

1. The sequence is to classify each use case by materiality and reversibility, assign an oversight mode, set the autonomy boundaries and escalation thresholds.
2. This article takes the control lens on human oversight; the cultural and capability dimension, how a firm builds reviewers with the skill and the confidence to challenge a model.
3. The binding constraint is that effective oversight competes for the scarcest resource in the firm, expert attention.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the live workflow with a named on-call owner: exercise the oversight (review, challenge and, where warranted, overturn AI outputs at the defined points.
Second Line	Design the oversight framework: assign oversight modes by materiality, set autonomy boundaries and escalation thresholds, monitor the challenge rate for signs of rubber-stamping.
Third Line	Test whether oversight is real, not nominal: check that overrides are exercised and that approval, override and rollback authority is logged and reconstructable.
Board and Executive	Ensure the highest cost-of-error decisions retain an unconditional human failsafe, and hold accountability for AI-assisted decisions, personal where a regime such as SM&CR applies (see P05-ST02).
Cross-Functional	For systems that act, run the goal-alignment, guardrail and runtime-approval controls that keep humans able to intervene at speed.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST08

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Automation bias / over-reliance**	Design against it: give reviewers time, authority and an alternative to overturn outputs; embed domain experts across the lifecycle.
HIGH **Agentic action outside bounds**	Bound the autonomy and gate consequential actions: goal alignment, guardrails, runtime intent validation and human approval before irreversible actions.
HIGH **Irreversible decision delegated to AI**	Maintain an unconditional human failsafe for the highest-stakes decisions, regardless of AI accuracy (for algorithmic trading, the required trading limits, testing.
MEDIUM **Uncalibrated oversight**	Calibrate the oversight mode to the materiality and reversibility of each use case.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Oversight calibration	% of AI use cases with an oversight mode assigned by materiality and reversibility	high-stakes or irreversible decisions executed without the required human approval, per 1,000 gated decisions
Challenge rate	rate at which reviewers edit, override or reject AI output, trended	incidents traced to accepted but incorrect AI outputs
Autonomy-boundary coverage	% of AI systems with documented autonomy boundaries and escalation thresholds	autonomous actions taken outside defined bounds, per 1,000 eligible actions
Failsafe coverage	% of high cost-of-error decisions with an unconditional human failsafe	irreversible or high-impact AI actions taken with no failsafe, and failed stop or rollback attempts per stop test

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

The boutique is well placed for the human element and poorly placed for the technical one: with few decisions and senior reviewers close to them, meaningful oversight is natural.

MID-SIZED

Capacity squeeze

Automation scales faster than the reviewers who can challenge it, and oversight quietly degrades into rubber-stamping.

LARGE AND GLOBAL

The large firm has the resource to engineer bounded autonomy and automatic escalation thresholds.



Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST08

P05-ST09

Explainability, Transparency and Recordkeeping: Making an AI Decision You Can Reconstruct

For: Risk and Compliance; Technology; Internal Audit

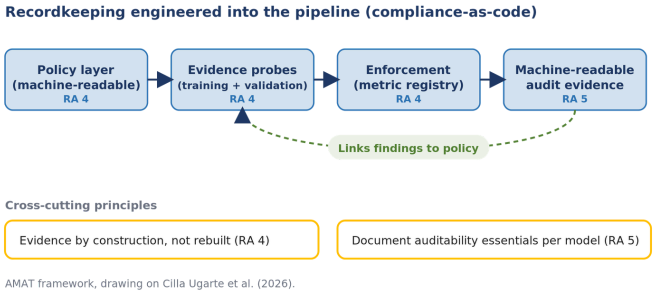
An AI decision a firm cannot explain or reconstruct is a decision it cannot defend. Explainability, transparency and recordkeeping are three connected controls with one practical test: can the firm show after the fact what informed an AI-assisted decision, to the standard the audience and the regulator require, and prove it from a record rather than from memory. Some supervisory authorities have identified limited explainability as a challenge for model validation, audit trails and accountability, and disclosure expectations are rising in parallel. In a regulated firm, a model that cannot be explained is harder to validate and audit, and a control whose evidence is reconstructed after the fact is one the firm cannot readily prove it ran.

72%

of S&P 500 companies disclosed AI-related risks in their 2023 annual-report risk factors (359 firms)

WHERE FIRMS GO WRONG

The failure begins with treating explainability as something to bolt on after a model is chosen, when the real decision is made earlier, between an inherently interpretable model and a black box explained after the fact. The second is over-trusting explanation tools, whose outputs can be technically valid yet hard for the recipient to interpret correctly. The third is that the record is not there: one research proposal argues that compliance is still commonly treated as a periodic, manual exercise, with evidence assembled by hand when an examination looms.



AMAT framework, drawing on Cilla Ugarte et al. (2026).

Takeaway. Explainability, transparency and recordkeeping are three connected controls with one test: the firm can show what informed an AI-assisted decision and prove it from a record. Decide explainability at model selection, calibrate explanations to the people who need them, and engineer compliance evidence into the pipeline so it exists by construction rather than by reconstruction.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST09

RECOMMENDED ACTIONS

1. **Choose the explanation approach at model selection**, preferring inherently interpretable models for high-materiality decisions.
2. **Calibrate explanations to the audience and the decision's materiality**, starting from the stakeholder's need.
3. **Where the firm builds or runs its own ML pipeline, engineer recordkeeping into it**, so compliance evidence is machine-readable and produced across the model estate.
4. **Decide before switch-on which AI outputs are records** under existing books-and-records rules, where they are kept and for how long.
5. **Document the auditability essentials** in model cards and datasheets: provenance, features, intended use, selection rationale and drift history.
6. **Disclose AI risks and deployment consistently** and substantiate every external AI claim across filings, RFPs and marketing.
7. **Treat unexplainability as a control gate** for high-materiality decisions, with supplementary explainability, a documented limitation and sign-off by the accountable owner and second line.
8. **State the assurance limit** where a model can only be tested by sampling, with the population, sample size and acceptance criteria set in advance.

SEQUENCE AND DEPENDENCIES

1. The sequence is choose-interpretable-where-you-can, then instrument the pipeline for evidence, then document the essentials, then disclose.
2. Model validation as a discipline is P05-ST04, and this article assumes rather than repeats it; the explainability and recordkeeping here are what validation and monitoring consume.
3. It depends on an honest reckoning with the accuracy-explainability tension: more accurate models are often less interpretable, though how large the trade-off is depends on the task.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Produce the explanation and the auditability documentation at build time, choosing the explanation approach at model selection and maintaining model cards.
Second Line	Set the explainability standard by materiality and audience, own the AI-disclosure and AI-washing controls.
Third Line	Assess whether models are genuinely auditable, insist on the documentation set, and where only stochastic sampling is feasible.
Board and Executive	Oversee the firm's AI disclosures and representations, scrutinize AI-specific management representations, and ensure AI claims are accurate and supported by the record.
Cross-Functional	Where the firm adopts a compliance-as-code pipeline, run it so that every control result links up to its policy and any finding is traceable end.



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST09

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Opacity impedes validation and audit**	Gate high-materiality decisions on explainability; no black box behind a material decision without supplementary explanation and a documented limitation.
HIGH **Unauditable for want of documentation**	Document the auditability essentials (data provenance, features, intended use, selection rationale, production and monitoring history) with model cards and datasheets.
HIGH **AI washing and inconsistent disclosure**	Disclose consistently and substantiate every AI claim; consider the recommended AI-definition and separate-reporting disclosures.
HIGH **Regulatory-requirement tension from opacity**	Treat explainability as a regulatory expectation, not an enhancement; build it in before deployment.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Explainability coverage	% of production models with an explanation method matched to decision materiality	high-materiality decisions made by unexplainable models
Auditability documentation	% of models with complete auditability documentation (provenance, features, intended use, selection, monitoring) and model cards	audit findings attributable to documentation gaps
Evidence-by-construction	% of compliance evidence generated automatically by the pipeline	evidence reconstructed manually before an examination
Audience calibration	% of AI use cases with explanations defined per stakeholder (client, PM, validator, regulator)	stakeholder or supervisor unable to interpret an explanation

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The obligation is mostly an ownership and vendor-management question: a firm that mostly deploys vendor AI should require proportionate explanations and the documentation its vendors can make available.

MID-SIZED

Capacity squeeze

The recordkeeping gap is at its sharpest: the firm has begun building models but its evidence discipline is ad hoc and assembled after the fact, so the priority.

LARGE AND GLOBAL

The large firm is advantaged but carries the heaviest surface: it may have the capability to pilot mechanistic interpretability and compliance-as-code on models it owns.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST09

P05-ST10

AI Inventory and Use-Case Classification: You Cannot Govern What You Have Not Counted

For: Risk and Compliance; Technology

A firm cannot govern, tier or report on AI it has not counted. The inventory is the foundation control of AI governance: a comprehensive, living register of every AI system the firm runs, including AI a vendor switches on inside a platform the firm already runs, with a classification that sorts systems by risk. Traditional IT asset lists miss most AI, and an incomplete inventory is a recurring concern in US bank examinations. The discipline is not to inventory everything with equal weight but to see everything, discover shadow and embedded AI continuously rather than by self-registration, and govern each use case in proportion to its risk.

29%
of organizations formally track where their AI adoption fails (2026 survey)

five
categories of AI routinely go unrecorded in traditional inventories (2026 vendor account)

WHERE FIRMS GO WRONG

The first failure is relying on self-registration and point-in-time lists, because the fastest-growing AI, embedded copilots and shadow experiments, is the least likely to be declared. The second is over-governing everything as if it were a model, which inflates control cost and creates bottlenecks. The third is inventorying without classifying: a list without a risk tier cannot say where to concentrate scrutiny, so the inventory and the classification are one control, not two.

A static asset list vs a living AI inventory

Dimension	Static list	Living inventory
Refresh	Point-in-time, manual	Continuously, auto-discovered
Scope	Formal projects only	Embedded, vendor, shadow, agents
Risk	No risk tier	Risk-tiered, with owner and lineage
Assurance	Blind to shadow AI	Reconciled to procurement and cloud

AMAT framework, drawing on Chen and Telang (2026), Evans, OneTrust (2026) and the MindForge Consortium (2026).

Takeaway. See everything, classify by risk, govern in proportion and keep the picture current, because a firm cannot govern, tier or defend AI it has not counted.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST10

RECOMMENDED ACTIONS

1. **Inventory every AI system with a named owner and risk tier**, including vendor-enabled features.
2. **Discover rather than rely on self-registration**, reconciling the inventory monthly against procurement, cloud spend and vendor change notices.
3. **Classify every use case by risk materiality**, rated inherent and residual, approving for production only within appetite.
4. **Tier controls to the rating**, governing at the use-case level rather than treating every tool as a model.
5. **Manage each use case across its lifecycle**, from approval to decommissioning, and keep the inventory current.
6. **Map the inventory to regulatory scope**, flagging EU AI Act high-risk systems and model-risk expectations.
7. **Put the inventory under a named control function**, with board assurance that it is current.

SEQUENCE AND DEPENDENCIES

1. The sequence is to discover and inventory first, then classify by risk, then tier controls, then govern the lifecycle, with reconciliation running continuously underneath.
2. The model-risk framework applied to the models the inventory catalogs, validation, monitoring and challenge, is P05-ST04, and this article assumes it.
3. The binding constraint is that completeness and currency are harder than the initial build: standing up an inventory is a one-off project.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Register the AI they build or adopt and keep their entries current, and flag new or shadow use so it enters the inventory rather than running.
Second Line	Own the inventory and the classification methodology, run scheduled discovery and reconciliation, and rate each use case by risk materiality.
Third Line	Independently validate the inventory against actual AI use, including shadow and embedded AI, comparing what is deployed to what the organization tracks.
Board and Executive	Assure itself that the AI inventory exists and is current, with material AI classified, as part of the board's AI oversight (see P05-ST02).
Cross-Functional	Reconcile the AI inventory with procurement, cloud, model and vendor registers, and with the DORA register of information where it applies.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST10

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Shadow and embedded AI off the register**	Discover continuously and reconcile to procurement; do not rely on self-registration.
HIGH **Incomplete or fragmented inventory**	Maintain a complete inventory under a named control function, linked to other registers.
MEDIUM **No risk classification**	Classify every use case by risk materiality before production and tier controls to the rating.
MEDIUM **Over-governance**	Govern at the use-case level, apply the tool-or-model test, and streamline lower-risk reviews; reserve deep control for higher-risk systems.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Inventory completeness	% of AI systems (incl. embedded, vendor-enabled, shadow and agents) captured in the inventory	AI systems found in use outside the inventory
Classification coverage	% of use cases with a risk-materiality rating	AI in production with no assigned risk tier
Control proportionality	alignment of control effort to risk tier (streamlined lower-risk vs deep higher-risk)	lower-risk uses stuck in heavy review, or higher-risk uses under-controlled
Inventory currency	% of entries updated on change or decommissioning within the defined window	decommissioned or changed systems still listed as live

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

The estate is small enough that a simple, well-maintained inventory can be genuinely complete, so the work is discipline rather than tooling.

MID-SIZED

Capacity squeeze

The estate and its shadow AI outgrow a manual list, and AI ends up in production off the register.

LARGE AND GLOBAL

The large firm has the resource for automated discovery and risk-materiality tiering but faces agent sprawl and thousands of use cases across entities, so its task is continuous discovery.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST10

AMAT

P05-ST11

Cross-Border AI Regulation: Governing One AI Estate Across Many Rulebooks

For: Risk and Compliance; Legal; Senior Leadership

A global asset manager runs one AI estate but answers to many rulebooks, and those rulebooks are diverging rather than converging. The EU regulates AI horizontally and prescriptively, the US relies on sectoral guidance and a shifting federal-versus-state patchwork, the UK applies its existing principles-based rules, and Asia's regimes span activity-specific expectations to horizontal statutes. Fragmentation is likely to persist, so the workable posture is a single high-bar governance program mapped to each jurisdiction's obligations, built on durable legislation rather than fragile executive action, and alert to the cases where satisfying one regime can breach another.

December 2027
is when the EU AI Act's high-risk duties apply, under the Act as amended, to any of a firm's AI systems within its high-risk categories

7%
of global turnover, or EUR 35 million, is the top EU AI Act penalty, reserved for prohibited practices

WHERE FIRMS GO WRONG

The internal failure is treating cross-border compliance as separate national checklists rather than one estate governed against many rulebooks, with material risk arising when a tool is scaled into new jurisdictions without re-testing. Identical rules are not applied identically, since most EU enforcement is delegated to member states. The subtler trap is conflict between regimes, most often between keeping records and deleting them, where books-and-records rules and data-protection limits apply to the same AI-drafted client communication.

The EU AI Act sets the cross-border baseline

2027 — from December: high-risk duties apply, under the Act as amended

EU baseline

7% — of global turnover or EUR 35m, for prohibited practices

Source: EU AI Act (Regulation (EU) 2024/1689), as amended by the Digital Omnibus on AI.

Takeaway. Do not wait for harmony; govern for divergence. Map one AI estate to every rulebook it touches and to where its data goes, run a single high-bar program set to the strictest regime, anchor compliance to durable legislation, and keep a live register of conflicts between regimes.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST11

RECOMMENDED ACTIONS

1. Maintain a jurisdiction-mapped AI inventory: start with the systems already used in more than one jurisdiction or with an EU nexus; build toward every AI system tagged with the regimes it touches, where its data is processed and stored, and the specific obligations each imposes.
2. Default the baseline to the EU AI Act where the firm has any EU nexus: start by setting that baseline for the systems with an EU nexus; build toward one high-bar governance program with thin per-jurisdiction overlays in place of separate national programs.
3. Base compliance on durable legislation, not executive orders or non-binding guidance that can be revoked; monitor the US federal-versus-state fight.
4. Screen for conflicting obligations across regimes and keep a register of cases where satisfying one rule risks breaching another, starting with record retention against data-protection limits.
5. Prepare any system in the EU AI Act's high-risk categories for the high-risk duties (technical documentation, documented human oversight) before they apply from December 2027, under the Act as amended, and reuse the package as the cross-border template.
6. Engage and support international coordination (IOSCO, FSB, ESMA, national regulators) to reduce fragmentation cost and gain early sight of change.

SEQUENCE AND DEPENDENCIES

1. The sequence is to map the estate to jurisdictions first, stand up the single high-bar program with overlays, anchor it to durable law.
2. The binding constraint is that this is a moving target: the US position in particular is being litigated.
3. Where a firm cannot yet resolve a genuine conflict between two regimes, the honest interim is to narrow the AI system's cross-border scope until it can.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Flag when an AI tool is used in a new jurisdiction or starts to process data in a new location.
Second Line	Own the jurisdiction map, the high-bar program and overlays, the conflict register, and the EU AI Act high-risk package.
Third Line	Test jurisdiction coverage and whether compliance rests on durable law rather than revocable guidance.
Board and Executive	Own the cross-border AI strategy and the choice of baseline, and confirm the firm is ready for the EU AI Act timeline.
Cross-Functional	Engage IOSCO, FSB and ESMA harmonization efforts and align the group's response across legal, compliance and technology.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST11

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Fragmentation compliance risk**	Keep a jurisdiction-mapped inventory, including where each system's data is processed and stored, and a high-bar program with per-regime overlays.
HIGH **Conflicting obligations**	Screen for and register cross-border conflicts; resolve each one case by case and keep the strictest-regime standard everywhere else.
HIGH **Building on fragile or contested law**	Base compliance on durable legislation and monitor the federal-versus-state fight.
HIGH **Extraterritorial exposure missed**	Default the baseline to the EU AI Act where any EU nexus exists; prepare in-scope systems for the high-risk duties that apply from December 2027.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Jurisdiction-map coverage	Share of the AI estate mapped to every regime it touches.	AI systems found in operation that were not on the jurisdiction map
Map currency	Cadence the jurisdiction map is refreshed against regulatory change	Elapsed time between a material regulatory change and its reflection in the map
EU AI Act high-risk readiness	Share of in-scope high-risk systems with complete technical documentation and documented human oversight	Overdue items in the high-risk package
Conflicting-obligation handling	Cross-jurisdiction conflicts logged in the conflict register	Unresolved conflicts past their review date

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Exposure is usually light because the firm operates in one regime, though a boutique that manages funds domiciled in another jurisdiction or has clients in other jurisdictions is not.

MID-SIZED

Capacity squeeze

The risk is that the firm expands cross-border faster than it builds its compliance map, so it should stand up the jurisdiction-mapped inventory and the EU baseline before it scales.

LARGE AND GLOBAL

The large firm carries the full weight of divergence and the real conflicting-obligation cases, so its task is the single high-bar program with overlays, active conflict-screening.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST11

P05-ST12

Incident Management and AI Failure Response: Detecting, Containing and Learning When AI Goes Wrong

For: Risk and Compliance; Operations; Technology

AI fails differently from the software a firm already knows how to run, so incident handling has to be extended for it. In a 2025 survey of large firms across sectors, 99% reported financial losses from AI-related risks, at an average of US$4.4 million among those that experienced them. The incident lifecycle covers detecting a failure, escalating, containing and responding to it, and learning from it, including near misses.

99%

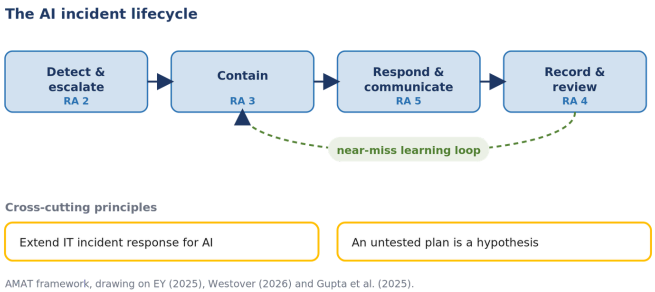
of surveyed large firms across sectors report financial losses from AI-related risks, averaging US$4.4 million among those affected (2025)

a fifth

of firms hold a tested AI incident response plan (2026 survey)

WHERE FIRMS GO WRONG

The first failure is reaching for the standard IT incident runbook: cybersecurity infrastructure translates to AI but must be extended to hallucination, drift, bias, prompt injection and agentic systems. The second is confusing a document with a capability, since a plan never exercised produces false confidence. The third is discovering mid-incident that the firm never captured enough of the model's state and decision context to reconstruct what happened.



AMAT framework, drawing on EY (2025), Westover (2026) and Gupta et al. (2025).

Takeaway. AI will fail, and the firms that handle it well treat incident management as a designed capability rather than a document: AI-specific runbooks, kill switches in front of irreversible actions, an incident record complete enough to reconstruct the decision, and a near-miss loop that turns each failure into a fix.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST12

RECOMMENDED ACTIONS

1. **Build AI-specific incident runbooks** rather than relying on standard IT runbooks, starting with the failure modes of the systems already in production.
2. **Detect and escalate**, starting with continuous monitoring on material and agentic systems and extending to every inventoried system.
3. **Install hard kill switches and agentic circuit breakers** before scaling autonomous systems, and test each stop in a drill.
4. **Define a quantified "significant AI incident" threshold** and log, from first use, a record complete enough to reconstruct any decision.
5. **Establish an AI incident-response protocol** inside the firm's existing incident process: containment, root cause, remediation and communication.
6. **Run near-miss reporting and post-incident reviews**, and test playbooks with tabletop drills.
7. **Have internal audit independently review incidents and remediation**, and report material AI incidents to the board.

SEQUENCE AND DEPENDENCIES

1. The sequence is to stand up AI-specific runbooks, instrument detection and escalation, put containment stops in place before scaling, establish the response protocol and the reconstructable record.
2. This article is the incident lifecycle, detection, escalation, response, near-miss reporting and post-incident review; the design of business continuity and operational resilience is P07-ST10.
3. The binding constraint is that the two failures most likely to turn a manageable event into a crisis are the untested plan and the unrecorded incident.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Detect and report incidents and near misses, operate kill switches, and keep the incident record for the systems they run.
Second Line	Own AI-specific runbooks, escalation thresholds and the incident-response protocol, and integrate incident capture into regulatory reporting.
Third Line	Independently review AI controls and the remediation of incidents, and test whether detection, containment and record controls operate as designed.
Board and Executive	Confirm material AI incidents are reported and that the firm is prepared, given a severe incident's board-level cost.
Cross-Functional	Align incident management with business-continuity and resilience design (P07-ST10), with the incident reporting processes of each regime the firm is under (Incident duties, Practical Applications).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST12

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Untested playbook / false confidence**	Exercise the playbook with tabletop drills and fix what breaks.
HIGH **AI-blind incident record**	Keep an aviation-black-box incident record and require vendor change notification.
HIGH **Standard IT runbook misfit**	Write AI-specific runbooks per failure mode and extend IR to agentic oversight.
HIGH **Detection-layer failure**	Test rules in a sandbox and tune for precision and latency before production.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Detection and escalation	% of AI systems under continuous monitoring with defined escalation	incidents detected late or by an external party
Containment readiness	% of autonomous / irreversible-action systems behind a kill switch	incidents where no stop was available
Stop time	elapsed time from the decision to stop to confirmed cessation of the system's actions, measured in a drill	stops that took longer than the firm's target, or did not hold
Record completeness	% of material AI systems logging the P05-ST09 decision record at the time.	incidents that could not be reconstructed or explained

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

A boutique runs on vendor AI it cannot instrument or fully investigate, so its exposure is the unnotified vendor model change it will not detect.

MID-SIZED

Capacity squeeze

Scaling agentic use outpaces manual incident response before the firm has AI-specific runbooks or kill switches, so it should put both in place before it scales.

LARGE AND GLOBAL

The large firm faces agent sprawl and multi-jurisdiction reporting, so its task is to integrate incident capture into its existing incident and post-trade reporting.



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST12

P05-ST13

Systemic AI Risk and Market Stability: When Everyone's AI Makes the Same Bet

For: Risk and Compliance; Senior Leadership; Board and Executive

AI in asset management is now deep enough, and similar enough across firms, that its biggest risks are no longer only firm-level. A 2026 study puts algorithmic trading at around 70% of US equity volume, with a single vendor holding about 92% of the data-center GPU market. When many participants run comparable AI on comparable data, the result is herding, model monoculture, algorithmic collusion and concentration in a few providers. Individually rational adoption produces collectively fragile markets, so a firm must reduce its own contribution to the monoculture, harden against correlated shocks it cannot prevent, and support the macroprudential response.

70%

of US equity volume is algorithmic trading, with machine learning embedded in 80 to 100 percent of trading firms' algorithms (2026 study)

92%

of the data-center GPU market is held by a single vendor (estimate)

WHERE FIRMS GO WRONG

The core failure is managing AI risk one firm at a time and being structurally blind to the system: when many firms adopt similar models trained on similar data, risks that look diversified inside one book are common across the market. Two traps follow. Widespread AI adoption can reduce apparent volatility while increasing tail risk, so a calm market can be a system growing more fragile underneath. And the transition may not be smooth, because pressure can push adoption and correlation past a tipping point in a discontinuous jump.

How firm-level AI risk becomes system-level risk



AMAT framework, drawing on Rodríguez Fernández, Funcas (2026) and the Bank of England (2025).

Takeaway. Systemic AI risk is the price of everyone buying the same good idea. Treat it as a measured exposure: know how much decision-making runs on shared models, hold real model, data and signal diversity, stress-test the scenarios where everyone moves at once, manage provider concentration and protect human judgment.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P05-ST13

RECOMMENDED ACTIONS

1. **Measure monoculture exposure**: how much decision-making runs on shared models and data, and concentration across models, data sources and vendors.
2. **Require model, data and signal diversity** for the most systemic decisions, with two or more structurally different model families and correlation limits.
3. **Stress-test for correlated scenarios**, such as a common signal or a provider outage, not just idiosyncratic ones.
4. **Manage provider concentration and keep the exit open**, with a tested exit for every provider treated as a systemic dependency.
5. **Preserve human judgment** against cognitive dependency and invisible groupthink through dependency limits, diversified AI consultation and override discipline.
6. **Engage the macroprudential agenda**, supporting regulator monitoring and AI-specific stress testing.

SEQUENCE AND DEPENDENCIES

1. The defining implementation challenge is that this is a collective-action problem.
2. This article is the market-level risk that arises from correlated adoption; the firm-level discipline of validating and monitoring a single firm's models is P05-ST04.
3. A firm can measure its monoculture exposure, diversify its models and data, stress-test correlated scenarios, and manage provider concentration, but it cannot prevent a market-wide event by itself.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Watch for crowding and correlation in their own strategies, diversify signals and models, and flag when a trade looks consensus across the market.
Second Line	Own model, data and signal diversity requirements, map which vendors share an underlying model provider, and own correlated-scenario stress testing.
Third Line	Challenge the assumption that firm-level diversification holds at the system level, and test whether monoculture and concentration exposures are actually measured.
Board and Executive	Own the firm's contribution to market-wide risk and its third-party concentration, and engage the macroprudential and regulatory agenda.
Cross-Functional	Coordinate with industry peers and supervisors on shared AI risk, including data sharing and common stress scenarios, because prevention is collective.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST13

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Procyclical flash / correlated-shock events**	Stress-test common-signal and provider-outage scenarios; know the market circuit-breakers.
HIGH **Algorithmic collusion / emergent manipulation**	Surveil for AI-driven collusion and manipulation; apply impact-based controls alongside intent-based ones.
HIGH **False comfort / hidden tail risk**	Measure tail risk and monoculture exposure, not observed volatility alone.
HIGH **Cognitive dependency / skill erosion**	Set dependency limits and keep override discipline so independent judgment survives.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Monoculture exposure	A tracked monoculture indicator (adoption x correlation x feedback)	Return / position correlation with peers observed in stress
Model & data diversity	Number of structurally different model families and data sources for strategic decisions; for bought AI.	Single-model dependence found in an event
Provider concentration	Share of critical AI capability from the top provider; tested-exit status	Dependency or outage events traced to a concentrated provider
Correlated-scenario coverage	Percent of material strategies stress-tested against common-signal and provider-outage scenarios	Correlated exposures discovered only in a live event

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

The boutique's exposure is inherited: it runs on a small number of vendor models it cannot instrument or diversify.

MID-SIZED

Capacity squeeze

The risk is that the firm scales correlated AI strategies faster than it builds diversity discipline, so it should set diversity rules (model families where it builds.

LARGE AND GLOBAL

The large firm has the resources for model diversity, monoculture metrics, and stress testing, but its own scale makes it part of the concentration problem.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P05-ST13

Pillar 6 Executive Summary

For: Senior Leadership; People and Change; Board and Executive · Also relevant to: Investment Teams, Risk and Compliance, Technology, Finance

THE ARGUMENT

In this guide's view, AI transformation turns as much on people as on models. Pillar 6 is the human dimension of the guide: the literacy every role now needs, the redesign of jobs before the retraining of people, the change-management discipline that turns usage into adoption, the new AI roles, the culture that keeps human judgment and challenge alive, and the leadership and talent strategy behind it all. In this guide's view the evidence points one way across the pillar: talent and behavior, more than technology, decide whether AI is adopted safely, and firms do better when they treat the people work as the transformation itself rather than the soft add-on to it. The evidence has stated limits: program and adoption figures are firm- or self-reported, no survey with a stated sample covers job-loss concern among long-only staff, and no source uses the guide's firm-size tiers.

62%
of executives in BCG's 2025 cross-industry survey cite a talent and AI-skills shortage, the most-cited barrier to AI value

30pp
gap in a 2025 Writer survey: 75% of the C-suite but 45% of employees say AI adoption succeeded

11 mo
average time for financial-services firms to fill AI roles needing regulatory knowledge, against 5 for general ML (one 2026 assessment)

WHY NOW

The people agenda has moved from optional to supervised. Where the EU AI Act applies, the firm must take measures to support AI literacy among the staff who operate AI; the SEC's March 2025 roundtable called the AI and data skills gap a resourcing risk, and participants in CFA Institute's 2024 roundtables expressed concern that under-resourcing holds back safe adoption. Hybrid talent is scarce and, on the evidence reviewed, concentrated in the larger firms, leaving smaller firms more exposed to implementation risk and vendor dependence; one 2026 assessment puts the average time to fill AI roles that need regulatory knowledge at eleven months. In this guide's view the workforce is a serious constraint on safe AI adoption.

RISKS TO AVOID

Access without activation: buying tools and tracking course completions while real capability, measured by active use, never forms. **Retrain without redesign**: upskilling people into unchanged jobs, so the expected productivity never appears and the entry-level pipeline quietly erodes. **Titles and dashboards without mandates**: declaring new roles and change programs done on a rising usage curve, while leaders overestimate progress and unaddressed fear can harden into quiet resistance.

Bottom line. AI transformation is a people transformation. In this guide's view the constraint is often not the model but the literacy, the redesigned roles, the culture of challenge and the scarce hybrid talent that make AI safe and productive, and every one of those is a leadership choice, not an HR by-product. Firms that redesign the work before retraining, run adoption as disciplined change, and compete for talent on meaning as well as money are better placed to convert AI capability into realized value; firms that buy tools and count completions are likely to stay in pilots. Take the people with you.

Chapter Leads, Pillar 6: Konstantin Boehmer

Explore Pillar 6 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P06-summary

PRIORITY MOVES FOR LEADERSHIP

1. **Mandate baseline AI literacy across every function**: where the EU AI Act applies, take the measures it asks for to support literacy among the staff who operate AI, teach the concrete core skills, and measure active use, not course completions.

2. **Redesign the job before you retrain the person**: redraw roles function by function along the roadmap, reskill into the new shape, and protect the entry level that grows future judgment.

3. **Run adoption as a structured change program that builds trust**: appoint an owner, build a protected champion network, make managers the multipliers, name and answer displacement fear, and reward changed behavior over rising usage.

4. **Tier training and the new roles to risk and outcome**: cut function-and-risk training pathways that teach judgment, gate the riskiest tools behind certification, and make the new AI roles real with mandates, resource and career paths.

5. **Engineer human challenge and accountability into the process**: require an independent view before the model is consulted on material decisions, protect diversity and psychological safety, and keep a named human accountable.

6. **Win talent on value and educate the board to govern**: map the hybrid skills you need and choose build, borrow or buy, compete on meaningful work over salary, and give directors enough AI understanding to test what they are told.

IN THIS PILLAR

P06-ST01	AI Literacy Across the Organization: From Access to Activation	172
P06-ST02	Role Redesign and Future Skills: Redrawing the Job, Not Just Retraining the Person	175
P06-ST03	Change Management and Adoption Strategy: Getting AI Past the Pilot and Into the Work	178
P06-ST04	Training by Function and Risk Level: Matching the Curriculum to the Job and the Stakes	181
P06-ST05	New Roles AI Creates: Product Owners, Model Stewards, Prompt Engineers and Workflow Designers	184
P06-ST06	Human Judgment, Challenge and Accountability: Keeping the Human Edge as AI Scales	187
P06-ST09	Workforce Impact and Capacity Planning: Planning the Shape of the Workforce, Not Just Its Size	190
P06-ST10	Leadership Education and Board Training: Educating Boards to Govern AI, Not to Build It	193
P06-ST11	Hiring and Sourcing AI Talent: Winning Scarce People When Everyone Wants Them	196
P06-ST12	Managing Fear, Resistance and Trust: Making It Safe Enough to Adopt	199

Explore Pillar 6 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P06-summary

AMAT

P06-ST01

AI Literacy Across the Organization: From Access to Activation

For: All Staff

Baseline AI literacy is the floor under every other people move in AI transformation: a workforce that cannot use AI confidently and critically cannot redesign its roles, challenge its outputs or govern its risks. The gap that matters is between the staff who can reach an AI tool and the smaller share who use it in daily work, and a program that tracks only course completions closes the wrong one. Where the EU AI Act applies, firms must take measures to support literacy among the staff who operate or use AI systems on their behalf, which strengthens the case for an all-staff baseline. The recommended approach mandates literacy across every function, teaches prompt, data and critical-evaluation skills, starts each person from where they are, delivers it continuously rather than annually and measures applied use.

61%

of surveyed US asset management firms rank AI awareness and education as their top change priority (2025)

60%

ceiling on staff with AI access who use it in daily work: fewer than 60%, roughly flat year on year (cross-industry, 2026)

WHERE FIRMS GO WRONG

Firms deploy tools and licenses and assume capability will follow, yet insufficient worker skills remain the biggest barrier to integrating AI into workflows. Many still equate a finished module with a literate employee. Left to self-selection, access also amplifies existing skill advantages, so the capable pull further ahead while others fall behind.

Access is not the same as activation

+50% one-year growth in AI access — **the activation gap** — **<60%** of those with access use AI daily (flat YoY)

Source: Deloitte AI Institute, State of AI in the Enterprise, 2026 (cross-industry survey).

Takeaway. Stop counting course completions and start closing the gap between access and activation. Mandate literacy across every function, teach the core skills, deliver it continuously and measure whether people use AI well.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST01

RECOMMENDED ACTIONS

1. Name an owner for the literacy program. Start by covering the staff who operate or use AI systems and the functions where usage is lowest; build toward baseline AI literacy mandated across every function, not just technology.
2. Teach a concrete core skill set: prompt literacy, data literacy, and the ethical, critical evaluation of AI outputs, and the firm's approved-use and data-handling boundaries.
3. Start by replacing the annual generic module with short, role-relevant modules on the tools already deployed; build toward continuous, role-adaptive delivery (just-in-time micro-modules, on-demand assistants, spaced reinforcement).
4. Address the risks and concerns people actually have alongside the benefits, and secure active managerial support as a key adoption lever.
5. Measure applied use and its quality, not course-completion rates; keep completion as an administrative record and treat a high completion rate with low usage as a red flag.
6. Embed literacy into onboarding so new joiners start AI-literate, and refresh the curriculum, and the expertise of the people who teach it, as the tools change.
7. Start by tracking activation by function and seniority; build toward a program managed equitably so AI access does not widen the gap between the most and least capable staff.

SEQUENCE AND DEPENDENCIES

1. In this guide's view, managerial engagement is among the most important enablers, alongside curriculum quality.
2. Start from the approved tools the firm already has.
3. Function-specific and risk-tiered training, with deeper curricula for portfolio managers, risk, compliance and audit calibrated to the risk of the use cases they touch, is P06-ST04.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Reach every function, not just technology: front-office, operations and distribution staff need baseline prompt, data and critical-evaluation literacy to use AI safely in daily work.
Second Line	Build first-line-compliance literacy so risk and compliance staff understand model behavior well enough to supervise it; the deeper risk-tiered curriculum is P06-ST04.
Third Line	Internal audit needs enough AI literacy to assess AI-augmented processes and to audit the literacy program itself on activation, not completion.
Board and Executive	Sponsor and fund the firm-wide program and own the firm's measures to support a literate workforce.
Cross-Functional	A named program owner, typically in People and Change or the function responsible for workforce learning, coordinates the literacy program, its delivery model.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST01

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Access without activation**	Design for productive engagement and measure activation, not access; target the functions where usage is lowest.
HIGH **Completion-only literacy**	Measure applied use and engagement, not completion; treat high completion with low usage as a red flag.
HIGH **Passive reliance**	Teach critical evaluation as a core literacy skill; pair literacy with the challenge culture of P06-ST06.
HIGH **Unapproved AI feature or data disclosure**	Teach the approved-use and data boundaries before activation; record approved features in the AI inventory (P05-ST10).

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Program reach	Share of all staff (all functions) with baseline literacy assigned and a named program owner	Functions with no baseline literacy coverage
Activation and quality of use	Active, sustained use of approved AI tools among trained staff with access to them.	Trained staff with access who never use AI in their work
Core-skill competence	Demonstrated prompt, data and critical-evaluation competence, including verifying and challenging a work-relevant AI output (assessed, not attended).	Staff passing completion but failing an applied assessment
Equity of literacy	Spread of activation across functions and seniority	Widening usage gap between the most and least capable staff

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

For a boutique the advantage is reach: with a small headcount, firm-wide literacy is quick to achieve and easy to keep current.

MID-SIZED

The mid-sized firm has a real opportunity: it is large enough to fund a dedicated, role-adaptive program and small enough to reach every function quickly.

LARGE AND GLOBAL

The large firm has the resource to run a scaled internal academy with continuous refresh, but in this guide's view its characteristic risk is completion-only literacy.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST01

AMAT

P06-ST02

Role Redesign and Future Skills: Redrawing the Job, Not Just Retraining the Person

For: People and Change; Senior Leadership

AI is changing the tasks and skills within existing asset management roles, while the balance between redesign, recruitment and workforce reduction varies by firm and the longer-term employment outcome remains uncertain. As machines absorb information gathering and routine processing, the value of analysts, portfolio managers, operations and compliance staff shifts toward framing questions, directing the machine and judging its output. Capturing that value means redesigning the job around AI and reskilling deliberately into it. A large 2025 survey ranks the talent and AI-skills shortage as the top barrier to AI value, and a 2025 hedge-fund working paper associates generative-AI use with higher returns, most strongly where AI talent was already in place.

62%

of surveyed executives name a shortage of talent and AI skills as the top barrier to AI value (cross-industry, 2025)

6%

had meaningfully begun upskilling their workforce

WHERE FIRMS GO WRONG

Firms treat reskilling as a substitute for redesign, retraining people in place rather than rethinking who does what once AI is in the workflow. They also let the entry level erode silently: in US cross-industry studies cited by EY, junior employment fell at AI-adopting organizations while senior employment held, which is not an asset management measure but a pattern to watch, since it removes the base from which future senior judgment is grown. Few firms plan for how scarce the hybrid domain-plus-AI skill set is.

How core investment roles can be redesigned

Dimension	Legacy role	AI-redesigned role
Research analyst	Gathers and processes data	Directs and validates AI output
Portfolio manager	Runs research and monitoring	Frames questions; judges change
Operations	Executes routine processing	Supervises AI; owns exceptions
Risk & compliance	Manual checking and review	Domain judgment plus AI fluency

AMAT framework, drawing on Caridi et al. (2026), EY (2026), KPMG International (2026), CFA Institute (2026) and Chartered IIA (2026)

Takeaway. Start with the role, not the training, on the next function on the transformation roadmap, and keep the workforce plan alive as both the tools and the roles keep moving.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST02

RECOMMENDED ACTIONS

1. Redesign each AI-affected role around AI before reskilling into it, starting with the task split for one workflow; do not bolt AI onto an unchanged job.
2. Start by mapping the skill gap for the functions next on the transformation roadmap; build toward a mapped gap, with defined hybrid capabilities, for every redesigned role.
3. Make upskilling current staff the primary talent pathway, blended with targeted external hiring for scarce hybrid skills.
4. Start by tying reskilling for the first redesigned roles to named business objectives; build toward role-based curricula and visible AI career pathways across the firm, not a generic training catalog.
5. Build prompt literacy and AI-output verification into every role that uses AI, as a core operating competence, with reviewers given the competence, time and authority to override or escalate.
6. Start by tracking entry-level intake and progression in the functions being automated and a supervised development task per workflow; build toward redesigned junior roles, so automating junior tasks does not hollow out future senior judgment.
7. Start with a combined human-plus-AI workforce plan for the functions already using AI; build toward a firm-wide total-workforce plan, refreshed every 6-9 months, that keeps role redesign ahead of tool rollout.

SEQUENCE AND DEPENDENCIES

1. The execution risk firms often underweight is the hand-off between domain experts and technical teams: the redesign stalls there unless the firm deliberately builds the translator capability that connects them.
2. The order in the Recommended Actions, role first and tools second, holds only if it is enforced; RA 7's workforce plan is where that happens.
3. New AI-specific roles are P06-ST05 and workforce sizing is P06-ST09; this article assumes both.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the redesign of front-office and operations roles, run AI-output verification, and build prompt literacy where AI is used.
Second Line	Reskill risk and compliance into hybrid roles that pair domain judgment with AI fluency.
Third Line	A Chartered IIA white paper produced with Diligent describes internal audit effort moving from data gathering toward analysis and challenge.
Board and Executive	Own the workforce-transformation strategy, treat talent as a leading constraint on AI value, and protect the entry-level pipeline against short-term automation savings.
Cross-Functional	The accountable business-function head owns the redesigned role and the live AI-assisted workflow, including staffing for review, escalation and updates.



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST02

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Hybrid capability gap stalls workflow integration**	Make upskilling the primary pathway and grow hybrid talent deliberately; map the gap before buying tools.
HIGH **Reskilling without redesign**	Redesign the role first; treat reskilling as the second step, targeted at the redesigned job.
HIGH **Entry-level erosion breaks the pipeline**	Redesign and protect entry-level roles; measure pipeline health, not only near-term headcount efficiency.
MEDIUM **Hybrid-talent flight**	Build AI career pathways and a strengthened value proposition; use internal mobility to retain reskilled talent.

METRICS SCORECARD

AREA	LEADING	LAGGING
Role redesign (control)	Share of AI-using roles with a redesigned job definition before tool rollout	Tool deployments made against unchanged role definitions
Changed work	Share of selected workflows in which AI is used in regular production work under the redesigned task split	Change against the firm's own baseline in time to a validated investment memo and in rework per workflow
Verification competence	Share of users in AI-using roles meeting an observed, role-specific verification standard	Errors found after review in sampled AI-assisted output
Reskilling coverage	Share of staff in redesigned roles who have completed role-based reskilling	Roles still operating on pre-AI skill profiles

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

A boutique's advantage is agility: with few people to reskill and no legacy job architecture, roles can be redrawn quickly.

MID-SIZED

The mid-sized firm has a real opportunity: it is large enough to fund a structured reskilling program and small enough to redesign roles across functions in one coordinated pass.

LARGE AND GLOBAL

The large firm can afford to scale structured reskilling, interdisciplinary teams and internal mobility, so its constraint is coordination rather than budget.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST02

AMAT

P06-ST03

Change Management and Adoption Strategy: Getting AI Past the Pilot and Into the Work

For: Senior Leadership; People and Change

The hard part of AI in asset management is not building the pilot but getting the whole firm to change how it works. Firms invest heavily in tools and employees experiment, then the shift to embedded, enterprise-wide use often stalls for people-and-change reasons even when the pilot works technically. This discipline covers how to structure the program, deploy peer champions, embed AI in everyday workflows, communicate the change and design incentives and metrics that make adoption stick. Run as a genuine change program, adoption is far more likely to compound; run as a tooling rollout with a usage dashboard, it tends to produce activity without change.

45%

of employees believe their organization has successfully adopted AI, against 75% of the C-suite (self-reported, survey of 1,600 knowledge workers, cross-industry, 2025)

WHERE FIRMS GO WRONG

Firms measure activity and declare victory, when the real question is whether AI has become central to how core work is done. Business and technology teams pull apart, which strands investment when budgets tighten. And change programs route around skeptics instead of engaging them with data and peer demonstrations in small forums.

Deploying AI vs embedding AI

Dimension	Deploying (tech lens)	Embedding (change lens)
Focus	Roll out the tools	Change how work is done
Success measure	Logins and usage time	Behavior change / adoption
Who drives it	IT and vendors	Champions and managers
Skeptics	Ignored or overridden	Engaged with evidence
Outcome	Usage up, impact flat	Adoption sticks

AMAT framework, drawing on BCG (2025) and Van Rossum / Lead with AI (2026)

Takeaway. Run adoption as a structured program with an operating structure, an influence-based champion network and managers as multipliers, and above all measure whether behavior has changed rather than whether usage has gone up.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST03

RECOMMENDED ACTIONS

1. Start by naming an adoption-program lead, heading a small center of excellence, and a first group of champions for the tools already approved; build toward this guide's three-layer structure: an AI lab or center of excellence, a champion network, and the everyday tools where work changes.
2. Start by selecting champions by influence in the teams already piloting AI, at 30 to 60 minutes a week; build toward a firm-wide peer champion network placed by influence (network analysis where available) and designed against burnout.
3. Make frontline managers the multipliers and convert early adopters into peer coaches; engage skeptics with evidence, do not override them.
4. Start by agreeing stakeholder roles and a senior sponsor for the current program; build toward business and technology aligned on agile value streams with senior oversight (AM transformation).
5. Measure behavior change, not logins; design incentives that reward good adoption behaviors and avoid a cost-cutting framing.

SEQUENCE AND DEPENDENCIES

1. In this guide's judgment, built on the source guide's start-small, find-connectors and embed-then-support approach, the operating structure and the champion network are best established before any push for scale.
2. This article is the embedding engine and owns the methodology, the champions and communities of practice, the communication of change to employees, and the incentives and metrics.
3. Because adoption and sustained usage behave as distinct processes (application 6, an association only), build measurement and adjustment in from the start.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the behavior change: managers model AI use, champions embed it in real workflows, and desks adopt AI-assisted ways of working as the norm.
Second Line	Ensures the adoption program stays inside firm-approved tools and data boundaries, and that incentives do not push staff toward unsafe or non-compliant use.
Third Line	On a risk-based plan, samples live AI-assisted workflows to test that the named owner, approvals, review evidence and fallback exist and operate.
Board and Executive	Sponsors the change as a cultural shift, funds the change-management work rather than only the tools, and reads behavior-based adoption measures, not usage dashboards.
Cross-Functional	An executive sponsor provides resources and settles cross-functional conflicts, and the adoption-program lead (RA 1) keeps one plan across People and Change, the business and Technology.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST03

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Measuring activity, not adoption**	Measure behavior change and workflow norms; treat the C-suite-versus-employee perception gap as a correction signal (this guide's recommendation).
HIGH **Pilot-to-enterprise stall**	Run adoption through a structured operating model and a deliberately built champion network.
HIGH **Unowned live workflow and single-champion dependence**	Name the workflow owner and confirm the inventory entry and approved data boundaries before first use; where outputs reach clients or regulators, a person reviews them.
MEDIUM **Business-IT disconnect**	Align business and IT early on agile value streams with clear stakeholder roles and senior oversight.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Adoption depth	Share of a quarterly sample of recurring work products produced through the agreed AI-assisted workflow	Change against the team's pre-AI baseline in time to complete.
Perception alignment	Employee-versus-C-suite adoption-success gap	Trend in the gap over time
Champion network	Champions as a share of users; lead-to-champion ratio	Champion retention past six months; number of other teams that independently repeat a documented, quality-checked workflow improvement
Manager activation	Share of managers visibly using AI	Team-level adoption where managers lead

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Approach: A handful of informal champions and leaders who visibly model use can carry a small firm; Do Now: Name a few peer champions, protect their time.

MID-SIZED

Opportunity: In this guide's view, a mid-sized firm is large enough to fund a formal champion program with a small center of excellence and small enough for one adoption-program lead.

LARGE AND GLOBAL

Opportunity: Scale supports the full three-layer model; Do Now: Run the AI lab or center of excellence, place champions across the organization by network analysis, and design against champion burnout.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST03

AMAT

P06-ST04

Training by Function and Risk Level: Matching the Curriculum to the Job and the Stakes

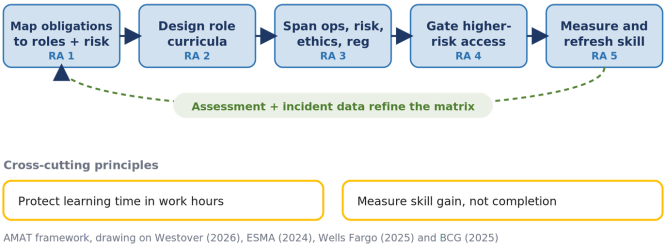
For: People and Change

Baseline literacy gets everyone to a common floor; what comes next is training calibrated to what each function does with AI and to the risk of the use cases it touches. A compliance officer under a conduct regime and a portfolio manager running a research copilot face different obligations and different failure modes. The recommended approach maps obligations and risk to roles, builds pathways by function, covers a tool's risks, ethics and regulatory limits as well as its use, gates the highest-risk tools behind certified training and measures competence, not completion. For investment firms providing retail investment services, ESMA treats MiFID II knowledge-and-competence expectations as already applying to staff who use AI to give clients product information.

WHERE FIRMS GO WRONG

The dominant failure is the one-size-fits-all module, which wastes experienced staff's time and leaves exposed staff under-trained. Programs also teach tool mechanics alone, leaving staff unable to recognize data-integrity problems, bias and unintended consequences. And access to higher-risk tools is often certified on completion alone, which does not show that a person can tell when the output is wrong.

Building training that matches the job and the stakes



AMAT framework, drawing on Westover (2026), ESMA (2024), Wells Fargo (2025) and BCG (2025)

Takeaway. Treat role and risk as the two axes of a single curriculum: map obligations to roles, cut distinct pathways, teach judgment as well as mechanics, gate the riskiest tools and measure competence rather than completion.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST04

RECOMMENDED ACTIONS

1. Start by mapping obligations and risk tiers for the roles that already use AI; build toward a regulation-by-role-by-risk matrix across jurisdictions as the spine of every training pathway.
2. Start with distinct pathways for the control functions and the roles using higher-risk tools; build toward distinct curricula for every function (investment, risk, compliance, audit, operations, client-facing), not one shared course.
3. Make every pathway cover operation, risk, ethics and regulation, including prompt craft, verification and red-flag recognition tied to the firm's AI policy.
4. Gate higher-risk AI tools behind certified training (a scenario-based assessment and the business owner's documented approval) and IT access controls, with exceptions recorded, access revoked on role change or expiry, and point-of-use reminders and control intensity scaled to risk.
5. Measure competence, not completion, with short role-specific simulations. Start by assessing skill on the higher-risk pathways; build toward periodic recertification for higher-risk tasks, incident data trended alongside as an association, protected in-work learning time and pathways tied to career paths.

SEQUENCE AND DEPENDENCIES

1. The matrix is a prerequisite, not a parallel workstream: pathways cannot be designed coherently until obligations and risk tiers are mapped to roles.
2. Assign one accountable business owner for each higher-risk AI use case and one program owner for the training matrix, usually People and Change working with Risk or Compliance.
3. Two conditions determine whether the curriculum sticks.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Train investment and client-facing staff to use AI decision-support and sanity-check its outputs, and hold client-facing roles to the same knowledge-and-competence bar that applies without AI.
Second Line	Control-function training is the priority tier: curricula here must go beyond mechanics to risk, ethics and regulatory implications.
Third Line	Internal audit needs its own AI competence to assure AI controls; the audit-specific training profile and roadmap are developed in Pillar 7 and are cross-referenced here.
Board and Executive	Board and leadership AI education is the subject of P06-ST10 and is not developed here; this article stops at the function-and-risk training layer beneath it.
Cross-Functional	Run the regulation-by-role-by-risk matrix, protected learning time and career-linked pathways as shared infrastructure so every function draws from one coherent program rather than parallel, inconsistent ones.



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST04

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **One-size-fits-all training**	Build a regulation-by-role-by-risk matrix and cut a distinct pathway per function from it.
HIGH **Ungated higher-risk access**	Require certified training (a scenario-based assessment and the business owner's documented approval) before use, backed by IT access controls and combined into a technical access gate.
MEDIUM **Mechanics-only curriculum**	Make every pathway span operation, risk, ethics and regulation, with verification steps and red-flag recognition tied to the AI policy.
MEDIUM **Completion standing in for competence on higher-risk pathways**	Measure skill with role-specific simulations and sampled independent review rather than tracking completion, trending incidents alongside as an association; track the rate at which reviewers edit.

METRICS SCORECARD

AREA	LEADING	LAGGING
Coverage by role (control)	Share of AI-exposed roles with a role-mapped pathway assigned	Share completing their role-based pathway
Higher-risk access control	Share of higher-risk-tool users certified before access	Access-gate exceptions and related incidents
Competence, not completion	Diagnostic assessment scores by topic and role	Pass rates on role-specific simulations, repeated after training; errors found by independent review of sampled higher-risk AI-assisted work.
Behavioral effectiveness	Share of sampled higher-risk AI-assisted work showing documented verification by the user	Change in on-the-job behavior, including staff reporting changed practice

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Advantaged

Opportunity: AI gains landed hardest with less-experienced staff in one customer-service field experiment, so in this guide's view a lean bench may benefit most; Do Now.

MID-SIZED

Opportunity: In this guide's view, a mid-sized firm is large enough to fund role-based pathways and small enough to map every function into the matrix in one coordinated pass.

LARGE AND GLOBAL

Advantaged

Opportunity: Dedicated learning and control functions can run distinct pathways per function; Do Now: Operate the matrix across jurisdictions and enforce certified gates for every higher-risk use case.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST04

AMAT

P06-ST05

New Roles AI Creates: Product Owners, Model Stewards, Prompt Engineers and Workflow Designers

For: People and Change; Senior Leadership; Technology

AI transformation does not only reshape existing jobs; it creates a small set of AI-specific roles. Two give established disciplines a new mandate: the AI product owner, accountable for the business value of an AI use case, and the model steward, who holds first-line ownership of model-risk oversight. Two have little precedent before AI: the prompt engineer, who works on model outputs, and the workflow designer, who shapes how people and agents work together. These roles fail when a firm adds a title without the budget, authority and career path that make it real. The fix is to define each role by the outcome it owns, give it a clear home, starting with a named cross-functional group, and keep regulatory accountability with senior management rather than letting it drift onto a new job title.

70%

of surveyed Italian asset managers have the established data and AI roles, against around 30% with a GenAI or prompt specialist (2025)

WHERE FIRMS GO WRONG

A common failure is creating the title without the mandate, so accountability blurs when a model underperforms. Firms also scale agents without anyone managing them as a workforce, and they confuse a new role with a redesigned one, which leads them to over-hire specialists they could have grown or to expect an existing team to absorb a new accountability in its spare time.

Four roles AI creates, and how each one fails

Dimension	What the role owns	Where it fails
AI product owner	Business value of the AI use case	No budget or decision rights
Model steward	Model-risk oversight in the first line	Blurred with the data owner
Prompt engineer	Quality of model outputs	Treated as a fad, not a skill
Workflow designer	Work around human + agent teams	Automates without redesign

AMAT framework, drawing on Deloitte (2026), EY China (2026) and KPMG (2026).

Takeaway. Define each role by the outcome it owns, resource it with a career path and keep regulatory accountability with senior management, because a new title carries ownership but not answerability.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST05

RECOMMENDED ACTIONS

1. **Define each role by the outcome it owns**, not its tools, with explicit decision rights in the job description.
2. **Give the roles one clear home**, even a small cross-functional group, building toward a center of excellence run as an internal market where scale justifies it.
3. **Name an AI product owner** accountable for the value each material AI use case must realize, using two-way secondment to build the hybrid profile.
4. **Give every material model and agent a first-line steward**, with independent challenge in the second line.
5. **Keep regulatory accountability with senior management**, and give each role budget, authority and a career path or do not create it.
6. **Grow the newer roles internally** where you can and buy in only what you cannot grow.

SEQUENCE AND DEPENDENCIES

1. The sequence that works is to define the role before hiring for it, situate it in a home before scaling it, and anchor accountability before either.
2. Given the people bottleneck described above, the realistic path is to grow the newer roles internally where the raw talent exists, as the learning.
3. Where a firm cannot yet fund a genuine role, the honest interim is one accountable AI lead and borrowed specialist skills, not half-titles that carry responsibility without authority.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the new roles that sit in the business, chiefly the AI product owners accountable for the value of each AI use case and the model.
Second Line	Owns the model-risk framework and risk tiers, independently validates and challenges the stewards' first-line assessments and escalates unresolved issues.
Third Line	Tests whether new AI roles carry real authority and budget, and whether accountability for each material model is clearly assigned, documented and evidenced rather than notional.
Board and Executive	Owns the decision to fund genuine roles and a center of excellence, and confirms the firm can name who is answerable for each material model.
Cross-Functional	Redesign of existing roles (P06-ST02), hiring, employer-brand and partner-talent strategy (P06-ST11), model risk management (P05-ST04) and the accountability regimes (P05-ST02) are cross-cutting and are referenced here.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST05

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Title without a mandate**	Define each role by the outcome it owns and give it decision rights, budget and a path; do not create a role you cannot make real.
HIGH **Accountability drifts onto a job title**	Keep accountability with senior management and document each role's responsibilities; the new role holds ownership, not answerability.
HIGH **Talent bottleneck strands AI in pilots**	Name an AI product owner and build the hybrid profile through two-way secondment so the business-technology gap is owned by someone.
MEDIUM **Unmanaged agent workforce**	Enter every material AI feature, vendor-enabled ones included, in the inventory before first use.
MEDIUM **CoE drifts from funded business problems**	Require a business sponsor for each use case, a simple service agreement and evidence of realized workflow benefits; adopt chargebacks.
MEDIUM **New role confused with role redesign**	Classify each need as new or redesigned, grow internally where possible, and treat existing-role evolution as P06-ST02.

BY FIRM SIZE

BOUTIQUE

A boutique borrows rather than builds, using fractional or embedded specialists and vendor support instead of dedicated hires.

AMAT

P06-ST06

Human Judgment, Challenge and Accountability: Keeping the Human Edge as AI Scales

For: Investment Teams; Risk and Compliance

As AI makes prediction cheap and abundant, the decisive capability in an investment firm is human judgment exercised through active challenge. The danger is not that AI replaces judgment outright but that it can erode it quietly: teams consulting the same models can converge on the same views, reviewers can rubber-stamp polished outputs and unused skills can fade. Preserving judgment, challenge and accountability is a deliberate design problem, not a matter of good intentions. Firms need to engineer challenge into the process, protect the cognitive diversity and psychological safety that make it work, and keep a named person accountable for every AI-assisted decision.

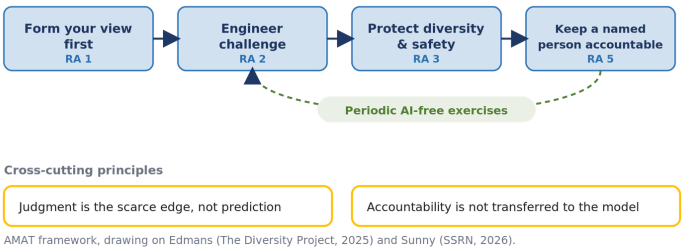
2.7%
of polled finance and accounting professionals would always trust AI agents to make decisions, judgment calls included (2025 webcast poll)

60%
would trust them only within a framework where people keep the judgment calls

WHERE FIRMS GO WRONG

Firms watch for visible replacement while the real damage is silent: collective decision quality can weaken even as individual productivity rises. Sustained use can dull critical thinking, and people tend to over-trust AI advice, especially on numbers. Challenge also does not happen on its own, since committee dynamics suppress dissent and a nominal human in the loop can become a formality under time pressure.

Engineering challenge into an AI-augmented process



AMAT framework, drawing on Edmans (The Diversity Project, 2025) and Sunny (SSRN, 2026).

Takeaway. Engineer challenge into the process, protect the diversity and safety that let it work, keep the skill alive and keep a named human accountable for every decision. Treat challenge as infrastructure rather than temperament.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST06

RECOMMENDED ACTIONS

1. Require staff to form and document their own view (initial thesis, key assumptions, main uncertainty) before consulting AI on any material discretionary decision.
2. Engineer adversarial challenge into the process. Start with a documented counter-thesis on material discretionary decisions; build toward an AI-free contributor who presents last on the highest-stakes ones, rotating the role, and a systematic record of disconfirming evidence.
3. Protect cognitive diversity and psychological safety as the conditions for challenge. Start by rewarding well-reasoned dissent, not just consensus; build toward measuring safety and dissent across seniority.
4. Run periodic AI-free analytical exercises to prevent cognitive deskilling and detect over-reliance early. Start with one exercise for the roles that rely most on AI; build toward exercises for every analytical role, with performance tracked over time.
5. Keep a named person explicitly accountable for every AI-assisted decision; the model is an input, never the decision-maker.
6. Evaluate process quality independently of short-term returns. Start by documenting the reasoning behind each material discretionary decision; build toward sampled error-catch and revision measures set against the firm's own baseline.
7. Guard against 'workslop' and uncritical acceptance: require substantive verification of AI output, not polish, before it is relied on.

SEQUENCE AND DEPENDENCIES

1. The dependency that decides whether any of this works is psychological safety, and it is slow to build and quick to destroy.
2. The boundary is worth restating because it is the pillar's most-confused line.
3. The baseline controls set out at the start of the Recommended Actions do not wait for it, but engineered challenge.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the judgment: form independent views before consulting AI, run engineered challenge on material theses, and treat AI output as an input to be interrogated.
Second Line	Guard against automation bias and over-reliance in the firm's own AI-assisted work, and set the expectation that AI outputs are challenged rather than trusted by default.
Third Line	Test whether challenge actually occurs, not whether a policy exists: sample decisions for a pre-AI view and a documented counter-thesis.
Board and Executive	Set the tone that judgment and dissent are valued, preserve human accountability for AI-assisted decisions (personal where a regime such as SM&CR applies.
Cross-Functional	Build the psychological safety, incentives and skill-retention exercises that make challenge possible, and align reward with well-reasoned dissent rather than surface consensus.



Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST06

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Silent homogenization of thinking**	Require an independent view before AI and, on the highest-stakes decisions, an AI-free contributor who presents last; measure diversity of input, not just agreement.
HIGH **Cognitive deskilling and over-reliance**	Mandate periodic AI-free exercises and use AI as devil's advocate; develop the complementary human capabilities AI cannot replace.
HIGH **Automation bias defeats human review**	Engineer active challenge into the review, not passive sign-off; track the rate at which reviewers edit.
MEDIUM **Challenge suppressed by hierarchy and incentives**	Build psychological safety and reward well-reasoned dissent; require a documented counter-thesis so challenge does not depend on courage.
MEDIUM **Accountability diffused into the model**	Keep a named human accountable for every AI-assisted decision; record the accountable decision-maker and retain evidence of their substantive assessment; treat human-in-the-loop as an accountability principle.
MEDIUM **Workslop and false confidence**	Require substantive verification proportionate to stakes; judge the content, not the presentation.

METRICS SCORECARD

AREA	LEADING	LAGGING
Independent input	Share of material discretionary decisions with a documented human view recorded before AI was consulted	Decisions where the first recorded view was AI-generated; convergence of documented pre-AI views across periodic look-backs
Engineered challenge	Share of material discretionary decisions with a documented counter-thesis.	Material decisions taken with no recorded dissent or challenge
Psychological safety	Measured team psychological-safety score; rate of recorded dissent from junior and lateral staff	Silence rate: senior views passing without challenge
Review challenge	Rate at which reviewers edit, override or reject AI output, trended	Material AI-assisted decisions approved with no recorded edit, override or challenge
Skill retention	Participation and performance in periodic AI-free analytical exercises	Decline in unaided analytical performance over time
Process quality	Documentation completeness of the reasoning behind material discretionary decisions	Share of sampled decisions where a material error in AI-generated analysis was found before approval rather than after.

BY FIRM SIZE

BOUTIQUE

The advantage is cohesion: a small team can institutionalize challenge and dissent directly, without committee machinery.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST06

AMAT

P06-ST09

Workforce Impact and Capacity Planning: Planning the Shape of the Workforce, Not Just Its Size

For: Senior Leadership; People and Change; Finance

The workforce question for AI in asset management is how work will be redistributed, not only how many jobs will be cut. Near-universal AI adoption is expected across financial services, and the workforce data show a change in composition: overall turnover falls while involuntary turnover rises, support and entry-level roles thin, and specialist, risk and compliance capability grows, although the source does not isolate AI as the cause. The recommended approach plans capacity: which work AI absorbs, where the pipeline is at risk and where capacity must be added. Tracking only headline headcount misses the disruption underneath.

95%
or more of financial-services employers expect significant AI adoption (2026)

4.9%
involuntary turnover across financial services in 2025, up from 3.5% in 2022, while overall turnover fell

WHERE FIRMS GO WRONG

Firms treat the question as a cut-or-not binary and manage to headline headcount, which understates the disruption underneath. They assume elimination and under-resource the transition, so they get neither the savings nor the capacity. And they thin the entry-level pipeline quietly: one 2026 set of interviews with firm leaders warns that curtailing entry-level hiring too aggressively risks capability gaps at mid-career level in five to ten years.

Headline turnover hides the shift

15.5% — overall FS turnover, 2025
down from 18.5% (2022)

the hidden shift

4.9% — involuntary turnover
up from 3.5% (2022)

Source: Aon, 2026 (4,200+ company database).

Takeaway. Plan capacity rather than count heads: map the work, route freed capacity into higher-value uses, plan the shift toward specialist and risk capability, plan the entry-level pipeline and measure the involuntary turnover the headline hides.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST09

RECOMMENDED ACTIONS

1. Build the capacity model on a map of automatable versus augmentable work, not a headcount target. Start by mapping the work in one function; build toward a firm-wide map that feeds the capacity model.
2. Plan freed capacity for redeployment into higher-value work and growth absorption, and resource the transition. Start by naming an owner and a redeployment path for the capacity the first AI use cases free; build toward a standing redeployment process with a quarterly review.
3. Plan the composition shift: size specialist, risk and compliance capacity to the workload the firm's AI deployments create; manage support and entry-level reductions as a planned change.
4. Plan the entry-level pipeline as a capacity variable (the redesign of junior roles is P06-ST02). Start by tracking junior hiring against plan; build toward intake planned against the firm's future requirement for experienced staff.
5. Track the disruption beneath the headline: split involuntary from voluntary turnover and monitor redeployment. Start by splitting the turnover figures the firm already holds and recording a reason for each planned movement; build toward tracking the redeployment rate.

SEQUENCE AND DEPENDENCIES

1. Sequencing matters: map the automatable and augmentable work before setting headcount targets.
2. Most managers will first gain AI capacity through features inside platforms they already use, so use the AI inventory (P05-ST10) to find where those features are live and compare business-user.
3. Workforce capacity planning sets the envelope; how roles are redrawn inside it is P06-ST02, and how the firm sources the specialist and AI talent the composition shift requires is P06-ST11.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Redeploy front-office and operations capacity freed by AI into higher-value work, and protect junior development so the pipeline is not thinned as a by-product of efficiency.
Second Line	Size oversight to the AI footprint: estimate review, third-party oversight, control testing and incident workload for the deployments in use.
Third Line	Internal audit capacity for AI assurance is a growing need with a real skills gap; the capability-planning detail is developed in Pillar 7 and cross-referenced here.
Board and Executive	Own the workforce-transition governance and the involuntary-turnover picture, and sponsor the dedicated AI leadership the composition shift increasingly requires.
Cross-Functional	The capacity model is jointly owned by Finance, People and Change, and the business, so headcount.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST09

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Managing to the headline**	Track the involuntary-versus-voluntary turnover split and the redeployment rate, not headline headcount.
HIGH **Junior intake falls below need**	Track junior hiring against plan and escalate shortfalls; the redesign of junior roles is P06-ST02.
HIGH **Overstated capacity released**	Pilot one workflow against baseline hours and output quality; net off review and rework before approving redeployment or reductions.
MEDIUM **Redeployment assumed, transition under-resourced**	Resource the change management that makes redeployment real, not just planned.

METRICS SCORECARD

AREA	LEADING	LAGGING
Turnover quality	Involuntary-versus-voluntary turnover trend	Exits by documented reason, including restructuring and verified AI-enabled change
Redeployment	Share of freed capacity with a redeployment path	Redeployment versus net-reduction ratio
Entry-level pipeline	Junior hiring against plan	Pipeline and leadership-bench health
Capacity coverage	Workflows with a recorded baseline of volume, staff hours, rework and control time	Net capacity released: baseline hours less review, exception-handling and correction time, checked with the workflow owner

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Advantaged

Opportunity: In this guide's view, a lean bench can redeploy freed capacity to absorb growth rather than cut; Do Now: Map automatable work and redeploy into higher-value activity.

MID-SIZED

Opportunity: In this guide's view, a mid-sized firm is large enough to fund a firm-wide capacity model and small enough for Finance, People and Change.

LARGE AND GLOBAL

Advantaged

Opportunity: In this guide's view, scale allows a planned composition shift toward specialists and dedicated AI leadership; Do Now.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST09

P06-ST10

Leadership Education and Board Training: Educating Boards to Govern AI, Not to Build It

For: Board and Executive; Senior Leadership

Boards and senior leaders do not need to build AI, but they do need to understand it well enough to govern it, and many still lack that understanding. Education has a narrow purpose: to give directors and executives enough grasp to set direction, ask the right questions and provide effective challenge, not to turn them into engineers. The greater danger is shallow engagement, where a polished quarterly briefing leaves a board feeling informed while unable to test what it is told. The recommended approach assesses the gap against what oversight requires, builds access to expertise, embeds AI in the governance model and has senior leaders visibly own it.

Nearly three quarters

of boards are perceived to have only moderate or limited AI expertise (cross-industry, 2026)

WHERE FIRMS GO WRONG

Firms treat AI as a siloed technology item that management and vendors own, so no one on the board is clearly accountable for oversight. They expect the board to catch up passively while senior leaders do not model the behavior themselves. And they let the oversight gap widen quietly, so it shows up only when something goes wrong.

Two separate cross-industry findings

~3 in 4

boards perceived to have only moderate or limited AI expertise

Cross-industry, KPMG (2026)

3x

more likely to strongly agree that senior leaders demonstrate ownership and commitment

AI high performers, McKinsey (2025)

Source: KPMG and INSEAD (2026); McKinsey (2025).

Takeaway. Give directors enough understanding to govern AI, not to build it: assess the gap, build access to expertise, embed AI in the oversight model, keep the education current and have the most senior people visibly own it.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST10

RECOMMENDED ACTIONS

1. **Assess the board's AI understanding** against what effective oversight of the firm's AI requires, then repeat periodically.
2. **Build access to expertise** through a standing education slot, an advisory council or a director with AI expertise.
3. **Embed AI oversight in the governance model** with assigned ownership, board touchpoints and a board-level governance framework.
4. **Make the education recurring** and keep it current with agentic and other evolving risks.
5. **Have senior leaders own and role-model AI use** so the education embeds.

SEQUENCE AND DEPENDENCIES

1. Sequencing matters: the board first assesses its current AI knowledge against what oversight of the firm's AI requires, proportionately to the firm's size, sector, strategy and technological maturity.
2. Baseline AI literacy for all staff is P06-ST01 and training tailored by function and risk level is P06-ST04; the accountability obligations the education serves are P05-ST02.
3. Building access to expertise (the levers above) is what makes the assessment actionable, and the guide treats it as a deliberate design choice rather than a reaction to an incident.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Benefits indirectly: a board that can challenge AI decisions sets clearer expectations for how AI is used and governed in the business.
Second Line	Briefs and educates the board on evolving AI risk, and depends on an AI-literate board to support proportionate control rather than defer to vendor assurances.
Third Line	Provides independent assurance that board oversight of AI is real and informed; the audit capability for AI is developed in Pillar 7 and cross-referenced here.
Board and Executive	Owns the education mandate: assess the gap, build access to expertise, embed AI in the oversight model, and role-model AI use so oversight is exercisable.
Cross-Functional	The board education program is jointly served by the corporate secretary, risk, technology and external advisors.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST10

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **False comfort from shallow briefings**	Replace passive briefings with recurring education calibrated to the oversight task; assess the board's ability to challenge, not just its attendance.
HIGH **Deferred ownership**	Agree who owns AI oversight, set clear management accountability and board touchpoints, and make the board's own capability development a standing governance item.
MEDIUM **Expanding oversight gap**	Provide continuous board education and extend it to evolving, agentic AI risks rather than treating it as a one-off.
MEDIUM **Education without leadership ownership**	Have executives learn hands-on, then visibly own and model policy-compliant AI use; treat leadership behavior as part of the program.

METRICS SCORECARD

AREA	LEADING	LAGGING
Board challenge competence	Directors who have completed a documented case-based challenge exercise on a real firm AI use case	Share of material gaps found in the exercise that are closed by the next exercise
Oversight effectiveness	Management AI decision papers that set out alternatives, evidence and a named escalation owner	Material board challenges that led to clarification, a changed decision or a documented risk acceptance
Educational currency	Board refresh scheduled after each material change in AI use, risk or regulation, including agentic systems	Next challenge exercise reflects the new exposure
Leadership modeling	Senior-team demonstrations of approved AI tools using synthetic or cleared information	Whether commitment is owned or only announced

BY FIRM SIZE

BOUTIQUE

Constraint: In this guide's view, a small board makes it hard to add a dedicated AI-expert director, and the AI footprint to oversee is simpler but still real.

MID-SIZED

Opportunity: In this guide's view, a mid-sized board is large enough to fund a recurring program and small enough to run it through existing meetings; Constraint.

LARGE AND GLOBAL

Opportunity: In this guide's view, scale supports dedicated board AI capability; Do Now: Consider an AI-expert director or a board technology and AI committee.



Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST10

AMAT

P06-ST11

Hiring and Sourcing AI Talent: Winning Scarce People When Everyone Wants Them

For: People and Change; Senior Leadership

Asset managers are competing for the same scarce AI talent, and many cannot simply outbid larger firms and big-tech employers. The shortage of hybrid talent, people who combine technical skill with real investment and market understanding, is holding back safe AI adoption. A firm does better not by paying the most but by being deliberate: mapping the scarce skills it needs, choosing build, borrow or buy for each, competing on a genuine value proposition rather than salary alone and pooling the scarce talent it has. External sourcing does not transfer accountability, so partner and vendor talent must meet the same competence bar as internal hires.

11 months

average time for financial-services firms to fill AI roles requiring regulatory knowledge, against 5 months for a general machine-learning role (one consultancy's citation of LinkedIn data, 2026)

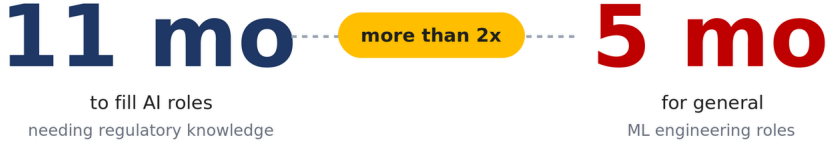
36%

of financial-services companies say total rewards need redesigning (2026)

WHERE FIRMS GO WRONG

Firms try to win talent on salary alone, which runs into a reward compression many cannot sustain. They overlook how geography may limit the talent they can draw on, and they blur sourcing with neighboring questions of which roles should exist and how many people the firm needs, so they argue about headcount when the real bottleneck is a proposition that attracts few candidates.

In one assessment, AI roles that need regulatory knowledge take over twice as long to fill



Source: The Thinking Company (2026), a readiness guide for banks and insurers, citing LinkedIn's 2025 Global Talent Report; financial-services AI hiring.

Takeaway. Map the hybrid skills the firm genuinely needs, choose build, borrow or buy for each, pool scarce talent and compete on a real value proposition, keeping the competence bar and accountability with the firm however people are sourced.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST11

RECOMMENDED ACTIONS

1. **Map the scarce hybrid skills** the firm needs and treat filling them as an AI-governance priority, not routine hiring.
2. **Choose build, borrow or buy** for each scarce skill, building toward a shared specialist pool that allocates scarce talent across the firm.
3. **Compete on a genuine value proposition**, not salary alone, building toward redesigned total rewards.
4. **Hold partner and vendor talent** to the same competence bar as internal hires, evidenced in proportion.
5. **Pipeline the slow, regulated-knowledge roles early** and cover the gap with interim or partner capacity.
6. **Match the sourcing play to firm size and location**: borrow and partner where scale and geography are against you.

SEQUENCE AND DEPENDENCIES

1. In this guide's view, start with the few workflows blocked by missing capability and decide, for each gap, whether to develop the skill.
2. Before signing with a partner, name the business owner and the cost of reviewing the work, and agree the work product, acceptance test, handover materials and exit date.
3. Cover a vacant critical seat as RA 5 sets out, and review the mix after the first delivery cycle.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the demand: defines the scarce hybrid/T-shaped skills each team needs, sponsors the build/borrow/buy choice.
Second Line	Challenges whether critical roles and externally sourced specialists have enough competence, independence and oversight, and escalates material gaps.
Third Line	Tests whether governance-critical roles are being filled and pipelined rather than left open.
Board and Executive	Owns the talent-sourcing strategy as a strategic risk, funds the value proposition and the shared specialist pool.
Cross-Functional	New-role definition (P06-ST05), workforce sizing and capacity (P06-ST09), technology build-versus-buy (P04-ST08) and the accountability regimes (P05-ST02) are cross-cutting and are referenced here, not re-derived.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST11

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Skills gap stalls safe adoption**	Map the scarce hybrid / T-shaped skills and treat filling them as an AI-governance priority, not routine hiring.
HIGH **Losing the pay war / reward compression**	Compete on a genuine EVP (mission, growth, fair and transparent reward), not salary alone; redesign total rewards.
HIGH **Concentration leaves small firms exposed**	Borrow and partner deliberately, and build the minimum internal capability to evaluate vendors rather than depend on them blindly.
HIGH **Key-person and partner dependence**	Name a deputy, keep a minimum handover pack (access, tests, runbooks), and require knowledge transfer before an engagement ends.
MEDIUM **Slow-to-fill governance-critical roles**	Pipeline the regulated roles early and cover the gap with interim or partner capacity; name a qualified interim owner and limit the affected use until review is competent.
MEDIUM **Geography may limit access to AI skills**	Use remote hiring, partnering and vendors where geography is against you, and locate teams near supply where you can.

1 more risk in the full article

BY FIRM SIZE

BOUTIQUE

A boutique is unlikely to win a pay war for scarce AI talent and, where local AI skills are scarce, may find them harder to reach.

MID-SIZED

A mid-sized firm has a real opportunity: it is large enough to fund a few targeted hybrid hires and a shared specialist pool.

LARGE AND GLOBAL

A large firm carries the advantage: AI expertise is concentrated in larger organizations, as the Fasken summary reports, so it can build a talent brand.

AMAT

P06-ST12

Managing Fear, Resistance and Trust: Making It Safe Enough to Adopt

For: People and Change; Senior Leadership

Many AI adoption problems in asset management turn on fear, resistance and trust, not only on technology. People are less likely to lean into a tool they are afraid of, do not trust or believe is there to replace them, and rollout discipline alone is unlikely to overcome that. An early priority is psychological safety, the belief that one can ask questions, experiment and admit mistakes without negative consequences; in one self-reported survey of 2,257 consulting-firm employees it was associated with whether people adopted AI tools at all. The recommended approach builds that safety, addresses displacement fear honestly, makes the tools trustworthy and works with professional identity, treating resistance as a signal to be read.

30

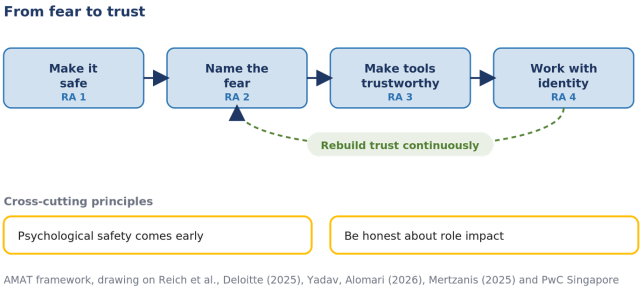
average layoff concern score (0 to 100) among surveyed hedge fund investment professionals in 2026, the highest since tracking began

44%

of 100 surveyed US buy-side analysts and PMs trust AI summaries as much as sell-side written summaries

WHERE FIRMS GO WRONG

Firms treat resistance as an obstacle to override rather than a signal to read. They leave displacement fear unspoken, where it can harden into quiet resistance, and they ignore professional identity and legacy culture, so a change that threatens what people are proud of is resisted even when it is rational.



AMAT framework, drawing on Reich et al., Deloitte (2025), Yadav, Alomari (2026), Mertzanis (2025) and PwC Singapore

Takeaway. Build psychological safety, name displacement fear and answer it with an honest role-impact statement and funded retraining, make the tools trustworthy and work with professional identity. A concern that is heard builds trust; one that is overridden goes quiet and erodes adoption.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P06-ST12

RECOMMENDED ACTIONS

1. **Build psychological safety early**, with leaders openly sharing their own AI mistakes and questions.
2. **Name displacement fear openly** with an honest role-impact statement and funded retraining.
3. **Make the tools trustworthy** through transparency, data protection and safeguards, and where feasible phased trials with results shown to users.
4. **Work with professional identity**: position AI as reinforcing expert judgment and work with senior professionals, not around them.
5. **Treat resistance as a signal**: invite confidential concerns, name a responder for each and report back.

SEQUENCE AND DEPENDENCIES

1. Psychological safety is an early priority, since employees who felt less psychologically safe were less likely to have adopted AI tools, so in this guide's view invest in safety.
2. This article is the fear, resistance and trust layer; it must not absorb its neighbors.
3. And because, in this guide's view, trust is built by evidence rather than assertion, limited trials with results shown to users should, where feasible.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Is likely to feel the displacement fear most directly; needs psychological safety to experiment and honest role-impact messaging backed by funded retraining.
Second Line	Depends on staff surfacing AI errors and doubts rather than hiding them, which requires psychological safety, and ensures trust-building does not overstate the tools' reliability.
Third Line	Tests whether the issues process, the approved-tool boundary and the evidence behind trust claims operate as described, using aggregated results where employee confidentiality applies.
Board and Executive	Sets the tone that makes it safe to raise concerns, funds real retraining behind an honest role-impact message, and treats trust as a governance issue.
Cross-Functional	The fear-and-trust work is jointly owned by People and Change, the business and Technology, coordinated with the P06-ST03 adoption program and P06-ST09 workforce planning.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST12

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Overriding resistance**	Treat resistance as a signal; invite confidential concerns, assign a responder, and report what changed or why a concern cannot be resolved; rebuild trust rather than pushing harder.
HIGH **Unaddressed displacement fear**	Name the fear openly, state what has and has not been decided on roles, and back it with funded retraining and transition support (P06-ST09).
HIGH **Unapproved experimentation**	Limit trials to approved tools and permitted data; require human review for material output, encourage prompt reporting of mistakes, and let the named owner suspend use.
MEDIUM **Distrust of the tools**	Invest in transparency, data protection and safeguards; where feasible, phase rollouts and show trial results to users.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Psychological safety	Team psychological-safety survey scores	Rate of good-faith AI issues reported and closed, tracked with confidentiality safeguards
Displacement fear	Affected roles covered by a published role-impact statement and a route for confidential questions	Change in a pulse score on understanding of role impact, after explanations and support
Tool trust	Trust in the AI tools among users	Trust as a cited barrier to use
Trust-building	Validated error and exception examples disclosed during trial	Sustained use in the approved workflow after trial results are shown to users

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Advantage: In this guide's view, leaders can often address fear and build trust face to face, and identity concerns are visible early; Do Now: Talk about displacement openly.

MID-SIZED

Opportunity: In this guide's view, a mid-sized firm is large enough to fund a structured trust program and small enough to coordinate it across functions and move quickly; Constraint.

LARGE AND GLOBAL

Opportunity: In this guide's view, scale supports structured psychological-safety and trust programs, phased rollouts and identity-aware change; Do Now: Run structured trust-building with safeguards and phased demonstrations.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P06-ST12

Pillar 7 Executive Summary

For: Risk and Compliance; Internal Audit; Technology · Also relevant to: Senior Leadership, Legal, Operations, Investment Teams

THE ARGUMENT

AI in regulated asset management must be tested, validated, audited and overseen across its lifecycle and its third-party estate, and Pillar 7 is the assurance that confirms the controls are operating: testing and monitoring, independent validation and audit, output, adversarial and bias checks, evidence trails, resilience, vendor oversight and inspection readiness. Where Pillar 5 defines the obligations, Pillar 7 is the evidence they are met. The lesson across the pillar is that assurance counts only when it can change an outcome and be shown on demand: a test that catches nothing, a validation that always passes, or a control no one can evidence gives no assurance.

>=1 in 10
samples hallucinated on three financial-language tasks (2025 survey)

14 of 23
audit experts interviewed named poor documentation the top auditability factor

~30
US advisers the SEC examined on AI in 2023; most lacked full policies (GAO)

WHY NOW

Assurance is moving from good practice to supervised expectation, most explicitly in the US: the SEC has named AI in its 2026 examination priorities, SEC officials told the GAO that most of the roughly thirty advisers examined in 2023 lacked comprehensive AI policies, and the SEC has charged advisers for AI washing. Australia's APRA, in a targeted review of selected large banks, insurers and superannuation trustees, called point-in-time assurance ill-suited to adaptive AI and found internal audit skills lagging, a benchmark for asset managers elsewhere rather than a rule. ESMA has set a 2027 supervisory priority that starts with how firms use AI, and Singapore's MAS has issued final AI risk management guidelines, not yet in force. From December 2027 the EU AI Act attaches conformity, documentation and retention duties to providers of high-risk systems.

RISKS TO AVOID

Assurance theater: tests, validations and audits that never change anything, a binder in place of evidence, controls that look present but do not bite. **Point-in-time assurance for a moving system**: testing once and assuming it holds, missing drift, vendor changes, new attack classes and silent degradation. **Unprovable governance**: controls that cannot be evidenced on demand, stale documentation, disclosures that outrun the systems.

Bottom line. Pillar 7 is the difference between a firm that says its AI is governed and one that can prove it: working, rehearsed controls, each sized to the stakes, rather than paperwork. Build the assurance now, as evidence, or discover under examination that the controls you claimed cannot be shown to work. Open limits: the SEC's 2027 priorities are unpublished, and no named asset manager's internal audit work on AI, production evidence pack or bias-testing program is documented in the sources.

Chapter Leads, Pillar 7: Declan Sheehy

Explore Pillar 7 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P07-summary

PRIORITY MOVES FOR LEADERSHIP

1. **Test before deployment and monitor after, as one lifecycle control**: set evidenced acceptance criteria before go-live, then monitor for drift and re-validate on material change.

2. **Make validation and audit independent and able to change outcomes**: give validators the standing to press for change or constrain use, with a recorded route for any override, and internal audit a wider scope and a faster clock, co-sourced around an internal chief audit executive where the firm can support one.

3. **Engineer output, adversarial and bias defenses into the systems**: ground and verify outputs, red-team the full kill chain, and run bias tests with counterfactuals and, where the firm has model access, representation-level analysis; treat control vectors as a promising lever to validate on the firm's own application.

4. **Keep an inspection-ready evidence trail and rehearse the exam**: capture audit trails, oversight records and accurate disclosures as a by-product, key the evidence to the AI inventory, and rehearse mock inspections with interviews.

5. **Extend assurance across the third-party estate and design for failure**: write provider-initiated change into vendor contracts, and build a fallback, a rehearsed stop with a named holder and deputy, and continuity for the providers you depend on. This pillar holds the guide's shared stop-and-revoke definition.

6. **Size assurance to the stakes, with a named owner for each system**: scale testing, validation and QA depth to what a failure can cost. Accountability sits with the firm's senior leadership everywhere, and is also personal where a regime such as SM&CR applies.

IN THIS PILLAR

P07-ST01	Testing AI Before Deployment	204
P07-ST02	Monitoring AI After Deployment	207
P07-ST03	Internal Audit Approach to AI	210
P07-ST04	Third-Party Vendor Due Diligence	213
P07-ST05	Independent Model Validation and Challenge	216
P07-ST06	Evidence, Documentation and Audit Trails	219
P07-ST07	Cybersecurity Testing and Adversarial Risk	222
P07-ST08	Data, Prompt, and Output Quality Assurance for AI	225
P07-ST09	Bias, Fairness and Unintended Consequences Testing	228
P07-ST10	Resilience, Fallback, and Business Continuity for AI	231
P07-ST12	Regulatory Inspection Readiness for AI	234

Explore Pillar 7 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P07-summary

P07-ST01

Testing AI Before Deployment

For: Technology | Risk and Compliance

The question is what "tested" means for an AI system before it goes into production. For a probabilistic system that can produce a fluent, confident and wrong answer, the familiar pass/fail of user acceptance testing is not enough. A pre-deployment campaign should contain explicit ground-truth acceptance criteria rather than a convincing demo, tests of what the system refuses to answer as well as what it answers correctly, and validation that holds out of sample rather than in the backtest. It should end in a documented, risk-differentiated review that no higher-risk system crosses without a record.

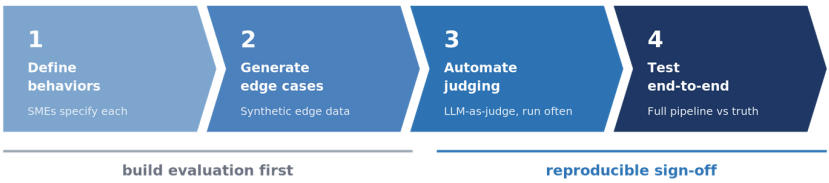
At least 10%

of samples showed factual hallucination on selected financial-language tasks (earlier research repeated by a 2025 survey; not a measured rate for asset-management deployments)

WHERE FIRMS GO WRONG

Firms treat testing as a checkbox, a demo-driven sign-off that never asks how the system behaves at the edges, on adversarial inputs or over time. Three blind spots compound: accuracy tested without boundaries, in-sample validation mistaken for robustness, and confident language that masks uncertainty.

Evaluation-driven development: a four-part pre-deployment test pipeline



AMAT framework, drawing on ZenML / BlackRock (2025).

Takeaway. "Tested" for AI is an evidenced claim that a system does what it should, refuses what it should not and holds up outside the conditions it was built in. Treat evaluation as the discipline development is organized around, not a gate at the end.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P07-ST01

RECOMMENDED ACTIONS

1. **Build the evaluation suite alongside the next system you deploy**, not as a final gate, and measure every change against it.
2. **Define ground-truth test sets** and make "verifiable against a trusted source" a formal acceptance criterion.
3. **Test boundaries as deliberately as accuracy**: out-of-domain and refusal suites, including over-refusal.
4. **Validate trading and signal models** out of sample and walk-forward, against realistic market impact, slippage and liquidity.
5. **Build a human-verification path** for every decision-relevant output, and test that reviewers can override and stop the system.
6. **Stress-test under messy and adversarial inputs**, feeding real production failures back into the suite.
7. **Run a risk-differentiated review**: fast-track lower-risk uses and evidence a sign-off before higher-risk go-live.
8. **Match the model to the task** and validate it against the ground-truth criteria.

SEQUENCE AND DEPENDENCIES

1. Sequence the evaluation asset ahead of the system, not behind it.
2. Keep the pre-deployment boundary clear and hand off cleanly.
3. Match the campaign's depth to the risk and to the model.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own and run the campaign: define ground-truth sets and acceptance criteria, build the evaluation-driven-development pipeline, and stress-test under messy inputs before requesting sign-off.
Second Line	Set the testing standard and the sign-off gate: mandate fact-grounding and boundary acceptance criteria, risk-tier the review, and require a human-verification path.
Third Line	Test that testing happened: verify test evidence, coverage and sign-off records rather than accepting that a system was validated.
Board and Executive	Require that no higher-risk AI reaches production without an evidenced testing sign-off, and fund the test infrastructure before an incident forces it.
Cross-Functional	Agree what “tested” and “done” mean before scale, and keep the pre-deployment boundary explicit: this article ends at go-live, and ongoing monitoring is ST02.



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST01

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Boundary failure**	Out-of-domain and refusal-boundary suites with refusal-rate and over-refusal measures; boundary enforcement as a formal acceptance criterion.
HIGH **In-sample illusion / overfitting**	Untouched final test window, sequential out-of-sample validation and look-ahead and survivorship checks before any signal goes live.
HIGH **Unrealistic execution assumptions**	Test against realistic market impact, slippage and liquidity; reject frictionless-only validation.
MEDIUM **Overconfident or unsupported outputs**	Test calibration, supported rationale and abstention before release; test that reviewers spot and reject unsupported claims.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Test coverage	Share of in-scope AI use cases with a defined ground-truth acceptance test set	Production incidents traced to a path never tested pre-deployment
Factual accuracy	Fact-grounding pass rate on the ground-truth set before release	Post-deployment factual-error incidents in decision-relevant outputs
Boundary enforcement	Share of out-of-domain and refusal-boundary tests the system correctly declines.	Boundary breaches in production (answered when it should have declined)
Out-of-sample robustness	Out-of-sample / walk-forward degradation versus in-sample performance	Live-versus-backtest performance gap after deployment

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Few models keep the testing burden small, but the firm depends on vendor systems it cannot fully inspect, so require the vendor's evaluation evidence.

MID-SIZED

Capacity squeeze

The model estate is scaling faster than test capability, so testing becomes ad hoc and inconsistent; stand up a shared evaluation harness and ground-truth sets and a risk-differentiated review process.

LARGE AND GLOBAL

Scale justifies building testing as a platform, an end-to-end framework every contributing team uses and a stress-testing environment that feeds messy, conflicting inputs and captures execution records.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST01

AMAT

P07-ST02

Monitoring AI After Deployment

For: Technology | Risk and Compliance

A model that passed every pre-deployment test can still be wrong next quarter, because the conditions it was validated against move. The central shift is from point-in-time assurance, a validation sign-off followed by a quarterly committee review, to continuous monitoring, because an adaptive system can drift materially between those cycles. The practical answer is a layered capability covering statistical drift, behavioral telemetry and fairness, with drift thresholds that trigger an owned action. Regulators in the US, EU and Singapore are converging on this lifecycle expectation, and the hard part is organizational rather than technical.

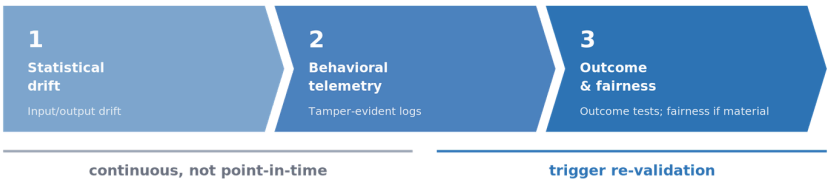
51%

year-over-year rise in recorded AI incidents in 2025 (AI Incident Database)

WHERE FIRMS GO WRONG

The dominant failure is treating deployment as the finish line and governing the model through the same periodic committee cadence as everything else. Firms also assume what passed once stays passed, when a prompt change or vendor update can shift behavior with no code change on the firm's side, and they watch uptime and latency while output quality degrades.

The three-layer continuous monitoring model



AMAT framework, drawing on Chen and Telang (2026), adapted to asset-management uses.

Takeaway. Passing the pre-deployment test is the start of assurance, not the end. Monitor continuously across three layers, set thresholds that trigger an owned action and manage each model across its lifecycle to retirement.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P07-ST02

RECOMMENDED ACTIONS

1. **Replace point-in-time sign-off with continuous monitoring**, starting with drift and output-quality monitoring on higher-risk and client-facing models.
2. **Define drift thresholds** that trigger an owned action rather than a dashboard nobody owns.
3. **Run hallucination and output-quality regression** on every release and every vendor change notice.
4. **Monitor asset-management-specific behavior**: mandate adherence, style drift, execution behavior and client-outcome fairness.
5. **Capture behavioral telemetry** in tamper-evident logs alongside infrastructure metrics.
6. **Feed production failures** and edge cases back into the pre-deployment suite.
7. **Assign monitoring ownership** across the three lines and move toward real-time surveillance.
8. **Manage the full lifecycle**, re-validating on a risk-based cadence and decommissioning models whose drift cannot be controlled.

SEQUENCE AND DEPENDENCIES

1. Start by turning monitoring into decisions, not displays.
2. Keep the boundaries clean so effort is not duplicated.
3. Before a vendor-enabled AI feature is switched on inside an existing platform, record its intended use, owner and risk tier in the AI inventory (P05-ST10).

WHO DOES WHAT: LINES OF DEFENSE

First Line	Run the monitoring: operate drift detection, behavioral telemetry and CI/CD regression, act on threshold breaches.
Second Line	Set the monitoring standard and the thresholds, define what triggers re-validation, and own control-effectiveness testing; the independent challenge of the model is ST05.
Third Line	Test that monitoring works: verify coverage, thresholds and that breaches produced action, using the telemetry and logs.
Board and Executive	Require continuous monitoring of higher-risk AI and evidence that assurance is no longer lagging deployment; fund the shift from periodic committee to standing surveillance.
Cross-Functional	Agree thresholds, ownership and re-validation cadence across the lines, and keep the boundary explicit: monitoring is the control lens, the observability platform is P04-ST09.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST02

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Point-in-time assurance ill-suited**	Continuous, cross-domain assurance with capable second and third lines and proportionate lifecycle assessment.
HIGH **Silent quality regression**	Re-run the firm's regression benchmarks on every release and every vendor change notice and on a fixed cadence, and compare with the prior version.
HIGH **Mandate / style drift**	Periodic mandate-adherence and style-drift testing with controls against adverse client and market outcomes.
MEDIUM **Monitoring the wrong signals**	Start with the signals the workflow exposes (sampled output accuracy, user correction rates, failed calls, vendor version identifier); where the firm controls the application, add token use.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Drift detection	Share of production models with statistical drift detection and defined thresholds	Drift events that reached client or market impact before detection
Output quality	Continuous hallucination / quality-regression pass rate in CI/CD	Post-deployment output-quality incidents after a model, prompt or data change
Mandate adherence	Frequency of mandate-adherence and style-drift testing per model	Mandate or style breaches attributable to model drift
Control effectiveness	Median time from alert to owner acknowledgment and to containment on higher-risk workflows	Confirmed material errors that reached a client or market before containment, and repeat incidents after remediation

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

For low-impact vendor-supported tools, use documented periodic output sampling and change checks; where a single workflow is autonomous or client-facing.

MID-SIZED

Capacity squeeze

The model estate has outgrown periodic review but the platform to automate monitoring is not yet in place, so prioritize continuous monitoring on the highest-risk and client-facing models first.

LARGE AND GLOBAL

Automate continuous monitoring for higher-risk workflows across business lines, while retaining lighter, documented monitoring for low-risk uses.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST02

AMAT

P07-ST03

Internal Audit Approach to AI

For: Internal Audit

Internal audit is being pulled into AI assurance from three directions at once: supervisors expect the third line to be able to assure AI, many audit and risk functions lack the skills and tools to assess agentic and automated systems, and AI is entering the audit process itself. The function becomes both an assurer and a significant user of AI. The response is to audit at three levels, the model, the system around it and the governance that oversees it, to size the operating model to the firm, including co-sourcing, and to refresh the plan at the cadence AI actually moves at. The function's own AI must be governed to the standard it demands of others.

WHERE FIRMS GO WRONG

The scope is too narrow, with assurance confined to the model while access controls, monitoring and governance go unaudited. Third-line teams are asked to challenge systems they were not trained to assess, so assurance lags deployment. And annual cycles assume a slow-moving environment, so findings can be stale on delivery.

Three-level scope for internal audit of AI



Source: Gomez et al., Arcadia Impact, 2025.

Takeaway. Internal audit needs a wider scope, a faster clock and the capability to use both: audit at three levels, co-source where depth is missing, refresh the plan at the cadence AI moves and keep a human in the loop on the function's own AI.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST03

RECOMMENDED ACTIONS

1. **Confirm the audit universe against the AI inventory** and expand scope beyond model-level checks, building toward a board-approved plan across all three levels.
2. **Adopt a co-sourced model** with an internal Chief Audit Executive and external specialists for depth.
3. **Differentiate audit frequency by risk layer**, with fast-moving controls semi-annual or quarterly.
4. **Review the AI audit plan at least quarterly** in dynamic environments, with defined triggers.
5. **Set a tiered information-access model** so auditors can reach sensitive technical evidence safely.
6. **Require methodology disclosure** from external AI or bias audits, and decline any provider will not disclose.
7. **Test model logic, not just outputs**, for higher-risk models.
8. **Govern the function's own AI** with an inventory, named owners, auditor-in-the-loop review and periodic independent review.

SEQUENCE AND DEPENDENCIES

1. Sequence the build rather than attempting comprehensive coverage on day one.
2. Independence is the control that makes everything else count, and it is a design choice, not a default.
3. Regulatory context also shapes where audit evidence will be expected.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own AI controls and provide the evidence auditors test; maintain the AI inventory, including shadow projects, so the third line is auditing reality, not paperwork.
Second Line	Coordinate with internal audit on scope and methodology without ceding independence; the third line tests whether second-line AI controls are designed and operating effectively.
Third Line	Owns the AI audit plan across all three levels, sourcing (co-sourced with an internal CAE, or an outsourced start that builds toward one), differentiated frequency.
Board and Executive	Set the mandate and fund the capability; require the CAE to report how internal audit is evolving with AI and to concentrate the plan on higher-risk.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST03

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Model-only assurance**	Mandate a three-level board-approved audit plan.
HIGH **Assurance lags deployment**	Build/co-source capability; confirm the audit universe against the AI inventory before planning.
HIGH **Stale annual cycle**	Differentiate frequency by risk layer; review the plan at least quarterly; trigger reviews on vendor and incident events.
HIGH **Governance-technical disconnect**	Corroborate with operational/system evidence via controlled access; for vendor AI, start with contractual evidence and record the assurance limitation.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
AI audit coverage	Share of higher-risk AI models in the board-approved plan.	Share of material AI control failures that audit independently detected or confirmed.
Plan currency	Days since last risk-assessment refresh	Findings traced to stale-plan gaps
Assurance timeliness	Time from AI go-live to first third-line review	Control weaknesses undetected between cycles; median days from a material AI change or incident to audit response.
Audit efficiency	AI-assisted reporting cycle time	Rework from AI errors in audit output

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

Start outsourced/co-sourced with a named owner; co-source one higher-risk deep-dive.

MID-SIZED

Capacity squeeze

Co-sourced with a permanent internal CAE; guard against assurance lagging deployment.

LARGE AND GLOBAL

Largely in-house with continuous auditing; manage model sprawl and assurance fatigue.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST03

AMAT

P07-ST04

Third-Party Vendor Due Diligence

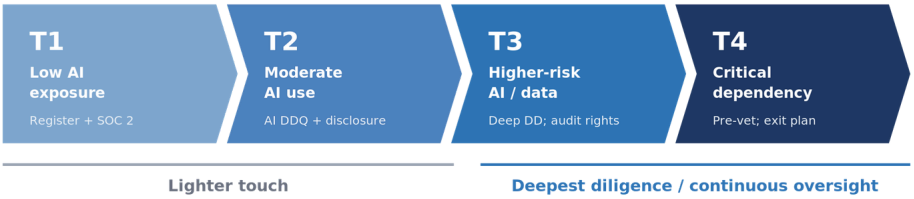
For: Risk and Compliance | Technology | Senior Leadership

Much of an asset manager's AI risk now sits outside the firm, inside its vendors. AI is arriving embedded in the platforms managers already use, often without disclosure, and traditional third-party risk management was not built to see it. Australia's prudential regulator found AI embedded in vendor platforms, opaque upstream dependencies and contracts that lag practice, and the firm, not the vendor, remains accountable. The discipline is to find vendor AI, tier it by risk, put the right terms in the contract and keep testing that they hold, extending the same rigor to fourth parties and foundation models.

WHERE FIRMS GO WRONG

Third-party risk management is still built for a pre-AI world: SOC 2 reports and generic questionnaires cannot show how a vendor uses AI or what data it relies on, so diligent firms come away with a false sense of coverage. Contracts lack the provisions that matter, and accountability sits with the deployer for design decisions made by a vendor.

Risk-based vendor-AI due diligence



Takeaway. Vendor due diligence for AI is a different discipline aimed at a moving target. Find the AI first, tier it, write the visibility and control you need into the contract, and keep testing that the vendor honors it.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST04

RECOMMENDED ACTIONS

1. **Build an AI vendor register**, starting with vendors behind critical and client-facing processes, mapped through to foundation models and fourth parties.
2. **Stress-test existing AI contracts** against five items: model-update notice, audit rights, incident timelines, data-handling triggers and termination or portability.
3. **Risk-tier vendors** by AI type, data sensitivity and business impact, with AI-specific due diligence on the material few.
4. **Set the AI clause standard** for new and renewed contracts, including disclosure, change notice, bias-audit rights, limits on training with your data and indemnification.
5. **Monitor and enforce contracts**: exercise audit rights and set escalation triggers rather than relying on signed paper.
6. **Assess concentration** and verify that exit plans for critical AI providers are executable.
7. **Extend model-risk validation** to material vendor models.

SEQUENCE AND DEPENDENCIES

1. In this guide's reading, the register comes first because contract stress-testing, risk-tiering and concentration assessment all start from knowing the population of vendor AI.
2. Vendor due diligence overlaps the build-buy-partner decision (P04-ST08) and data-vendor strategy (P03-ST10); this sub-topic owns the risk-control view of an existing or prospective vendor relationship.
3. Smaller firms often cannot compel foundation-model providers to accept bespoke audit or notification terms, so proportionality matters.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the vendor relationship and the AI register entry, including embedded AI; gather the AI-specific due diligence and performance evidence.
Second Line	Set the risk-tiering method, the due diligence standard and the AI clause standard for agreements; independently challenge tiering decisions, exceptions and accepted gaps.
Third Line	Independently test whether vendor-AI oversight is designed and operating effectively; ST03 is the audit view, this sub-topic is the control it tests.
Board and Executive	Set concentration risk appetite and hold accountability for outsourced AI, personal where a regime such as SM&CR applies; the firm, not the vendor.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST04

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Hidden vendor AI**	Contractual AI-disclosure + active discovery (register + scanning).
HIGH **Opaque fourth-party chain**	Map register through to foundation models and fourth parties; document opacity.
HIGH **Contract gaps**	Stress-test every AI contract against the five-item checklist; negotiate material gaps first, with compensating controls and a recorded risk acceptance where terms cannot be changed.
HIGH **Inherited AI risk**	AI-specific due diligence and risk-tiering; require control evidence pre-onboarding.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Discovery	Share of AI vendors mapped to foundation models and fourth parties, or with a documented gap.	Vendor-AI incidents traced to undiscovered / undisclosed AI
Contract coverage	Share of critical AI contracts carrying the five required provisions	Findings caused by missing contract provisions
Risk-based diligence	Share of high-tier AI vendors with AI-specific due diligence completed	High-impact vendor AI onboarded without AI-specific DD
Enforcement	Audit rights exercised on critical AI vendors per year	Service-level or compliance breaches undetected until failure

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Start with a named owner, an AI register, approval before new AI features are used and an executable disablement route.

MID-SIZED

Capacity squeeze

Risk-tier and run AI-specific DDQs; guard against sprawl outpacing the function.

LARGE AND GLOBAL

Dedicated TPRM with AI addenda and pre-vetting at scale; manage fourth-party opacity and concentration.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST04

AMAT

P07-ST05

Independent Model Validation and Challenge

For: Risk and Compliance | Technology

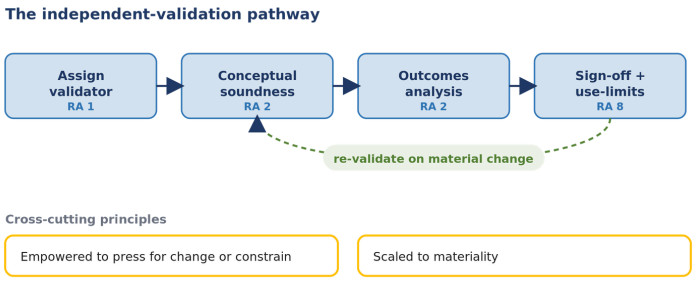
The question is who tested an AI model, and whether they were independent enough and had the standing to press for changes or constrain its use. Validation by the build team lacks that independence. The discipline that matters is effective challenge: expert, independent and empowered critical review across the model lifecycle, scaled to the damage the model can do. A validation should cover conceptual soundness, outcomes analysis and ongoing monitoring, across the model, the technical systems around it and the governance over it. With no settled standard for AI validation, firms must document their own methodology and equip validators to probe systems they cannot simply read.

28%

of surveyed firms had a program to test or validate the outputs of the AI tools they use, 10 points more than in 2024 (self-reported, 2025)

WHERE FIRMS GO WRONG

Firms stand up a validation step but staff it from the team that built the model, or give it no standing to constrain a deployment, so the model ships regardless. They also treat a model evaluation as a full validation, leaving surrounding organizational controls unexamined, and accept a generic benchmark score in place of testing on the firm's own investment task.



Takeaway. Validation is worth the name only when the person doing it is independent, expert and able to press for change or constrain use. Engineer independence into the second line, scale challenge to materiality and treat an honest "cannot validate" as a finding.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P07-ST05

RECOMMENDED ACTIONS

1. **Make validation independent and empowered**, in the second line, with authority to press for changes or constrain a deployment.
2. **Validate against three components**: conceptual soundness, outcomes analysis and ongoing monitoring.
3. **Right-size validation to materiality**, scaling the depth of challenge to the model's impact and risk tier.
4. **Validate all three layers**: model behavior, surrounding technical systems and governance, including human review.
5. **Build model-challenge tooling**, benchmarking against alternatives and re-testing on your own investment tasks, never accepting a generic benchmark score instead.
6. **Staff multidisciplinary validation** with reviewers who did not build the model.
7. **Document methodology**, roles, authorization chains, escalation paths and thresholds.
8. **Treat un-validatable opacity as a finding**: constrain use, add monitoring or reject the model.

SEQUENCE AND DEPENDENCIES

1. The failure that recurs is a validation function that reports to, or is staffed from, the team it is meant to challenge, which no amount of interpretability tooling will fix.
2. This sub-topic is the independent-challenge lens: the model risk management framework and its governance sit in P05-ST04.
3. Validation is normally completed before first use, but for AI it cannot stop there, because model-level assurance degrades and the model drifts.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Build validatable models: document design, assumptions and limitations for conceptual-soundness review, and provide evidence, but do not self-certify; validation is someone else's job.
Second Line	Own independent validation and effective challenge: run the three components, right-size to materiality, staff multidisciplinary teams, and hold the authority to press for changes or constrain.
Third Line	Assure the challenge: test that validation was genuinely independent, covered all three components and had teeth, rather than confirming a report exists.
Board and Executive	Require independent, empowered validation for higher-impact AI, see every decision to proceed over a validator's objection on a higher-impact model.
Cross-Functional	Agree independence, risk-tiering and what "validated" means across the lines, and keep the boundary explicit: this is independent challenge, the MRM framework is P05-ST04.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST05

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Model-only validation**	Validate all three layers (model, technical systems, governance), including single-model dependence and tool permissions; re-validate as model-level assurance degrades.
HIGH **Opacity blocks conceptual soundness**	Benchmarking and structured re-testing for conceptual soundness, interpretability tooling where the model can be inspected; constrain use where opacity cannot be reduced.
HIGH **Generic-benchmark over-reliance**	Independent, investment-specific validation; never substitute a generic benchmark score for the firm's own investment-specific test.
MEDIUM **Validator competence gap**	Multidisciplinary validation teams (technical, domain, compliance, ethics); invest in challenge capability.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Challenge effectiveness	Share of material validation findings resolved or risk-limited before go-live	Models deployed with unresolved material findings; material failures in production that validation should have caught.
Independence	Share of higher-impact models validated by a function independent of the build team with authority to press.	Deployments that proceeded over an unresolved validation objection
Component coverage	Share of higher-impact use cases with conceptual soundness and outcomes analysis complete and an approved production monitoring plan.	Production failures traced to a skipped validation component, and material failures that a monitoring alert or planned re-validation.
Right-sizing	Alignment of validation depth to model materiality/risk tier	High-impact models found to have received only light-touch validation

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

A separate in-house validation function is usually unaffordable, but independence can be bought, so commission external independent validation for high-impact and client-facing models and right-size the depth to each model's.

MID-SIZED

Capacity squeeze

An in-house validation capability is forming but its independence is frequently nominal, with the same small team building and reviewing, so the priority is organizational separation and empowerment.

LARGE AND GLOBAL

A dedicated second-line validation function exists, so the risk shifts from capacity to culture, a process that produces validation reports without genuine challenge.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST05

P07-ST06

Evidence, Documentation and Audit Trails

For: Internal Audit | Risk and Compliance

When an AI system causes a loss, the first question is not what went wrong but whether the firm can show what happened, and many cannot. When an auditor or a court asks which version was live, which data influenced it and who approved its release, the governance record is often missing, because evidence was reconstructed after the fact from decks and manual reports disconnected from the code. The fix is to produce evidence as a byproduct of AI pipelines rather than assemble it ex post. That means a tamper-evident audit trail binding technical artifacts to the governance decisions behind them, machine-readable evidence, living model cards and a design that scales to production volume.

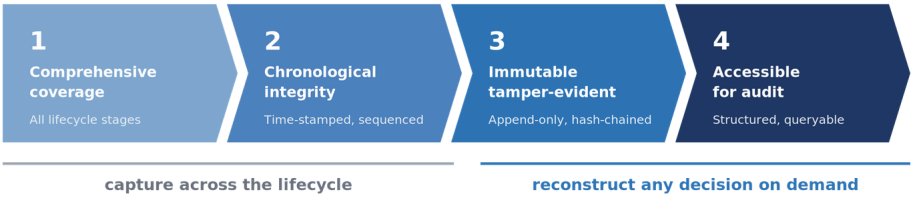
14 of 23

practitioners interviewed on AI auditability cite inadequate documentation as the most frequent factor (61%)

WHERE FIRMS GO WRONG

Firms treat compliance as a documentation exercise rather than an evidence problem, producing registers and decks that describe an intended system rather than the one that ran. They also trust static model cards to stay true as the system drifts, and expect an audit trail to explain causes when it evidences only sequence.

What an AI audit trail must be (four properties)



Takeaway. Passing the evidence test is an engineering choice made early, not a document written late. Build the trail, generate machine-readable evidence as a byproduct, keep cards living and keep the record of what happened distinct from the explanation of why.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P07-ST06

RECOMMENDED ACTIONS

1. **Build a tamper-evident audit trail** binding technical artifacts to the governance decisions behind them, starting with higher-risk models.
2. **Record the minimum at each lifecycle stage**, from training-data provenance to the approved-version-to-deployment binding, and vendor model versions and configurations.
3. **Generate evidence as a pipeline byproduct**, version-controlling the assessment plan alongside the code.
4. **Adopt a recognized machine-readable evidence standard** linking controls to the policy and risk each satisfies.
5. **Produce a per-run audit trail** for LLM outputs that feed decisions, with reviewer decisions and reasons.
6. **Keep model cards and datasheets living**, tied to the trail, and retire the point-in-time template.
7. **Document model design, assumptions and limitations** to a standard that meets recordkeeping and supports external audit.
8. **Plan for audit-trail scale** with schema, retention tiers and query tooling sized for production volume.

SEQUENCE AND DEPENDENCIES

1. One published proof of concept (a small library, confidentiality out of scope) suggests a tamper-evident trail can be added as a small, reusable layer within existing workflows.
2. Sequence documentation reform around the pipeline, not the auditor's calendar.
3. Coordinate ownership so the trail is not orphaned.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Produce the evidence: instrument capture in the pipeline, register artifacts and approvals as trail events.
Second Line	Set the documentation and trail standard, define the firm's minimum per-stage capture and the machine-readable evidence format, and own the boundary with explainability (P05-ST09).
Third Line	Rely on and test the trail: confirm it is comprehensive, chronological, immutable and accessible, and use it to reconstruct decisions rather than accepting reconstructed narratives.
Board and Executive	Fund evidence infrastructure from the outset, not after an incident; require that higher-risk AI decisions are reconstructable and that audit-preparation cost is falling.
Cross-Functional	Agree schema, retention and "complete" across the three lines before scale; keep the record-of-what-happened (here) distinct from the explanation-of-why (P05-ST09).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST06

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Reconstruction gap**	Tamper-evident audit trail binding artifacts to governance decisions; the firm's minimum per-stage capture.
HIGH **Documentation as an afterthought**	Generate machine-readable evidence as a pipeline byproduct; version-control the assessment plan with the code.
HIGH **Stale or drifting documentation**	Tie documentation to the trail; refresh on change; measure coverage against an auditability framework.
HIGH **Governance detail omitted**	Capture governance checkpoints (approvals, conditions, rationale) as first-class trail events.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Trail coverage	Share of higher-risk models with a complete lifecycle audit trail	Incidents where the live version or approver could not be reconstructed
Evidence automation	Share of compliance evidence generated as a pipeline byproduct	Audit-preparation effort per examination
Documentation currency	Share of model cards / datasheets tied to the trail and refreshed on change	Findings of documentation that no longer matches the live model
Reconstruction	Share of sampled material decisions for which an independent reviewer retrieves the complete evidence pack (version, data.	Material incidents and audit requests where the firm could not establish the live version.

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Few models keep the evidence burden small, but the firm depends on vendor documentation it cannot fully verify.

MID-SIZED

Capacity squeeze

Models are scaling faster than evidence infrastructure, so documentation gets reconstructed ex-post and the reconstruction gap widens; begin with a central AI inventory (P05-ST10).

LARGE AND GLOBAL

Many models and continuous operation can generate millions of records and fragmented documentation, so standardize model cards, datasheets and system cards, automate capture.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST06

AMAT

P07-ST07

Cybersecurity Testing and Adversarial Risk

For: Technology | Risk and Compliance

For AI, the security test that matters most is what an attacker can do after slipping in a malicious prompt. Documented attacks on LLM applications now behave like malware, chaining initial access, privilege escalation, persistence, command-and-control, lateral movement and action on objective across a recognizable kill chain. For an asset manager the exposed surface includes retrieval corpora, agent tool access and the older machine-learning models that move money. Assume initial access and red-team the rest, with canary tokens for cheap detection and least-privilege tool access to limit the blast radius. Treat AI cyber risk as part of enterprise cyber resilience, not a separate technical silo.

fifteen of twenty-one

documented attacks on LLM applications in 2025-26 reached four or more kill-chain stages, up from none in 2023 (36 incidents, Feb 2023 to Jan 2026)

WHERE FIRMS GO WRONG

Firms import the wrong mental model, equating prompt injection with SQL injection when it behaves like a general malware-execution mechanism. They look for the attack in the wrong place, since in retrieval systems the document corpus is the attack surface and a single malicious document can be enough. And they overlook older adversarial attacks on the models that already move money.

The promptware kill chain, tested across all stages



AMAT framework, drawing on Brodt et al. (2026); seven stages condensed to four.

Takeaway. Shift the security question from whether an attacker can get in to what they can do once they are in. Test across the full kill chain, put canary tokens and least-privilege access in first, and keep red-teaming continuous.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST07

RECOMMENDED ACTIONS

1. **Red-team the full kill chain**, not just prompt injection, starting with applications that retrieve external documents or touch order and client systems.
2. **Enforce least-privilege tool access per step**, with each agent holding its own identity, a named owner and a tested revocation path.
3. **Deploy canary tokens** in sensitive data and prompts for deterministic detection.
4. **Treat every retrieval corpus with external or user-writable content as an attack surface** and test it with malicious documents.
5. **Test multi-agent configurations directly** and constrain agent-to-agent calls.
6. **Red-team against the agent's action space**, building toward continuous post-deployment red-teaming.
7. **Add jailbreak and data-exfiltration tests** to the security suite for GenAI applications.
8. **Put AI cyber risk into enterprise stress testing**, resilience programs and board oversight.
9. **Test defenses against AI-enabled impersonation** with independent confirmation, second approvers and deepfake scenarios.

SEQUENCE AND DEPENDENCIES

1. Sequence the controls by cost and value rather than trying to deploy the whole catalog at once.
2. Put the AI security tests where the security function already works.
3. Finally, treat the retrieval and integration architecture as a security-testing input, not a fixed given.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the red-team program, least-privilege tool configuration, canary deployment and RAG source-trust controls; refresh the reference-attack set and test multi-agent topologies directly.
Second Line	Set the standard that AI applications are tested across the full kill chain before deployment; set the release rule for red-team findings.
Third Line	Test whether adversarial AI testing is designed and operating, whether findings are remediated, and whether coverage keeps pace with new exploits.
Board and Executive	Treat AI cyber exposure as a board-level, financial-stability matter; require cyber stress testing that includes AI, evidence that higher-risk applications are red-teamed.
Cross-Functional	Keep P07 testing in step with the P04-ST07 security architecture it tests; coordinate incident response (P05-ST12) and resilience (P07-ST10) so a detected compromise has a rehearsed.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST07

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **RAG / indirect injection**	Test the corpus with malicious documents; source-trust and retrieval-time controls.
HIGH **Lateral movement / one-to-many spread**	Constrain agent-to-agent calls; test multi-agent topologies; per-step tool limits.
HIGH **Persistence and C2**	Canary tokens; memory reset / isolation; monitor for command-and-control patterns.
HIGH **Data exfiltration via jailbreak**	Jailbreak / exfiltration tests in the GenAI suite; tie tests to reachable data classes.

4 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Kill-chain coverage	Share of AI applications red-teamed across all kill-chain stages	Incidents reaching a stage never tested
Least privilege	Share of agents on per-step, least-privilege tool access	Blast radius of the worst simulated compromise
Detection	Canary-token coverage of sensitive data and prompts, with alert routing tested	Mean time to detect a simulated exfiltration
RAG exposure	Share of external/user-writable corpora with source-trust controls and malicious-document testing	Successful indirect-injection findings in production

4 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

A small AI-application estate keeps the attack surface manageable, but the firm relies on vendor security and cannot red-team a vendor model's internals.

MID-SIZED

Capacity squeeze

Retrieval and agentic applications are scaling faster than AI-specific security testing, so the RAG corpora and agent tool-access expand the blast radius before anyone has red-teamed the full kill chain.

LARGE AND GLOBAL

Many integrated AI applications, coding assistants and multi-agent systems create real lateral-movement and one-to-many propagation risk, so run continuous action-space red-teaming, memory isolation and command-and-control monitoring.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST07

AMAT

P07-ST08

Data, Prompt, and Output Quality Assurance for AI

For: Technology | Risk and Compliance

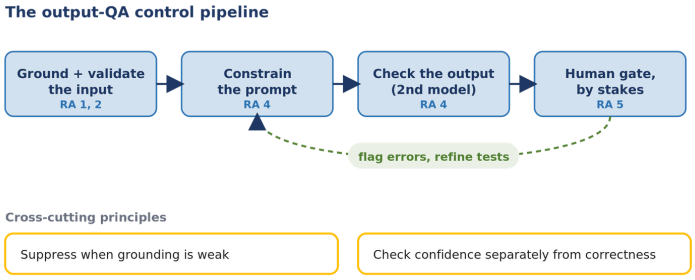
An AI output is not usable until it has been checked, because these systems produce fluent, confident text that is wrong often enough, and convincingly enough, that acting on it unverified is a control failure rather than an efficiency gain. Quality assurance has to cover the data and context that go into a model, the prompts that operate it and, above all, the outputs, with hallucination detection and controls sized to the damage a wrong answer can do. As models improve they hallucinate less often but more convincingly, so reviewer vigilance degrades over time. The answer is systematic testing against a known standard, built as a layered pipeline.

Nearly 60%

of statements from leading AI chatbots were false or misleading in NewsGuard's 2025 analysis, up from roughly 10%

WHERE FIRMS GO WRONG

Firms treat quality assurance as a reviewer glancing at the answer, which cannot catch the property that makes AI output dangerous: confident error. Reviewer vigilance also degrades as models improve, because the remaining errors are harder to spot. Firms underinvest in the input data and context, the prompts and the outputs at the same time.



Takeaway. Stop asking whether the answer looks right and start asking whether it has been checked against something that is. Ground the input, refuse to answer on thin evidence, check the output and gate the high-stakes cases through a human.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P07-ST08

RECOMMENDED ACTIONS

1. **Define ground-truth test sets** for the highest-stakes use cases, building toward a versioned, held-out set for each.
2. **Ground outputs in verified, date-stamped sources** and suppress generation when the grounding is insufficient.
3. **Score confidence separately from correctness**, requiring the model to express uncertainty and decline when it cannot answer from evidence.
4. **Add automated output checks**, from schema validation and figure checks to cross-validation and an independent second model.
5. **Require human review** before any AI output is used or sent externally, with a hard gate for high-stakes outputs and tracking of reviewer override rates.
6. **Require vendors to evidence** their hallucination testing and write the terms into the contract.

SEQUENCE AND DEPENDENCIES

1. Build the ground-truth test sets first, because everything downstream depends on having a standard to test.
2. Give each live AI-assisted workflow, including vendor-enabled AI features, one named owner before it is used.
3. Expect the hardest cases to be open-ended synthesis and non-determinism.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the pipeline: name an owner for each live AI workflow, ground inputs, constrain prompts, run the automated output checks, and review outputs before use.
Second Line	Set the output-QA standard by stakes, require ground-truth testing and confidence checks for high-stakes use cases, and confirm the human gate is real rather than nominal.
Third Line	Assure that quality assurance happened and had teeth: test that ground-truth sets exist, that suppression and gates operate, and, using the review-integrity rate.
Board and Executive	Require output QA for client-facing and investment-decision AI, treat a human review gate that has become a rubber stamp as a failure rather than a success.
Cross-Functional	Agree what "checked" means and who owns each layer across the lines, and keep the boundary explicit: this is output QA.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST08

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Overconfident hallucination**	Check confidence separately from correctness and record confidently wrong answers as their own failure category.
HIGH **Inaccurate output in investment decisions**	Test outputs against a ground-truth set sized to the stakes; place a hard human gate on high-stakes outputs.
HIGH **Hallucination when grounding is weak**	Ground outputs in verified sources; enforce a retrieval-quality threshold that suppresses generation when grounding is insufficient.
HIGH **Acting on unverified output creates liability**	Require verification before use; add a dual verification step and a feedback channel for the most consequential outputs.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Grounding and citation	Share of high-stakes outputs with every material claim traced to a verified source	Incidents traced to an ungrounded output that was acted on
Hallucination rate	Error rate of AI outputs measured against the ground-truth test set	Production errors traced to a use case with no ground-truth set
Production error rate	Share of sampled high-stakes outputs found to contain a material error (a wrong figure.	Material errors that reached a client or a decision
Retrieval quality	Share of responses suppressed or escalated when retrieval grounding is below threshold	Hallucinations traced to generation on insufficient grounding

3 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

An in-house evaluation capability is usually unaffordable, so quality assurance is largely bought with the tool and vendor due diligence is the first control, though not the only one.

MID-SIZED

Capacity squeeze

An in-house capability is forming but cannot yet cover every use case, so the priority is triage: define ground-truth test sets for the highest-stakes client-facing and compliance outputs first.

LARGE AND GLOBAL

Eval harnesses, secondary-model checks and prompt libraries are affordable, so the risk shifts from capability to scale.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST08

P07-ST09

Bias, Fairness and Unintended Consequences Testing

For: Risk and Compliance | Investment Teams

Bias in AI belongs in the assurance program as a testable failure mode that scales. Removing protected data fields and instructing a model to be fair does not remove bias, because models infer sensitive attributes from proxies such as name, address or education; in controlled credit tests reported in 2025, open-source language models reproduced, and at times exceeded, documented human lending disparities. Testing that inspects only outputs will miss this. Effective testing uses counterfactuals, probes for proxies and, where the firm has model access, examines internal behavior. Because competing fairness definitions cannot all be met at once, the firm must choose a standard per use case and keep it current through a lifecycle.

6.5

basis-point human lending gap used as the baseline; four of five open-source models tested exceeded it in controlled credit decisions (2025)

WHERE FIRMS GO WRONG

Firms mistake surface controls for remediation: stripping protected inputs and prompting the model to be fair can mask bias without removing it. Most have never chosen what fairness means for a given use case, though the formal definitions trade off against one another. And they audit only the finished model's outputs, missing bias that entered through historical data and proxy variables.

Common surface controls vs what bias testing requires

Dimension	Surface controls	Counterfactual testing
Protected data	Remove race/gender fields	Test for proxies (name, zip)
Method	Prompt: be unbiased	Counterfactual: vary one attribute
What it sees	Outputs only	Outputs under matched inputs; internal logic where access exists
Fairness standard	Implicit / assumed	Explicit, chosen, documented

AMAT framework, drawing on Cook and Kazinnik (2025) and Krause and Krause (n.d.).

Takeaway. Test deeper than surface controls: use counterfactuals, probe for proxies and interrogate internal behavior where you can. Choose a fairness standard per use case, document the trade-off with legal input and keep it current through a lifecycle.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P07-ST09

RECOMMENDED ACTIONS

1. **Test with counterfactuals**, changing one protected attribute at a time for client-facing and credit-adjacent models.
2. **Choose and document a primary fairness metric** per use case, with legal input, recording the accuracy-versus-fairness trade-off.
3. **Test for proxies** such as name, address and education rather than just removing protected fields.
4. **Run bias management as a lifecycle**, not a one-off pre-launch check.
5. **Diagnose at the representation level** where you have model access, and re-audit any remediation for unintended effects.
6. **Hold bias-audit reporting to a disclosure-ready standard**, and require the same of vendors.
7. **Monitor input data and live outcomes** on a defined schedule, including outcome disparity and override rates by group.

SEQUENCE AND DEPENDENCIES

1. Bias testing is not a task a pure technology team can own.
2. Decide early how the firm will get the data it needs to test.
3. Sequence the work so that fairness is decided before scale, not retrofitted after it.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Own the counterfactual test evidence and the fairness-metric documentation for the models they run; surface proxy findings rather than sanitize inputs and move.
Second Line	Set the fairness-standard policy, choose metrics with legal for each use case, and independently validate fairness; own the disclosure-ready bias-audit standard, including for vendors.
Third Line	Test whether bias controls are designed and operating, that fairness choices are documented, and that remediation was re-audited; follow the model-logic testing cross-referenced from P07-ST03.
Board and Executive	Hold accountability for bias, personal where a regime such as SM&CR applies; require evidence that higher-risk models are tested and that a fairness trade-off.
Cross-Functional	Convene the interdisciplinary review (technical, legal, domain, social) that no single function can staff alone; align the governance calibration to the chosen fairness definition.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST09

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Proxy discrimination**	Test for proxies with counterfactuals and correlate probing, not field removal.
HIGH **Surface fix masks bias**	Test with counterfactuals, and at the representation level where the firm has model access; treat prompt fixes as insufficient.
HIGH **No explicit fairness standard**	Select and document a primary fairness metric and trade-off, with legal.
HIGH **Historical-data bias at scale**	Establish data lineage; rebalance pre-training; monitor input data.

3 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Testing coverage (control coverage)	Share of client-facing and credit-adjacent models with counterfactual bias testing	Bias incidents traced to untested models
Fairness standard (control coverage)	Share of in-scope models with a documented fairness metric and trade-off	Findings where no fairness standard was on record
Proxy exposure (control coverage)	Number of proxy correlates probed per model	Post-deployment disparities from undetected proxies
Remediation quality	Measured disparity reduction after intervention	Side effects introduced by remediation

6 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

A small model estate makes the fairness surface manageable, but reliance on vendor models means the firm cannot inspect internals, so seek contractual bias-audit rights and disclosed methodology.

MID-SIZED

Capacity squeeze

Client-facing AI is scaling faster than fairness governance, so choose and document a primary fairness metric per use case and stand up a small interdisciplinary review (technical, legal.

LARGE AND GLOBAL

In-house teams can run counterfactual and representation-level testing and full disparate-impact analysis directly.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST09

AMAT

P07-ST10

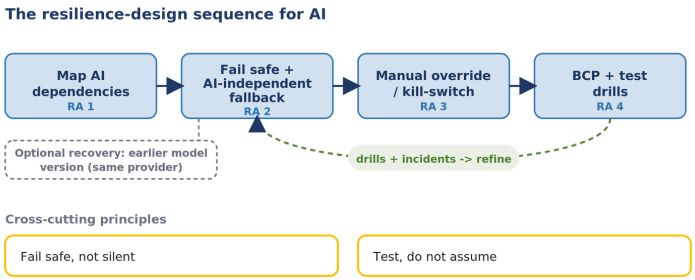
Resilience, Fallback, and Business Continuity for AI

For: Operations | Technology | Risk and Compliance

AI earns a place in critical workflows by being available and reliable, so the question is what happens when it is not. Resilience for AI is a design discipline: build systems that fail safe, keep a fallback that does not depend on the AI, give a human a tested way to stop the system quickly and plan business continuity around the AI dependencies the firm has quietly acquired. That includes failures that come from a provider's decisions rather than the model. Supervisors are already expecting firms to map the dependencies behind important services and to plan for third-party failure, and shared-provider concentration means one outage can reach many firms at once.

WHERE FIRMS GO WRONG

The dominant failure is treating an AI system as though it will always be there: firms retire or de-skill the manual path it replaced and never build a fallback. They also leave systems that act without a fast, tested way for a human to halt them, which converts a model error into an uncontrolled event.



Takeaway. Resilience is a design choice made in advance: a fallback that runs without the AI, a stop a person can pull quickly and has rehearsed, and a continuity plan that treats the firm's AI and shared providers as the dependencies they are.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST10

RECOMMENDED ACTIONS

1. **Map the AI dependencies** behind each important business service, including third-party AI and cloud providers, starting where an AI acts or sits on the critical path.
2. **Design to fail safe** with a credible fallback that does not depend on the AI, trained, exercised and sized to the service's impact tolerance.
3. **Implement a tested kill-switch** for every AI system that can act, held by a named role and deputy and rehearsed through intervention drills.
4. **Cover AI failure modes and provider failure** in business continuity and disaster recovery, including a model that is available but wrong.

SEQUENCE AND DEPENDENCIES

1. Sequence the work from dependency mapping outward, because the fallback, the override and the continuity plan all depend on knowing where the AI actually sits.
2. The recurring obstacle is that the manual path an AI replaced is allowed to atrophy, so budget to keep a fallback trained and exercised rather than assuming staff can revert.
3. Treat a manual fallback as a staffing plan: size it to what the service must deliver within its impact tolerance or recovery target.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Build the fallback, the fail-safe defaults and the kill-switch into each AI-dependent service, keep the manual path trained, name the stop holder and deputy.
Second Line	Require dependency mapping and tested fallbacks for critical AI, require the inventory entry and named owners before first use, set the operational-resilience standard.
Third Line	Assure that fallbacks and kill-switches exist and are actually tested, that dependency maps are current, and that continuity plans are exercised rather than filed.
Board and Executive	Own the accountability that cannot be outsourced, personal where a regime such as SM&CR applies: require resilience for AI-dependent critical services.
Cross-Functional	Build ST10's design (fallback, override) and P05-ST12's response (incident playbook) together, and agree who owns each AI dependency and its contingency.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST10

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **No fallback for a critical AI**	Maintain a trained, exercised fallback that does not depend on the AI (scaled-down rules process or manual procedure), sized for the minimum safe service capacity.
HIGH **No fast human stop**	Embedded intervention capacity, an emergency stop or pause with responsibility assigned, with activation protocols that are clear and tested regularly, held by a named role and deputy.
HIGH **Third-party and cloud concentration**	Map and diversify concentrated third-party dependencies; tested exit strategies and tested recovery with alternative arrangements explored where possible (both technology-neutral expectations).
HIGH **Provider-initiated failure**	Contract for version pinning, notice before a version is retired and the right to disable the AI feature (P07-ST04, RA 4).
MEDIUM **Continuity plan ignores AI dependencies**	Map the dependencies behind each important business service, including any AI dependency, and assess their controls.

METRICS SCORECARD

AREA	LEADING	LAGGING
Fallback readiness	Share of critical AI-dependent services with a trained, exercised fallback.	Missed fund dealing, NAV, reconciliation or client deadlines while running in degraded mode.
Human override	Share of acting AI systems with a tested kill-switch and a recent intervention drill	Events where a misbehaving AI could not be stopped quickly; actions left unreconciled after a stop, drills included
Recovery performance	Time from a simulated AI failure to the fallback operating within the service's impact tolerance or recovery target.	Live events where the fallback or the stop took longer than in exercises
Dependency coverage	Share of important business services with AI/third-party dependencies mapped and assessed	Outages traced to an unmapped AI or provider dependency
Concentration	Number of critical services reliant on a single shared AI/cloud provider without a contingency	Disruption caused by a shared-provider outage with no contingency
Near misses	Near misses found in drills or live use (a fallback or stop that worked only in part.	Live failures of a fallback or stop that an earlier near miss had shown

1 more metric in the full article

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST10

AMAT

P07-ST12

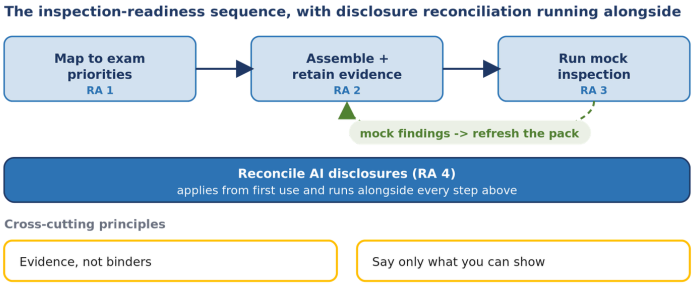
Regulatory Inspection Readiness for AI

For: Risk and Compliance | Legal | Senior Leadership

Supervisors are starting to examine how firms govern AI as well as whether they use it, most explicitly in the US, where the SEC has named AI-related areas in its examination priorities. The question is whether a firm could withstand that examination today. Inspection readiness is the discipline of being able to show, on request, that AI is governed and controlled: a maintained evidence pack keyed to the firm's AI inventory, accurate disclosures about how AI is actually used and a rehearsed ability to answer questions before they are asked. Examinations may be moving toward more data-driven requests, so the evidence has to be retrievable, not just well formatted.

WHERE FIRMS GO WRONG

Firms treat inspection readiness as a document-retrieval exercise done the week the examiner calls, so a request that should take hours takes weeks. They also let what they say about their AI drift from what it does, with marketing or disclosures claiming a capability or level of oversight the operation cannot substantiate.



Takeaway. Readiness is a standing discipline, not a fire drill: a current evidence pack mapped to what supervisors ask for, mock inspections that find the gaps first and AI disclosures that never claim more than the firm can demonstrate.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST12

RECOMMENDED ACTIONS

1. **Run a gap analysis** of AI policies and procedures against the published examination priorities and applicable documentation obligations.
2. **Maintain a current AI evidence pack**, one file per system keyed to the AI inventory and built on the audit trails.
3. **Run periodic mock inspections** with interview rehearsals, treating every gap or slow response as a finding to remediate.
4. **Reconcile every external AI claim** against what the systems actually do, so the firm never describes a capability it cannot demonstrate.

SEQUENCE AND DEPENDENCIES

1. Sequence the work from the gap analysis, because it tells the firm where it is not ready and turns readiness into a prioritized list rather than an open-ended worry.
2. The recurring obstacle is documentation debt: the design rationale and testing evidence a supervisor expects were not captured at the time and cannot be reconstructed convincingly later.
3. Co-locate readiness with the obligations work in P05-ST05 so the two stay aligned: P05-ST05 defines what the rules require, ST12 keeps the firm able to demonstrate it.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Capture the evidence as work happens (design rationale, testing, oversight records) and keep the records for the systems it owns.
Second Line	Own the readiness standard and the gap analysis against examination priorities, challenge the quality and completeness of the evidence.
Third Line	Assure that the readiness is real: test that the evidence pack is current and retrievable across both lines and that mock-inspection findings are remediated.
Board and Executive	Treat an AI examination as a matter of when, not if; require a standing readiness capability and accurate disclosures.
Cross-Functional	Keep ST12 readiness aligned with the P05-ST05 obligations, the ST06 audit trails and the P05-ST10 AI inventory.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST12

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Not examination-ready**	Gap-analyze against the published examination priorities and run periodic mock inspections; remediate findings before a real exam.
HIGH **Cannot produce the evidence a supervisor expects**	Maintain a current, retrievable evidence pack built on the ST06 audit trails, mapped to authorization and documentation obligations and kept to the retention standard.
HIGH **Misrepresenting AI use**	Reconcile every external AI claim, including DDQ and RFP answers, against actual system behavior; say only what can be demonstrated.
MEDIUM **Underestimating supervisory reach**	Prepare for data-driven, AI-augmented examination: evidence that withstands testing against data, including extracts from the inventory, system logs and oversight records, not just a policy on file.
MEDIUM **Unassessed incidents**	Hold the incident log, the classification decision for each incident and a record of incidents considered and not reported, with the reason, in the evidence pack (P05-ST12).

METRICS SCORECARD

AREA	LEADING	LAGGING
Readiness coverage	Share of published examination priorities with a mapped, current evidence response	Examination requests that could not be met from the existing pack
Evidence retrieval	Time to assemble a supervisor's evidence request from the pack.	Requests where assembly took weeks rather than hours
Mock inspection	Mock-inspection cadence and share of findings remediated before the next cycle.	Real-examination findings that a mock inspection should have caught; material mock-inspection findings that recur at the next cycle
Disclosure accuracy	Share of external AI claims reconciled against actual system behavior	Misrepresentation or disclosure findings raised at examination
Board sight	Share of higher-risk AI systems on which the board or a board committee received a report.	Examination findings that the board was not told of a material AI risk or incident
Human review integrity	Rate at which reviewers edit, override or reject AI output, trended	Material AI errors reaching an investment, client or regulatory decision despite review



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P07-ST12

AMAT

Pillar 8 Executive Summary

For: Senior Leadership; Board and Executive; Finance · Also relevant to: Strategy, Technology, Risk and Compliance, Cross-Functional

THE ARGUMENT

Pillar 8 is the execution discipline that turns strategic ambition into sustained operating capability: the roadmap, the use-case portfolio, ROI and impact metrics, the funding model, scaling from pilots to production, the operating cadence, maturity benchmarking, industry collaboration, and the agentic future. The through-line across the pillar is that AI value is realized only when a firm treats scaling and value as disciplines rather than outcomes: it reallocates spend rather than adding to it, agrees how to measure value before it scales, benchmarks maturity honestly, and earns autonomy as governance keeps pace. In this guide's view, many firms are still in the gap between pilots and production, which is why the advantage remains open to those that move deliberately now.

81% / 14%
of industry respondents in CCAF's 2026 financial-services survey were adopting AI; 14% viewed their deployment as transformational

60%
of companies globally generated no material value from AI despite investment (BCG, 2025 survey)

7%
of asset managers have deployed agentic AI to date (EY, base not stated); 78% of 100 firms were exploring it (EY, early 2025)

WHY NOW

Adoption is now widespread, with 81 percent of industry respondents in CCAF's 2026 survey of financial-services firms adopting AI, so simply using AI no longer separates a firm from its peers; in this guide's view, advantage has moved to execution and realized value. Measurement is the weak point: in the same survey, 55 percent of industry respondents find measuring AI value challenging. Meanwhile the rules for AI are being written now and agentic interest is running ahead of deployment.

RISKS TO AVOID

The pilot trap: mistaking a portfolio of demonstrations for transformation and staying stuck in the gap between pilots and production, so the execution gap never closes. **Unmeasured, unfunded value**: scaling without the measurement to prove returns, or self-scoring maturity from opinion, so real value is cut because it cannot be shown. **Autonomy outrunning control**: funding without discipline and scaling agents faster than governance, reliability and a tested failure plan can keep pace.

Bottom line. Pillar 8 is the difference between running pilots and running an AI-native firm. In this guide's view, these are disciplines rather than outcomes: reallocate rather than add, measure before scaling, benchmark maturity honestly, and earn autonomy only as governance keeps pace. The evidence is uneven: no cited asset-management survey establishes a realized AI ROI percentage, no asset manager in the evidence base publishes an AI-attributed alpha figure or a meeting cadence, collaboration evidence does not show that participation changes final rules, and size-band results do not map to AUM tiers.

Explore Pillar 8 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P08-summary

PRIORITY MOVES FOR LEADERSHIP

1. **Fund transformation by reallocation, not addition**: shift the technology budget toward transformation rather than adding to a starved base, commit investment at levels associated with higher maturity, cost the full ownership, and tie continued funding to measured outcomes.

2. **Build foundations and pilot under control, then scale as a discipline**: run pilots under first-use controls, put clean data and redesigned workflows in place before scaling, agree one value method and a shared platform, govern with fast stop-or-scale calls, and control cost and value before scaling agents.

3. **Manage AI as a governed portfolio, not a collection**: see which use cases deliver value, rank them on value against complexity, concentrate on the one or two domains that compound, and retire the ones that do not with equal willingness.

4. **Measure value honestly before you scale**: agree baselines and a measurement method first, build wider infrastructure as you scale, track all four value drivers rather than the countable one, attribute alpha with care, and make every claim assurance-ready.

5. **Benchmark maturity to find the real gap**: assess against one credible model on evidence, not opinion, across firm-wide, data-readiness and assurance lenses, benchmark assurance externally, and turn the baseline into the staged roadmap.

6. **Run a standing cadence and earn autonomy deliberately**: write down decision rights, set the firm's own cadence and a board agenda beat, engage the forums writing the rules, and move up the agentic autonomy bands only as governance and reliability keep pace.

IN THIS PILLAR

P08-ST01	AI Transformation Roadmap	240
P08-ST02	Use Case Prioritization and Portfolio Management: Choosing and Governing the AI Bets That Pay	243
P08-ST03	Proving the Payoff: Benefits Tracking and ROI Measurement for AI	246
P08-ST04	Productivity, Alpha, Risk and Client Outcome Metrics: Measuring What AI Actually Delivers	249
P08-ST05	Funding the AI Transformation: The Model and the Investment Case	252
P08-ST06	Scaling from Pilots to Production: Turning Proven AI into Production Capability	255
P08-ST07	Enterprise AI Operating Cadence and Governance Forums: Building the Rhythm That Keeps AI Moving	258
P08-ST08	Industry Collaboration and Common Standards: Shaping the Rules Rather Than Reacting to Them	261
P08-ST10	Knowing Where You Stand: AI Maturity Models and Benchmarking	264
P08-ST12	The Agentic Turn: Autonomous Workflows and the AI-Native Firm	267

Explore Pillar 8 online
The full pillar summary, every sub-topic article and the evidence behind it.
amatguide.org/amat-guide#P08-summary

P08-ST01

AI Transformation Roadmap

For: Senior Leadership | Technology

An AI transformation roadmap is not a technology project plan with an end date. It is a continuously governed redesign of the operating model, sequenced so that foundations come before scale. In a 2026 survey of 352 financial-services respondents, 81 percent were adopting AI at some level but only 14 percent regarded their deployment as transformational, a 67 percentage point execution gap. The evidence points to four disciplines: a business-led strategy that leads rather than follows tooling decisions, foundations built before pilots, whole domains redesigned rather than bolted onto legacy processes, and a plan re-cut on a cadence short enough to keep pace with the technology.

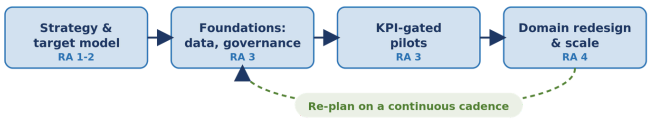
81%
of 352 surveyed financial-services respondents adopt AI at some level (CCAF, 2026)

14%
regard their AI deployment as transformational to strategy and competitive advantage

WHERE FIRMS GO WRONG

The most common internal failure is not a shortage of AI activity but the absence of a plan to direct it; only 50 percent of the AI leaders in one 500-executive survey had an implementation plan with timelines, milestones and metrics. Asset managers also spend on average 60 to 80 percent of their technology budget running the business, so change stays underfunded, and launching hundreds of disconnected use cases dilutes effort further. A roadmap written once and filed is obsolete on contact, because the pace of change outruns the plan.

A sequenced AI transformation roadmap



Cross-cutting principles

- Business-IT co-ownership of every stage
- One senior owner empowered to hold scope

AMAT framework, drawing on Grant Thornton (2026), Alpha FMC (2026), SEI (2026), EY (2025) and McKinsey (Godsall et al. 2025; Soller and Quensel 2025).

Takeaway. Set the strategy before the tools, build the foundations before the pilots, redesign whole domains instead of automating tasks, fund change by reallocating from legacy run costs, and keep re-planning on a cadence that matches the technology.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST01

RECOMMENDED ACTIONS

1. Set a business-led AI strategy with measurable objectives before any technology or vendor decision.
2. Establish the target (and, for multi-country firms, the global) operating model, with decision rights and rules for local exceptions, before starting the platform build; refine the detail as the roadmap re-plans.
3. Start by checking data quality and governance for the first candidate use case before any pilot; build toward the full sequence: data quality and governance first, standardized workflows second, KPI-gated pilots third, scale last.
4. Start by redesigning one high-potential domain end to end; build toward a focused set of whole domains rather than a scattered portfolio of use cases.
5. Write the plan down with explicit phases, milestones, named owners and KPIs.
6. Start by baselining the run and change split of technology spend; build toward a technology budget reallocated from run-the-business toward change-the-business.
7. Start by measuring delivery-cycle length on one program; build toward cycles short enough for the roadmap to iterate at the pace of the technology.
8. Name one senior owner empowered to hold scope. Start by fixing the date of the first formal re-plan; build toward a regular re-planning cadence.

SEQUENCE AND DEPENDENCIES

1. The binding constraint on delivery is ownership.
2. Firms frequently fail to fully decommission legacy systems even after modernization, leaving bloated application portfolios and limited efficiency gains, as McKinsey's 2025 analysis finds.
3. Finally, the roadmap should express a firm-specific AI identity rather than a generic template.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns domain redesign and value streams; adopts vendor-standard processes by default and escalates only documented, business-justified exceptions; with Finance or the program office.
Second Line	Stands up governance and responsible-AI structures early rather than as a retrofit, and gives independent challenge on how AI use cases are risk-classified and controlled (Pillars.
Third Line	Independently audits the governance and control framework around the roadmap, including whether the benefit reporting produced by the first line and Finance can be relied.
Board and Executive	Owns the vision and the business-led strategy and appoints senior owners who can assertively challenge business-side scope preferences that would undermine the business case.
Cross-Functional	Maintains full business-IT collaboration with clear value-stream roles; IT-led programs without genuine business commitment put the business case at risk.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST01

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Execution gap**	Move prioritized domains through KPI-gated stages from pilot to scale; do not let initiatives stall at pilot.
HIGH **No strategy or plan**	Set a business-led strategy and commit it to a written plan with milestones, owners and KPIs.
HIGH **Run-the-business budget trap**	Run a deliberate budget reallocation program shifting spend from run to change.
MEDIUM **Bolt-on trap**	Redesign prioritized domains end to end; treat tools as inputs, not the strategy.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Strategy and plan quality	Share of AI initiatives mapped to a written plan with milestones and KPIs	ROI realized versus the approved business case
Budget mix (run vs change)	Quarterly run/change split of technology spend	Share of spend on change-the-business
Delivery cadence	Average delivery-cycle length	Time from approval to first measured benefit, with change failure and incident rates
Domain focus	Number of domains under end-to-end redesign	Value captured per domain

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

The differentiation window is open but narrowing, so a small firm should move deliberately and anchor its decisions in a firm-specific AI identity.

MID-SIZED

Capacity squeeze

One mid-sized pension investor recognized its in-house capability was insufficient and, rather than building everything itself, partnered externally to build a business case, roadmap, responsible-AI framework and early prototypes.

LARGE AND GLOBAL

In McKinsey's 2025 account, one firm above one trillion dollars in AUM rebalanced its technology budget from roughly 80 percent run-the-business toward 70 percent change-the-business over five years and compressed.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST01

P08-ST02

Use Case Prioritization and Portfolio Management: Choosing and Governing the AI Bets That Pay

For: Senior Leadership | Investment Teams

Every function in a firm can now name AI use cases it would like to run. The scarce discipline is choosing among them and managing the set as a portfolio rather than a scatter of pilots. In KPMG's 2026 survey of Dutch asset managers, only about a quarter could see which use cases deliver value, and departmental initiatives often bypass change governance. In McKinsey's 2024 account of financial services, the institutions that appear to be gaining momentum narrow to a small number of high-value domains and scale those deliberately. The test is not how many use cases a firm is running but whether it can say which ones are worth running.

a quarter

of surveyed Dutch asset managers had a clear view of which use cases deliver measurable value (KPMG, January 2026)

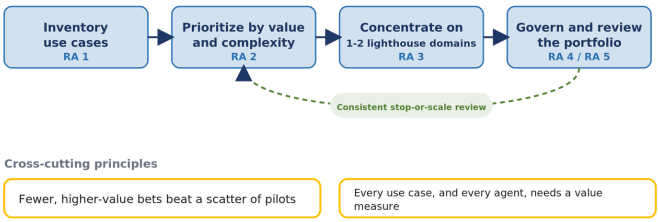
42%

reviewed their use cases consistently

WHERE FIRMS GO WRONG

The dominant failure is a portfolio that no one actually manages. Use cases are generated bottom-up by individual departments, bypass change processes and leave leadership with little visibility into their impact. Without a shared way to see value and a cadence to act on it, weak use cases are neither scaled nor killed; they persist, and launching many pilots at once works against impact.

The use-case portfolio cycle



AMAT framework, drawing on KPMG Netherlands (2026), Assogestioni and Bain (2025), McKinsey (2024) and EY (2026).

Takeaway. Managing AI as a portfolio rather than accumulating it as a collection is what turns a wide opportunity set into realized return: rank honestly on value against complexity, concentrate on the one or two domains that compound, and scale and retire with equal willingness.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P08-ST02

RECOMMENDED ACTIONS

1. Manage AI use cases as one governed portfolio with a consistent intake, so initiatives are captured and assessed rather than bypassing change governance. The intake covers AI capabilities embedded in existing research, portfolio, order-management, client and service-provider platforms, not only internal pilots.
2. Start by scoring the use cases already running on value at stake versus implementation complexity; build toward funding the quick wins and the deliberate strategic bets from one shared ranking that passes risk and control gates, and parking the rest.
3. Start by naming the one lighthouse domain with the strongest case; build toward one or two lighthouse domains, clusters of related use cases, rather than spreading across dozens of pilots.
4. Start by agreeing value-linked evaluation criteria for the use cases already running; build toward a reusable delivery playbook and a consistent review cadence that scales or stops use cases on the evidence, with a named authority to suspend or retire each one.
5. Require a per-dollar value measure for every agent at approval, scale and renewal; do not let agent count stand in for return.

SEQUENCE AND DEPENDENCIES

1. The first practical move is to sequence adoption so the portfolio earns the right to grow.
2. The second is to make prioritization a shared, cross-functional act.
3. The third is to accept that the portfolio is a living object.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the intake and runs the prioritization; proposes use cases against value criteria and captures the reusable delivery playbook.
Second Line	Independently challenges the proposed risk tier, material-change classification and adequacy of controls, and sets escalation requirements (P05-ST08).
Third Line	Tests that portfolio decisions rest on defined, value-linked criteria and a genuine stop-or-scale cadence, not advocacy.
Board and Executive	Concentrates resources on fewer, higher-value domains, resists the scatter, and names an accountable portfolio owner.
Cross-Functional	Defines and prioritizes use cases across lines of business with shared KPIs and reusable, risk-tiered playbooks.



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST02

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Ungoverned bottom-up sprawl**	Route initiatives through a governed intake and a single managed portfolio with shared value criteria.
HIGH **Scatter across too many pilots**	Concentrate on one or two lighthouse domains; cap active bets and force trade-offs.
HIGH **Uncontrolled embedded or retired AI**	Record embedded AI in the inventory before use (P05-ST10), obtain vendor change notices through P07-ST04, identify the business owner and disable route.
MEDIUM **Prioritizing by availability, not value**	Rank on value at stake vs complexity with explicit, value-linked criteria before funding.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Portfolio value visibility	Use cases with a defined value measure and approved baseline at intake	Share of portfolio spend in use cases whose realized value meets or beats their approved baseline
Portfolio governance cadence	Review cadence maintained; share of use cases reviewed each cycle	Use cases scaled, stopped or retired on review evidence, and spend released from stopped use cases
Concentration	Share of resource on the top 1-2 domains	Share of pilots in the lighthouse domains converted to production value
Agent value	Value measure and total cost estimate defined per agent at approval	Realized value per dollar per agent against its approval baseline

BY FIRM SIZE

BOUTIQUE

Exposed

A small firm cannot fund a broad portfolio: pick one or two high-value use cases and prove value before adding more.

MID-SIZED

Capacity squeeze

The mid tier has more demand for use cases than it can fund and govern at once, so a selective set is the sensible norm.

LARGE AND GLOBAL

Advantaged

The largest firms have the resources to run a broader portfolio, so the discipline at this scale is value visibility rather than volume.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P08-ST02

AMAT

P08-ST03

Proving the Payoff: Benefits Tracking and ROI Measurement for AI

For: Senior Leadership | Finance

AI adoption has run ahead of impact, and proving what AI returns is harder still. In CCAF's 2026 survey, 55 percent of financial-services respondents found measuring AI value challenging, even at the scaling stage, and value that cannot be evidenced is hard to defend. The evidence favors building measurement capability early, tracking value across several economic dimensions rather than one ROI number, and using different yardsticks and horizons for generative and agentic AI, with a payback horizon set in each business case.

55%
of financial-services respondents find measuring AI value challenging (CCAF, 2026)

60%
of companies globally were not generating any material value from AI despite substantial investment (BCG, 2025)

WHERE FIRMS GO WRONG

The first failure is measuring activity instead of outcomes: documents processed are easy to report, cost avoided or revenue added is harder. The second is the wrong yardstick and clock, since most organizations in one 2025 survey reached satisfactory AI ROI in two to four years, and generative and agentic AI return value on different timescales. The third is attribution, because AI arrives alongside process redesign, data work and organizational change.

Two ways to read AI's return

Dimension	Efficiency-only ROI	Multi-bucket value view
Cost	Labor time saved	Cost avoidance, quantified
Value	Efficiency only	Plus shareholder value and growth
Risk	Not counted	Plus risk reduction, quantified
Horizon	7-12 months (AWS, via Deloitte)	Firm-set horizon, tracked (Deloitte 2025: 2-4 years)

Source: AMAT framework, drawing on MIT Sloan / Vanguard (2025), Deloitte (2025) and AWS (2024, cited by Deloitte).

Takeaway. Measure before you scale: build measurement infrastructure, track value in defined buckets anchored in task-level baselines, judge generative and agentic AI on their own horizons and make every claim assurance-ready.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST03

RECOMMENDED ACTIONS

1. Establish before-and-after baselines for AI-touched workflows and agree the measurement method early, and build the wider measurement capability as the program scales; in CCAF's 2026 survey, measurement difficulty stayed high at the scaling stage and was not reduced by higher maturity.
2. Start by tracking the largest one or two use cases in defined economic buckets; build toward tracking value across the portfolio in those buckets (cost avoidance, growth/shareholder value, risk reduction, operational efficiency).
3. Start by baselining one workflow at the task level; build toward before-and-after baselines (headcount, volume, unit cost, accuracy) across AI-touched workflows so cost avoidance is computed, not estimated.
4. Use different frameworks and time horizons for generative versus agentic AI, and set a firm-specific payback horizon in each business case; in Deloitte's 2025 survey, most organizations reached satisfactory ROI within two to four years.
5. Count intangible, growth and risk value alongside efficiency, and define wins as net value created across cost, growth, risk and service quality, not as revenue growth or cost savings alone.
6. Start by keeping documentation for the value claims that go to the board; build toward assurance-ready evidence (audit trails, documentation) for every material claim, the scope this guide sets, so claims about AI withstand scrutiny.

SEQUENCE AND DEPENDENCIES

1. The first practical point is sequencing: baseline before you build.
2. The second is that measurement must include the full cost of getting to value.
3. Value tracking is continuous, not a one-off business case, and it needs named owners close to the workflows being measured.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns task-level baselining and use-case value tracking; captures the before, records accuracy and volume, and reports value in buckets rather than activity.
Second Line	Challenges attribution and metric integrity; ensures risk-reduction value is quantified and that value claims are not overstated.
Third Line	Independently assesses, within its risk-based plan, the design and operating effectiveness of the controls that make value claims evidenced, traceable to a P&L effect and reconstructable.
Board and Executive	Sets a payback horizon in each business case, defines wins as net value across cost, growth, risk and service quality rather than efficiency alone.
Cross-Functional	Provides the enterprise value-tracking infrastructure and the common bucket definitions so value is measured consistently across the firm.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST03

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Measurement blind spot**	Put before-and-after baselines and an agreed measurement method in place before scaling, and build the wider measurement infrastructure as the program grows.
HIGH **Narrow-metric undervaluation**	Track value across defined buckets and define wins as net value across cost, growth, risk and service quality, not efficiency alone.
HIGH **Inflated or double-counted benefit**	Finance reconciles unique benefit entries to the baseline and realized P&L, keeps capacity separate from cash savings, and checks full implementation and run costs.
MEDIUM **Horizon and expectation mismatch**	Set a firm-specific payback horizon in each business case and use different frameworks for generative versus agentic AI.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Value realization	Share of sampled AI-assisted tasks meeting the agreed quality and unit-cost baseline	Realized annual net benefit against the approved no-AI baseline
Benefit validation	Share of business-owned benefit claims with a validated before-and-after sample	Share of forecast cash benefit realized in the period
Payback horizon	Time-to-value milestones met	Months to satisfactory ROI
Value composition	Share of tracked value from growth and risk, not efficiency alone	earnings before interest and taxes (EBIT) impact attributed to AI

1 more metric in the full article

BY FIRM SIZE

BOUTIQUE

A simple operating estate makes AI's contribution easier to isolate: baseline one use case at the task level before deploying and read value directly.

MID-SIZED

Capacity squeeze

The constraint is capacity: moving from activity metrics to a genuine value-tracking infrastructure with limited people and tooling.

LARGE AND GLOBAL

In CCAF's 2026 survey, 76 percent of large firms reported difficulty measuring AI value (CCAF's own size band, not an AUM band).

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST03

P08-ST04

Productivity, Alpha, Risk and Client Outcome Metrics: Measuring What AI Actually Delivers

For: Senior Leadership; Investment Teams; Finance

AI could materially change asset management economics, yet many firms still measure it too narrowly to prove value or steer investment. This article sets out how to measure AI across its four value drivers, productivity, alpha, risk reduction and client outcomes, pairing each with leading and lagging indicators and, where the evidence allows, a benchmark. The central discipline is to instrument the AI-specific and forward-looking metrics that standard financial reporting misses, and to baseline each process before deployment so impact is measured rather than estimated. The evidence base for productivity is strong; for alpha, risk reduction and client outcomes it is thinner, and the firm's own instrumentation must fill the gap.

25 to 30%

estimated productivity uplift for European asset managers across the value chain (McKinsey estimate, cited by EFAMA 2025)

80%

efficiency lift estimated for agentic systems on high-impact use cases, against about 20% for prior-generation task tools

WHERE FIRMS GO WRONG

The most common failure is narrowness rather than the absence of metrics. In ThoughtLab's 2025 survey of 500 wealth and asset management firms, AI leaders measured value across four categories while other firms tended to measure only standard financial metrics. Even firms that measure broadly tend to instrument productivity well and the other three drivers poorly, because alpha, risk reduction and client outcomes demand attribution judgments that most frameworks avoid.

Four-category AI metrics framework

ThoughtLab 2025: leaders gave more attention to all four categories than other firms.



Source: adapted from ThoughtLab, 2025.

Takeaway. Measure across all four drivers, baseline before deploying and attribute alpha rigorously enough to survive challenge. Productivity can be measured with confidence now; for the other three drivers the firm must build the instrumentation the external evidence does not yet supply.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P08-ST04

RECOMMENDED ACTIONS

1. Start by adding the AI-specific category to the existing scorecard; build toward a four-category metric set, strategic, people, innovation and AI-specific, and retire scorecards that track financial metrics alone.
2. Baseline every target process before deployment (headcount, task volume, unit cost, cycle time, error rate, total effort) on definitions agreed with Finance, so impact is measured, not estimated.
3. Split every driver into leading and lagging indicators; never manage on lagging outcomes alone.
4. Start by reporting a sampled output-quality and error measure, the share of eligible work handled with AI and the frequency of AI-assisted tasks for the largest deployment; build toward these AI-specific metrics as headline KPIs for every deployment.
5. For alpha, start by recording AI-assisted decisions beside a comparable non-AI process, using a pre/post comparison only to test feasibility; build toward isolating AI-attributable excess return from market beta and sentiment against a control before claiming an uplift.
6. Track the rate at which reviewers edit, override or reject AI output, trended, and investigate whenever it falls to or near zero for a sustained period, a possible sign of rubber-stamping.

SEQUENCE AND DEPENDENCIES

1. The hardest part of measurement is not choosing indicators but capturing an honest baseline.
2. Metrics such as model accuracy, adoption and cycle time live in different systems from the finance ledger.
3. Once a tool is live, the pre-AI state is gone, so baselining has to happen before deployment and be owned by someone other than the team promoting the tool.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns operational productivity and adoption metrics: cycle time, STP rate, model accuracy and usage for the workflows it runs.
Second Line	Sets risk-measurement expectations, reviews exceptions and independently challenges the integrity of reported risk, error and AI-alpha outcomes.
Third Line	Audits the integrity of the metric set itself: whether baselines are real, attribution is defensible, and reported gains reconcile to source systems.
Board and Executive	Owns the enterprise value-realization scorecard across the four drivers and holds the mandate to fund or stop on the evidence.
Cross-Functional	Finance and Technology jointly maintain the baselines, data feeds and metric definitions so the numbers are comparable across functions.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST04

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Narrow-metric blind spot**	Adopt the four-category metric set and make AI-specific metrics headline KPIs.
HIGH **Unproven pilots stall**	Baseline each process before deployment and define the success metric before the pilot starts.
MEDIUM **AI-exposure return mistaken for alpha**	Attribute against a control or counterfactual; separate AI-attributable return from market beta.
MEDIUM **Lagging-metric over-reliance**	Pair every lagging indicator with a leading one; review leading indicators on a faster cadence.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Productivity	Task cycle time; tasks per person; STP rate; adoption rate	Unit cost per task; cost-base reduction
Investment performance (alpha)	Validated research coverage per analyst; ideas progressing to a decision; signal quality after human review	Risk-adjusted excess return attributable to AI (vs control)
Risk reduction	Anomaly-detection coverage; exceptions auto-cleared; control-test pass rate	Incident rate; error rate; control-failure count
Client outcomes	Onboarding time; personalization coverage; response latency	Client satisfaction; retention

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Track a minimal metric set for one or two material vendor-enabled workflows: appoint a first-line owner and record completed-task time.

MID-SIZED

Capacity squeeze

The constraint is capacity to instrument all four drivers at once, so the natural starting point is productivity, the driver with the strongest evidence (see Practical Applications).

LARGE AND GLOBAL

Large firms have the scale to instrument AI-specific telemetry, model accuracy, adoption and usage, across the portfolio of deployments.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST04

P08-ST05

Funding the AI Transformation: The Model and the Investment Case

For: Senior Leadership | Finance

Funding AI transformation is not primarily a spending problem; it is an allocation and measurement problem. In NVIDIA's 2026 survey, 83 percent of financial-services firms expected AI spending to increase, yet among large asset managers the statistical relationship between technology spend and productivity is close to zero. Three disciplines recur in the evidence: reallocating the technology budget from running legacy systems toward genuine transformation, building business cases on the fully loaded cost of ownership rather than headline license fees, and tying funding to measured outcomes. The funding figures managers publish are plan envelopes and combined spend lines, so none isolates the cost of AI. How a firm funds AI matters as much as how much it spends.

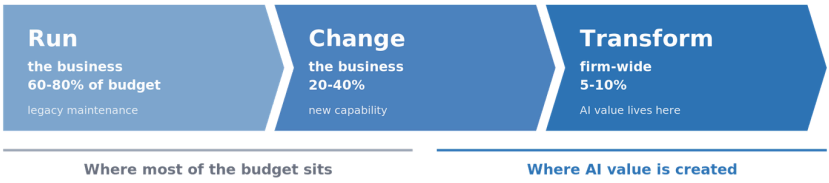
83%
of surveyed financial-services firms expected AI spending to increase in 2026 (NVIDIA)

1.3%
R-squared between technology spend and productivity across the large asset managers in McKinsey's panel

WHERE FIRMS GO WRONG

The internal problem is not underspending but spending that has come apart from value: technology grew at 8.9 percent a year against 4.5 percent for total costs, yet its link to productivity is almost nonexistent. With 60 to 80 percent of technology budget spent running the business and only 5 to 10 percent of total spend reaching firm-wide transformation, an AI program is underfunded by construction. Business cases then under-cost AI by omitting governance, change and failure-and-recovery costs.

Where the technology budget sits vs where AI value is created



Source: McKinsey, 2025.

Takeaway. The board's task is not to approve a number but to insist on the discipline that turns it into value: reallocate rather than add, cost AI fully and tie continued funding to measured outcomes.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST05

RECOMMENDED ACTIONS

1. **Reallocate the technology budget** by baselining the run-to-change split and setting an explicit reallocation target, funded by simplifying the legacy estate.
2. **Set a material funding floor tied to ambition**, starting with a bounded first tranche sized to it and managing the investment adaptively.
3. **Build every business case on fully loaded total cost of ownership**, including governance, change, failure recovery and ongoing run costs.
4. **Fund responsible AI as a value enabler**, with a responsible-AI line in every AI business case.
5. **Choose a funding model that fits the operating model**: fund the shared platform centrally and use cases from the business units that own them.
6. **Tie funding to measured outcomes**, reinvest the gains and stop funding what does not demonstrably work.

SEQUENCE AND DEPENDENCIES

1. The first practical point is that, for many firms, funding AI is a reallocation problem as much as a shortfall.
2. The second is to start the business case with a modest, bounded diagnostic.
3. Fund in tranches tied to evidence rather than as a single annual commitment, keep a contingency for the ongoing run costs that capital-expenditure-only models miss.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the use-case business cases and their fully loaded costs; builds each case on complete total cost of ownership and runs it against measured value rather.
Second Line	Prices AI's own risk into the investment case, challenges the assumptions behind the committed risk reserve, and challenges cost control as programs scale.
Third Line	Tests that claimed returns are measured rather than asserted and that implementation and ongoing run costs were not omitted from the case at approval.
Board and Executive	Owns the funding-commitment and run-to-change reallocation decisions and the responsible-AI investment level and, with Finance, the sizing and release of the committed risk reserve.
Cross-Functional	Funds and stewards the shared data and model platform centrally so it is not orphaned between business-unit budgets.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST05

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Spending without reallocating**	Baseline the run-to-change split and approve an explicit reallocation target alongside the initial AI investment decision; do not hold a bounded, controlled first tranche for legacy decommissioning.
HIGH **Run-the-business trap**	Fund legacy decommissioning and simplification as an explicit, ring-fenced line, not a residual.
HIGH **Under-costed business case**	Mandate a fully loaded TCO covering every cost component in each case, compared with expected value.
HIGH **Unpriced AI risk**	Use a risk-adjusted ROI that prices both risk reduction and risk increase, and hold a committed risk reserve.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Budget allocation	Run-to-change ratio trend	Share of tech spend on firm-wide transformation
Investment intensity	AI share of technology or digital budget	Advanced-maturity and profitability outcomes
Responsible-AI funding	Responsible-AI spend, planned vs actual	Responsible-AI maturity score; earnings impact
Business-case integrity	Fully loaded TCO captured in every case	Variance of realized vs projected cost

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Exposed

Absolute spending levels are harder to reach here: the roughly 100,000-dollar annual spending level that CCAF's 2026 survey associates with advanced maturity is proportionally harder for a boutique to reach.

MID-SIZED

Capacity squeeze

The constraint is funding capacity: paying for the shared capability no single desk will fund.

LARGE AND GLOBAL

Advantaged

The opportunity is reallocation at scale, as the trillion-dollar case's shift from about 80 percent run-the-business to about 70 percent change-the-business shows.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P08-ST05

AMAT

P08-ST06

Scaling from Pilots to Production: Turning Proven AI into Production Capability

For: Senior Leadership | Technology

AI adoption is now widespread; far fewer asset managers run AI in production. In ESMA's 2026 survey, investment managers reported among the lowest rates of a fully rolled-out use case of any sector surveyed, at roughly one in five. The gap is an operating-model problem: pilots stall, banking evidence shows, without a shared platform, an agreed value method and cross-functional ownership, and when a broken workflow is automated instead of redesigned. Scaling is a distinct discipline, sequenced after clean data and standardized workflows and governed by fast stop-or-scale decisions.

53%

of EU investment firms and investment managers reported no AI use cases, a gap ESMA links largely to firm size (ESMA, 2026)

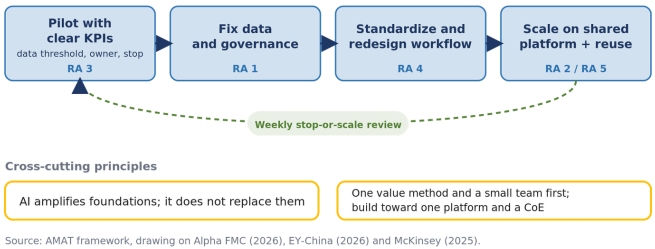
11%

of 500 US technology leaders had agentic systems in production, while 38% were piloting (Deloitte 2025, cross-industry)

WHERE FIRMS GO WRONG

The central problem is treating scaling as the automatic sequel to a good demo. In a 2025 EY and IIF survey of banks, more than half of AI projects stalled in the pilot phase, and the cause is rarely model quality but the absence of a unified data and model platform, an agreed value method and a cross-functional center of excellence. Firms also pile up isolated proofs of concept that never cohere into enterprise value and layer AI onto broken workflows.

The scaling sequence: from controlled pilot to production



Source: AMAT framework, drawing on Alpha FMC (2026), EY-China (2026) and McKinsey (2025).

Takeaway. Build the capability to scale rather than accumulating pilots: sequence clean data and redesigned workflows first, stand up shared foundations, and put cost and value control in place before scaling autonomous agents.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P08-ST06

RECOMMENDED ACTIONS

1. Sequence the investment: pilot only on data that meets a defined quality threshold and with clear KPIs; put data quality and governance in place and standardize the target workflow before scaling AI.
2. Stand up the three scaling foundations: start by agreeing one use-case value method and naming a small cross-functional team; build toward a shared data and model platform and an AI center of excellence before broad rollout.
3. Run a pilot, verify, promote cadence with a fast, evidence-based stop-or-scale decision and short business-feedback loops, at least weekly.
4. Redesign the workflow around AI rather than automating the existing one: start by mapping one end-to-end process before deploying AI in it; build toward redesigning end-to-end processes, not isolated tasks.
5. Build for reuse: start by documenting each proven use case as a reusable pattern; build toward a reuse library on the shared platform that cuts the marginal cost of each new one.
6. Govern agents from first use: give every agent a named owner, limits on its actions and a tested stop mechanism before it acts, and put cost, usage, risk and value governance in place before it scales.

SEQUENCE AND DEPENDENCIES

1. Funding is often the first practical obstacle, and McKinsey's asset-management research suggests it reflects how budgets are spent as much as how large they are.
2. Rewire the workflows, put a compatible architecture in place and stand up trustworthy governance before productionizing.
3. Do not wait for perfect conditions: begin with practical, proven use cases and embrace experimentation, convene cross-functional ownership, and have senior leaders visibly use the tools themselves.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns workflow redesign and adoption: runs the pilot, verify, promote cadence and captures proven use cases as reusable patterns rather than one-off builds.
Second Line	Establishes cost, usage, risk and value governance for scaled and agentic AI before scale, and guards against ungoverned, fragmented deployment across business units.
Third Line	Tests that claimed AI benefits rest on controlled evaluation rather than self-selected usage, and that scaled use cases remain documented and reconstructable.
Board and Executive	Holds the sequencing and funding discipline: treat scaling as a distinct capability, fund redesigned processes not tools, and hold the stop-or-scale line.
Cross-Functional	Provides the shared platform, the ROI method and the business-technology integration on which the entire scaling model depends.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST06

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Automating broken workflows**	Mandate zero-based, end-to-end workflow redesign before deployment; fund scaling only for redesigned processes.
HIGH **Premature agentic scaling**	Require per-agent ownership and a kill-switch from first use; a cost ceiling, a defined value measure, cost thresholds and drills before scaling; defer control design to Pillar 5.
HIGH **Deploying on weak data foundations**	Gate each pilot on a data-quality threshold for the data it uses, and gate scaling on firm-level data quality and governance.
HIGH **Proliferation without integration**	Manage a single portfolio of use cases with one value method from the start and retire duplicative pilots; build toward a shared platform and reuse library.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Pilot-to-production conversion	Use cases passing the stop-or-scale gate per quarter	Share of pilots reaching production
Reuse and unit economics	Reuse rate of shared components	Marginal cost per new use case
Time-to-value	Pilot-observed change in cycle time, errors and manual handling against the agreed baseline	Sustained net operational benefit in production, after vendor and oversight costs
Adoption depth	Share of target workflow volume passing validated quality thresholds during controlled rollout	Share of eligible workflow volume processed in production with AI within agreed quality and incident thresholds

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Exposed

In ESMA's survey, about 65 percent of micro firms had no active or planned AI use cases; in this guide's view.

MID-SIZED

Capacity squeeze

The bottleneck is the shared platform, value method and center of excellence that EY's banking evidence identifies, not more pilots.

LARGE AND GLOBAL

Advantaged

In ESMA's survey, about 96 percent of large firms were using or planning AI; industrialize with reusable recipes and centralized delivery.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P08-ST06

AMAT

P08-ST07

Enterprise AI Operating Cadence and Governance Forums: Building the Rhythm That Keeps AI Moving

For: Senior Leadership; Risk and Compliance

An AI transformation is a standing capability, not a project, and it needs a standing operating rhythm to match: a defined operating model, governance forums with clear ownership and a regular decision cadence. In McKinsey's 2024 study of banks and other financial institutions, roughly 70% of those running highly centralized gen AI operating models had moved use cases into production, against about 30% of those with fully decentralized models. A central team that commands rather than coordinates, however, risks becoming a bottleneck. The operating-model evidence comes mainly from financial institutions, four asset managers describe governance structures they have set up or announced, and none publishes a meeting frequency.

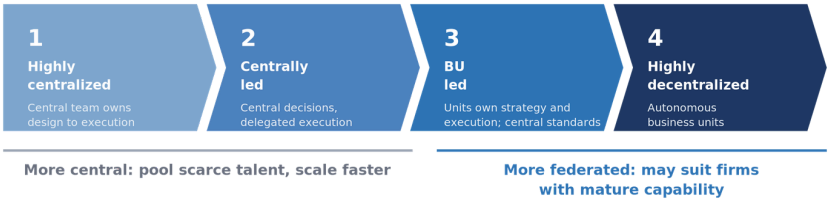
70%
of financial institutions with highly centralized gen AI operating models had moved use cases into production (McKinsey, 2024, banks and other financial institutions)

30%
of those with fully decentralized models

WHERE FIRMS GO WRONG

The common failure is to treat AI governance as a gate rather than a rhythm: a forum that convenes quarterly to bless projects, or a policy written once and shelved, is governance theater. The opposite over-correction is a central AI team that owns delivery end to end and becomes a bottleneck the business routes around. Both share a root cause, which is that ownership and decision rights are left undefined, so accountability diffuses and the rhythm never forms.

Four gen AI operating-model archetypes (centralization spectrum)



AMAT framework, drawing on McKinsey, 2024 (banks and other financial institutions).

Takeaway. An AI capability without a rhythm decays into silos and box-ticking. Choose how much to centralize, write down the decision rights, run a small coordinating team on a fixed drumbeat under a board that owns a real agenda, and keep reassessing as the firm matures.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST07

RECOMMENDED ACTIONS

1. **Choose an operating-model archetype** on the centralization spectrum and set explicit decision rights for steering, standard setting and execution.
2. **Fix a reporting drumbeat**, such as a weekly solution-team review run by a small coordinating group, building toward a central team that coordinates business-owned solution teams rather than commanding them.
3. **Define where board-level AI oversight sits** and set a cadence for the AI agenda: what is presented, by whom and how often.
4. **Give every AI asset and domain a single named owner**, a regular review and a change log.
5. **Reassess the operating model and its standards on a regular cycle**, planning a swing from central control toward a federated model as maturity grows.
6. **Match the funding mechanism to the centralization choice**, with flexibility to adapt as the firm matures.
7. **Track the cadence**, starting with adherence and time to decision, then a short metric set built into the transformation plan.

SEQUENCE AND DEPENDENCIES

1. In McKinsey's 2024 study, the financial institutions that handled this transition best already had a high level of organizational agility, letting them rework processes quickly and pool resources flexibly.
2. Where that agility is absent, in this guide's view a centrally led model that concentrates scarce talent and builds a repeatable delivery model is the safer start.
3. Funding is a practical constraint on the archetype choice, since a highly centralized model implies central funding while a federated one pushes funding into the business.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Runs the solution teams and the reporting drumbeat; owns execution and the progression of use cases through the cadence.
Second Line	Holds a standing seat in the governance forums; owns the risk-and-control standards and the sign-off that gates production.
Third Line	Assesses the design and operation of the AI governance forums, whether they own their decisions.
Board and Executive	Owns where AI oversight sits and the cadence of the board agenda, and is accountable for the decisions the forums make.
Cross-Functional	Maintains the shared architecture, data and control standards and the common vocabulary so steering, standard setting and execution stay aligned.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST07

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Unowned or theatrical governance**	Define where board and committee AI oversight sits and set a fixed agenda cadence (what, by whom, how often).
MEDIUM **Centralization friction**	Set decision rights up front; keep a central standard-setting spine if execution is later delegated; stage any change in the balance.
MEDIUM **Command-and-control drag**	Keep the central team small and coordinating; the business owns solution teams and execution.
MEDIUM **Uncontrolled vendor AI activation**	Record the feature in the AI inventory before use (P05-ST10); name a business owner (RA 4); check change notices and disable rights (P07-ST04).

METRICS SCORECARD

AREA	LEADING	LAGGING
Cadence adherence	Forum attendance; standing-agenda adherence; items tabled	Decisions made vs deferred per cycle
Scaling progression	Use cases entering the execution stage	Share of use cases reaching production
Decision velocity	Items tabled per cycle	Time from tabling to decision
Governance freshness	Scheduled model / standard reviews held on time	Overdue reviews; unreviewed AI assets



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST07

AMAT

P08-ST08

Industry Collaboration and Common Standards: Shaping the Rules Rather Than Reacting to Them

For: Senior Leadership; Board and Executive

No asset manager sets the rules for AI on its own. Firms that treat industry collaboration and standard-setting as a strategic activity, rather than an external compliance event, get the chance to shape standards before they are set. In the World Economic Forum's 2024 playbook for investors, 149 global institutional investors holding more than US$1.66 trillion signed a single investor statement on the EU AI Act. This article covers how to use the sector's collective weight, how to engage regulators and industry bodies while the rules are still being written, and how to manage the shared risks no single firm can control alone. It is the thinnest sub-topic in the pillar, so it leads on collaboration evidence and actions.

US$70 trillion

in assets represented by those behind the International Corporate Governance Network's AI engagement guide (WEF, 2024)

149

global institutional investors, holding more than US$1.66 trillion, signed a single statement on the EU AI Act

WHERE FIRMS GO WRONG

The default posture is to treat standards as something that happens to the firm: wait for the rule, then comply. That is a strategic error while the rules are still being formed, because it forfeits the chance to present the firm's perspective. The second gap is shared risk. Concentration on a handful of AI providers, correlated model behavior and third-party dependencies are sector-level exposures that no single firm can manage alone.

Two separate examples of investor coordination on AI

around US$70tn	over US$1.66tn
Investors leading the ICGN AI engagement guide: assets under management	149 global institutional investors signing an EU AI Act statement: assets

Source: World Economic Forum, 2024. Different groups; the amounts are not additive.

Takeaway. Treat collaboration as a capability: pick the forums that matter, engage before the rules are set, use collective influence, adopt the common floor standards and bring shared risks into the room. The evidence base is thin, but the choice is not: shape the standards, or comply with someone else's.

Read the full article online

Full evidence and references, case examples and implementation detail.

amatguide.org/amat-guide#P08-ST08

RECOMMENDED ACTIONS

1. Start by listing the regulators that bind the firm and the associations, coalitions and standard setters that influence market practice, ranked by relevance; build toward deep, sustained contribution in the few that matter rather than spreading thin.
2. Engage regulators and bodies while standards are still being drafted, and present the asset-management perspective rather than reacting after the rule is set.
3. Use collective weight, investor statements and coalitions, to contribute to standards a firm cannot shape as effectively alone.
4. Start by mapping the industry reference practices (e.g., acceptable-use, vendor-evaluation and adversarial-AI papers) to the firm's controls; build toward adopting the relevant ones and contributing to their next versions.
5. Push for shared evaluation benchmarks for AI and LLM use, and for coordinated domestic and international standard-setting to reduce fragmentation.
6. Collaborate on the shared and systemic risks that cross firm boundaries (provider concentration, correlated model behavior, shared-infrastructure vulnerabilities): start by raising one of them in a forum the firm already sits on; build toward standing collaboration on all three.
7. Respond to relevant consultations from IOSCO or national regulators when they are open, so the asset-management perspective is on the record before rules are finalized.

SEQUENCE AND DEPENDENCIES

1. The main practical constraints are sustained ownership, access to the relevant forums, and the capacity to turn technical positions into usable consultation responses.
2. Senior attention matters, but firms should also budget policy, risk, legal and subject-matter support for material engagements.
3. Standard-setting engagement only works when someone with authority contributes substantively and consistently.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Implements adopted standards in day-to-day workflows and feeds practical experience back into the firm's collaborative positions.
Second Line	Monitors regulatory developments, reviews proposed positions for compliance and conduct implications, and challenges the evidence behind requested policy changes.
Third Line	Provides assurance that adopted external standards are actually implemented and kept current, not just signed up to.
Board and Executive	Owns the decision to engage, which coalitions to join, and the use of the firm's collective influence; accountable for AI outcomes whatever standards it shares.
Cross-Functional	Keeps a single approved position log so the firm does not contradict itself across bodies, aligned with internal policy and strategy.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST08

KEY RISKS AND CONTROLS

RISK	CONTROL
MEDIUM **Limited influence over emerging standards**	Maintain a regulatory horizon scan and engage associations and regulators selectively while standards are in draft.
MEDIUM **Regulatory fragmentation**	Advocate for coordinated domestic and international standard-setting, aligned with bodies such as IOSCO and the FSB, to reduce fragmentation.
MEDIUM **Unmanaged shared or systemic risk**	Collaborate on shared risks through association and regulator forums, raising only aggregated observations that have been cleared for sharing.

METRICS SCORECARD

AREA	LEADING	LAGGING
Participation	Material consultations identified and assigned an owner	Share receiving a timely, substantive response
Influence	Specific, documented buy-side proposals submitted	Proposals acknowledged or reflected in consultation feedback or final language, with no assumption that participation caused the outcome
Standards adoption	Relevant reference practices mapped to the firm's controls	Improvements implemented, with evidence of testing or review



Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST08

AMAT

P08-ST10

Knowing Where You Stand: AI Maturity Models and Benchmarking

For: Senior Leadership | Risk and Compliance | Technology | Data and Analytics | Internal Audit

Most financial-services firms can now say they use AI. Far fewer can say, with evidence, how mature that use is, and in CCAF's 2026 survey 81 percent of respondents were adopting AI at some level while only 14 percent viewed their deployment as transformational. A maturity model is a diagnostic, not a scorecard: it forces an honest baseline, exposes capability gaps that adoption alone hides and converts a vague sense of progress into a staged roadmap. The article reads maturity through three lenses together, firm-wide maturity, data readiness and assurance maturity, because a single headline score misleads when it rests on weak data or immature controls. Benchmarking against an external framework stops a firm grading its own homework.

14%

of surveyed financial-services respondents viewed their AI deployment as transformational, against 81% adopting at some level (CCAF, 2026)

WHERE FIRMS GO WRONG

The first failure is treating a maturity model as a questionnaire to complete rather than an assessment to evidence; a firm that scores itself from opinion produces a comfortable number that can collapse when tested. The second is stacking frameworks: the developers of one 2026 model reviewed more than one hundred existing efforts and found inconsistent definitions, so the lesson is to choose one spine and apply it consistently. The third, and most consequential, is reading maturity through a single lens when the constraint often sits in data and assurance rather than in the models.

AI maturity spectrum (assessed per dimension)



Source: Carnegie Mellon SEI, AI Adoption Maturity Model (2026).

Takeaway. Pick one credible model, score it against evidence, read all three lenses together, benchmark assurance against an external standard and reassess on a cadence. The advantage lies in knowing, and being able to prove, exactly where the firm stands.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST10

RECOMMENDED ACTIONS

1. Choose one credible maturity model and apply it consistently across assessment cycles; avoid stacking or averaging multiple frameworks, which undermines comparability.
2. Run the assessment against evidence, not opinion: consolidate stakeholder input, metrics, tooling data, artifact reviews, and operational data.
3. Assess all three lenses (firm-wide maturity, data readiness, assurance maturity): start by running the data-readiness assessment alongside the firm-wide one, testing whether data is the binding constraint; build toward all three lenses assessed in parallel every cycle.
4. Benchmark assurance maturity externally: start by checking one material AI workflow's pre-deployment tests, monitoring records and change approvals against the firm's stated control requirements; build toward benchmarking the whole testing, audit and control practice against an external reference such as ISO/IEC 42001 rather than the firm's own history.
5. Use the baseline to build a staged roadmap: start by ranking the gaps the baseline exposes and sequencing the first fixes; build toward a roadmap that sets out the level, capability gaps, priority use cases and the sequence of fixes.
6. Give every maturity baseline a refresh date. Start by updating the critical-gap dimensions each quarter; build toward a full reassessment annually, with an ad hoc reassessment on material regulatory change.

SEQUENCE AND DEPENDENCIES

1. The first practical point is that the assessment must be resourced as an evidence-gathering exercise, not a workshop.
2. The second is sequencing across the three lenses.
3. The third is that maturity assessment is a governance and data discipline as much as a technology one, and it should be owned accordingly.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Supplies the operational evidence the assessment consumes, keeps the workflow and system population current, and owns closing the workflow and data gaps the maturity profile exposes.
Second Line	Sets the challenge criteria and tests whether management has justified its material scores and risks.
Third Line	Independently reviews the assessment process, the traceability of evidence behind scores and selected controls on a risk basis, without owning the maturity score.
Board and Executive	Approves the assessment scope, sets the cadence, insists on evidence behind every dimension score.
Cross-Functional	Runs the assessment across the three lenses, maintains the single spine model, and turns the baseline and gap map into the staged roadmap.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST10

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Vanity self-assessment**	Assess against consolidated evidence (metrics, tooling data, artifacts, operational data); require the evidence behind each dimension score.
HIGH **Unmeasured material foundation**	Assess all three lenses in parallel and rank gaps by consequence for each priority use case, not by the lowest score alone.
HIGH **Scope and evidence drift**	Freeze the assessment scope and system list for each cycle, reconcile it with the AI inventory (P05-ST10), record exclusions and missing evidence, and compare like-for-like populations between cycles.
MEDIUM **Framework stacking and drift**	Choose one credible spine and apply it consistently across assessment cycles.

2 more risks in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Assessment coverage and integrity	Share of material AI workflows in the inventory with an evidence-backed score	Assessment findings that survive audit or board challenge
Firm-wide maturity	Number of dimensions assessed against the spine model	Movement in maturity level between cycles, with scope held constant
Data readiness	Data-readiness dimensions assessed	Share of AI projects blocked by data-maturity gaps
Production conversion	Use cases passing readiness assessment before build	Share of AI reaching production

2 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

A simple estate makes a single-model assessment quick to run: adopt one credible spine and put the effort into the data-readiness lens, where, in this guide's view.

MID-SIZED

Capacity squeeze

The constraint is having the evidence base to score honestly: build the metrics and tooling data that a credible assessment consumes before assigning a level.

LARGE AND GLOBAL

The assessment is multi-dimensional across a complex estate, so read the dimensions against the use cases they affect rather than the average.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST10

AMAT

P08-ST12

The Agentic Turn: Autonomous Workflows and the AI-Native Firm

For: Senior Leadership | Board and Executive | Investment Teams

The next phase of AI in asset management is not a better chatbot; it is the agent, a system that plans, decides and acts across multiple steps rather than waiting to be prompted. In EY's early-2025 survey of 100 wealth and asset management firms, 78 percent were exploring agentic AI, yet just 7 percent of asset managers had deployed it. The agentic future will be paced by governance and reliability, not model capability: the constraint is whether a firm can let an agent act safely. Autonomy is a spectrum to be earned, with narrow, well-governed workflows first and oversight and failure response built before scaling. Only about one in five organizations has tested a response plan for when an agent acts wrongly.

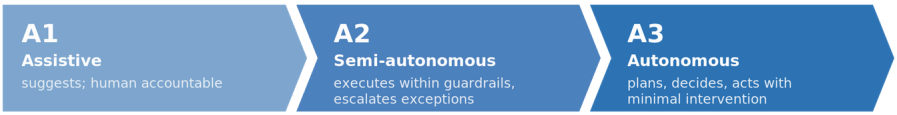
7%
of surveyed asset managers had deployed agentic AI, though 78% were exploring it (EY, early 2025, 100 wealth and asset management firms)

49%
of respondents using or assessing agentic AI said agents were already deployed (NVIDIA, 2026; vendor-published survey of 839)

WHERE FIRMS GO WRONG

The first failure is mistaking experimentation for capability: in McKinsey's 2025 global survey, 62 percent of respondents were at least experimenting with agents but only 23 percent were scaling one anywhere in the enterprise. The second is underestimating the reliability and skills gap that autonomy exposes, with performance and reliability issues (34 percent) and a lack of internal skills to manage and monitor agents (33 percent) the leading challenges in NVIDIA's survey. A firm that scales autonomy faster than it builds the capability to monitor it accumulates a hidden liability.

Agentic autonomy as three authority bands



Recorded separately, not on the scale: multi-agent or orchestrator; tool-using
Listed in the source without a place on the spectrum

Source: AMAT framework, drawing on World Economic Forum and Accenture, The AI Playbook for Financial Services (2026).

Takeaway. Treat autonomy as something to be earned: define the spectrum, pilot where supervision is cheapest, build failure response before scaling and close the skills gap. The advantage lies with the firms whose autonomy is the most governed, not those with the most autonomous systems.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST12

RECOMMENDED ACTIONS

1. **Adopt an explicit autonomy spectrum** (assistive, semi-autonomous, autonomous, multi-agent) and place every agentic use case on it before deployment.
2. **Start agentic pilots where autonomy is cheapest to supervise and highest in value**, such as client-support and operational or risk workflows.
3. **Build failure-response and oversight machinery before scaling autonomy**, including a tested response plan for when an agent acts wrongly.
4. **Close the reliability and skills gap** by naming who monitors each live agent and what they watch for before its remit expands.
5. **Convert pilots into scaled, governed workflows** by giving every pilot a dated scale-or-retire decision.
6. **Sequence toward multi-agent orchestration**, scaling one proven use case on a shared platform and expanding autonomy only as governance keeps pace.

SEQUENCE AND DEPENDENCIES

1. The first practical point is that agentic AI should be sequenced onto existing foundations, not bolted on as a novelty.
2. The second is that the move from single-function pilots to enterprise orchestration is an organizational change as much as a technical one.
3. The third is that the human-and-machine operating model is the live design question underneath all of this.

WHO DOES WHAT: LINES OF DEFENSE

First Line	Owns the earned-autonomy discipline in practice: classifies each agent in an autonomy band, runs supervised pilots, and keeps a human accountable for agent actions.
Second Line	Sets the guardrails and escalation thresholds for each autonomy level and challenges autonomy creep before an assistive agent becomes an autonomous one unnoticed.
Third Line	Tests whether agent oversight and failure-response plans actually exist and work; detailed agent assurance and testing sit in Pillar 7 (P07-ST01 for testing before deployment).
Board and Executive	Owns accountability for autonomous systems acting on the firm's behalf, personal where a regime such as SM&CR applies; sets the risk appetite for autonomy.
Cross-Functional	Designs the orchestration architecture, maintains the autonomy-spectrum standard, and coordinates the move from single-function pilots to enterprise orchestration.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST12

KEY RISKS AND CONTROLS

RISK	CONTROL
HIGH **Ungoverned autonomy**	Build and test a failure-response and escalation plan before scaling autonomy; match oversight to the autonomy band; name who can suspend each agent and revoke its credentials.
HIGH **Reliability and skills gap**	Invest in monitoring and the internal skills to supervise agents before expanding their remit.
HIGH **Unauthorized action and propagation**	Managed identity, least-privilege permissions, an approved tool list, transaction limits, action logs, pre-execution gates for material actions, and revocable access.
MEDIUM **Autonomy creep**	Classify every use case in an autonomy band and require a governance decision to move it up a band.

1 more risk in the full article

METRICS SCORECARD

AREA	LEADING	LAGGING
Autonomy governance	Share of agentic use cases classified in an autonomy band	Incidents from agents acting outside their intended band
Failure readiness	Response plans tested per autonomous use case	Time to contain an agent failure
Pilot-to-production	Pilots with a defined scale-or-retire decision	Share of agents scaled beyond one function
Reliability and skills	Monitoring coverage of live agents; agent-supervision skills in place	Reliability incidents per agent in production; unauthorized or failed actions per 1,000 executed actions

4 more metrics in the full article

BY FIRM SIZE

BOUTIQUE

Agentic capability will arrive mostly through vendors: choose one narrow, well-supervised use case, such as client-support drafting with human approval where the firm has a wealth-management client base.

MID-SIZED

Capacity squeeze

Focus on a small number of governed pilots in operational and risk workflows and invest early in the monitoring and skills that live agents require; in this guide's view.

LARGE AND GLOBAL

The largest firms have the resources to pursue multi-agent orchestration across functions but also the hardest governance problem: sequence deliberately, scale only proven use cases, and treat cross-functional coordination.

Read the full article online
Full evidence and references, case examples and implementation detail.
amatguide.org/amat-guide#P08-ST12

WHAT THE DIGITAL EDITION ADDS

- Every sub-topic article in full, with evidence and case examples
- Risks and controls, metrics, and guidance by firm size and line of defense
- The complete references for every figure and claim
- Updates as the guide evolves

REGISTER FOR UPDATES

amatguide.org/register

BUSINESS & ECONOMICS / INVESTMENTS & SECURITIES

Building the AI-Native Asset Management Industry

Almost every asset manager now has the same question on its desk: how to adopt AI safely, within regulation, and beyond the pilot stage. Most are still answering it alone, with duplicated effort, no shared standards, and pilots that stall before production.

This guide sets out one practical framework across eight pillars, from strategy and use cases to governance, assurance, and scaling. Its Chapter Leads are running AI transformation inside their own firms, and an Industry Review Panel tested the framework against the operational constraints that investment, technology, risk, and compliance teams face every day.

INSIDE THE GUIDE · 8 STRATEGIC PILLARS

01	Strategy, Vision, Leadership
02	AI Use Cases
03	Data, Knowledge, and Decision Intelligence
04	Technology Architecture, Infrastructure, and Model Layer
05	Governance, Risk, Regulations, and Controls
06	People, Talent, Culture, and Change Management
07	Assurance, Audit, Testing, and Third-Party Oversight
08	Scaling, Value Realization, and Industry Transformation

AMAT

An industry-led initiative. Built by practitioners.

amatguide.org

ISBN 979-8-9823233-0-9

9 798982 323309 >