
AI OPERATING MODEL STRATEGY

Institutional AI Execution: From Adoption to Impact

A cross-source synthesis on capturing AI value, drawn from the perspectives of McKinsey, BCG, and Goldman Sachs.

Executive Summary

Three major institutional intelligence sources — *McKinsey's People & Organizational Performance Practice*, *Boston Consulting Group's Asset Management Executive Perspectives*, and *Goldman Sachs Global Banking and Markets* — converged on a single structural conclusion in mid-2026: AI adoption is now widespread, but enterprise-scale value capture remains the exception.

The numbers make the gap concrete.

McKinsey's global survey finds that 70 percent of employees feel personally ready to work with AI, while only 27 percent of leaders believe their organizations are ready to make the shifts required, and just 11 percent of organizations have reached the reinvention horizon.¹

BCG, drawing on work with more than 2,500 clients, puts the unclaimed prize at 300 to 500 basis points across the P&L for asset managers, the difference between a 30 percent margin and a 40-plus percent margin business, while noting that only about 30 percent of firms have generated real AI value to date.²

Goldman Sachs frames the surrounding environment as industrial transformation unlike any in modern history: hyperscalers alone are projected to invest more than \$6 trillion in AI through 2030, within a total buildout estimated at \$7 trillion or more between 2026 and 2031.³

The gap between experimentation and execution is no longer a technology problem. It is an organizational design problem, a governance problem, and increasingly a capital deployment sequencing problem. McKinsey's central quantitative finding is that organizational readiness accounts for 48 percent of the difference between leaders who capture value from AI and those who do not, against 25 percent for personal readiness; the organization is nearly twice as decisive as the individual.¹

These findings extend and reinforce NextFi Advisors' published thesis on the Convergence Economy: that two cost inputs (cognitive labor and transaction coordination) are being compressed simultaneously, generating structural surplus that is real, persistent, and contested. The research examined here confirms that most institutions are still on the wrong side of that contest.

For institutional leaders, the immediate priority is to convert these signals into a phased execution plan that is commercially viable, regulator-ready, and operationally durable.

Section I: Source-Specific Signals

The Three Sources at a Glance

Source	Scope	Method / evidence base	Core claim	Headline number
McKinsey (July 2026)	Cross-industry, organizational	Global survey of 750 employees and leaders; 608 leaders on organizational readiness; data collected Feb–Apr 2026	AI does not create enterprise value simply because more people use it	70% personally ready vs. 27% organizationally ready
BCG (July 2026)	Asset management, operational	Work with 2,500+ clients over three-plus years; industry benchmarking	Agentic AI will separate those who implement from those who do not	300–500 bps of P&L value available
Goldman Sachs (2026)	Macro, capital markets	Firm research and transaction experience; letter from Dan Dees	The AI era is an industrial and capital-formation event, not a software story	\$7T+ buildout, 2026–2031

A note on method: McKinsey discloses that recruitment deliberately targeted respondents from organizations in more advanced horizons, and that its estimates "may not reflect overall market prevalence." The practical implication cuts in favor of this brief's argument rather than against it: the true share of institutions still in the earliest horizon is likely *higher* than the survey suggests.

McKinsey — Three Horizons of AI Transformation (July 2026)

McKinsey's People & Organizational Performance Practice, with QuantumBlack, AI by McKinsey, presents findings from a global survey of 750 employees and leaders across North America, Europe, Asia–Pacific, Latin America, and EEMEA, collected February to April 2026. The central observation is direct: AI does not create enterprise value simply because more people use it.

McKinsey groups organizations into three horizons based on how deeply AI is embedded in daily work:

Horizon 1: Enablement: individual employees are given access to general-purpose AI tools that assist with parts of their existing jobs.

Horizon 2: Automation: AI is used at scale to automate and improve *existing* end-to-end, cross-functional workflows. These organizations are still starting from the current state and optimizing it.

Horizon 3: Reinvention: roles, workflows, and operating models are redesigned from a blank slate around what AI makes possible.

Nearly 90 percent of organizations remain in the first two horizons. Only 11 percent of leaders place their organization in reinvention, and that group is by far the most likely to report enterprise value, at 48 percent, against 24 percent in automation and 13 percent in enablement.

Three findings deserve particular attention from institutional leaders:

- **The readiness gap is the headline.** Seventy percent of respondents feel personally prepared to adopt and use AI; only 27 percent of leaders believe their organizations are ready to make the people and culture shifts required. Employees are adapting faster than the institutions they work in.

Even within the reinvention horizon, 44 percent of leaders say their organizations are not ready, rising to 68 percent in automation and 84 percent in enablement.

- **Organizational readiness is nearly twice as decisive as personal readiness.** It accounts for 48 percent of the difference between leaders who report capturing value and those who do not, against 25 percent for personal readiness. Tool distribution is not the lever; workflow, operating model, leadership behavior, and culture are.
- **Workflow redesign pays from Horizon 1 onward.** Leaders are 5.3 times more likely to report enterprise value capture when workflows are redesigned than when they are left unchanged (32 percent versus 6 percent). Two further multipliers matter: 3.9x where leadership teams demonstrate high AI fluency (35 percent versus 9 percent), and 3.3x where leaders receive the support and training to build new skills (30 percent versus 9 percent).

McKinsey also surfaces an underweighted human capital dimension. AI transformation alters how people work, what they are responsible for, and how they perceive their value. Trust in the organization is the critical readiness factor across all three horizons. Middle managers report the highest anxiety (roughly one in four, against one in five individual contributors) and respondents reporting low trust in their organization's support are 1.5 times more likely to feel anxious about AI-related change.

McKinsey draws a distinction worth carrying into any transformation program: reducing anxiety and building trust are not the same thing, and reassurance that cannot hold over time erodes the second while only temporarily achieving the first.

The paper closes on an analogy that is useful in board settings. Early electrification manufacturers replaced steam engines with electric motors but left factories, workflows, and management systems unchanged; productivity gains stayed limited until factories were redesigned around electricity itself.

NEXTFI READ

The 5.3x workflow-redesign multiplier is the most actionable statistic in the McKinsey dataset, because it is available to institutions that have not yet left Horizon 1. **Value does not require reaching the reinvention horizon, but it does require redesigning.**

BCG — AI-First Companies Win the Future (July 2026)

BCG's July 2026 Executive Perspective targets asset management specifically, drawing on work with more than 2,500 clients over three-plus years. It is framed around the question executives are now bringing to BCG: two years of AI pilots, and the ROI is underwhelming. What are the best doing differently?

The industry context BCG establishes is as important as its AI prescription. Assets under management grew 11 percent to \$147 trillion in 2025, but more than 80 percent of revenue growth came from market performance rather than flows. Average management fees are down 23 percent since 2010, with fee compression continuing at 1 to 3 percent annually, and costs have outpaced revenues for fifteen consecutive years (cost CAGR 5.4 percent against revenue CAGR 5.1 percent). The AI question is

therefore not discretionary technology spending. It is the available answer to a structurally broken growth model.

The quantified prize

BCG sizes an AI-first asset manager at 300 to 500 basis points of value across the P&L: "the difference between a 30 percent margin and a 40-plus percent margin business." The components:

Lever	BCG estimate
Research coverage	2x–5x
Sharpe ratio improvement	5%–15%
Client coverage per relationship manager	3x–5x
RFP response capacity and speed	5x–10x
Speed to market for new products	2x–3x
Trade error reduction	50%–70%
Operations capacity unlock	55%–65%
Net cost reduction over three to five years	25%–30%
Incremental net inflows	0.5%–1.0% of AUM

BCG puts asset managers behind banks, payments, telecom, and software on AI maturity, but finds AI leaders generating 3.8x higher total shareholder returns. Roughly 30 percent of organizations have generated real AI value, up 2.2x from 2023, and the agentic share of AI value is expected to roughly double from 16 percent to 33 percent by 2028.

The end-to-end agentic model

BCG's scope covers the full value chain, with capacity-unlock estimates by a financial institution's function:

- Investment operations and HR up to 60 percent
- Marketing and finance up to 55 percent
- Investment research, portfolio construction and risk, and IT up to 50 percent
- Sales up to 45 percent
- Product, trading, and legal and compliance up to 40 percent
- AI and data governance up to 30 percent

Underneath sits a hub-and-spoke agent architecture; a research orchestrator routing work to specialized market-context, screening, fundamental-analysis, signal-generation, portfolio-construction, risk, and stewardship agents, fed by a unified investment knowledge graph with per-agent subgraph projections, MCP and API connectors to Bloomberg, Refinitiv, FactSet, Aladdin, ISS, and MSCI, and infrastructure-level requirements for immutable audit trail, multi-jurisdiction data residency, and information barriers between mandates.

Three plays and seven steps

BCG frames execution as three plays: 1) **Deploy** table stakes, 2) **Reshape** where value concentrates, and 3) **Invent** new models. BCG also gives executives the following seven steps:

1. Set ambition and tone from the top with CEO sponsorship and protected budget
2. Pick two or three workflows and go all-in with specific P&L targets

3. Invest in activation and literacy alongside technology
4. Fix the data the agents need; build governance in from day one
5. Tie AI outcomes to incentives and scorecards
6. Design the organization of the future with agent engineers embedded in the business

Two BCG lines are worth quoting directly to a board. On governance: *"Retrofitting is exponentially harder."* On ambition: *"Half-measures lead to pilot purgatory."*

Peer movement is already observable

Norges Bank Investment Management reports that half of its 700 staff code their own AI tools, with more than 213,000 hours saved. Bridgewater has incorporated machine learning as a "sub-advisor" and launched a \$2 billion AI fund. Man Group is building AlphaGPT with Anthropic.

Ninety percent of asset management organizations are considering agentic deployment within six to twelve months. Most consequentially for capital formation: institutional allocators are now asking about AI in due diligence questionnaires, and 12 to 15 percent of retail investors already use generative AI to support investing decisions.

NEXTFI READ

The DDQ signal converts AI capability from an efficiency story into a fundraising story. That is a materially different conversation with a CEO, and it is the argument most likely to unlock budget.

Goldman Sachs — Harnessing AI for the Real Economy (2026)

Goldman Sachs Global Banking and Markets, in a report opening with a letter from Dan Dees, Co-Head of Global Banking and Markets, situates AI not as a software productivity story but as an industrial infrastructure event: the AI era is driving industrial transformation unlike any in modern history — faster, broader, and more capital-intensive, creating new infrastructure and industry simultaneously.

Key factual anchors:

- Hyperscalers alone are projected to invest more than \$6 trillion in AI through 2030, many times the capital deployed to internet infrastructure during the dotcom era.³
- Total capital requirement for compute, power, and data centers is estimated at \$7 trillion or more between 2026 and 2031, with baseline aggregate estimates around \$7.6 trillion.³
- Global hyperscaler capex is projected to exceed \$760 billion in 2026, roughly \$2 billion per day. Global data center supply grew from 30 GW in 2019 to 57 GW in 2024, with a further 65 GW projected online by 2030.⁴
- Most of the capital, infrastructure, and M&A that will define the AI economy has not yet been deployed, built, or executed.

Scale in context — the argument the headline numbers obscure

Even at \$7 trillion-plus, US AI investment equals only about 1.2 percent of today's annual GDP. Goldman notes that while that share will rise if current projections are realized, even the most

optimistic scenarios keep it well below the railroad buildouts of the 1800s, which ran 3 to 4.5 percent of GDP.⁵ The figures are large in absolute terms and historically moderate in relative terms — a distinction that matters for any institution being asked whether it is underwriting a bubble.

Where the value actually lands

Goldman's most consequential framing for this brief's audience is proportional: software-as-a-service accounts for less than 0.5 percent of global GDP. The real economy — the other roughly 99.5 percent, from manufacturing and robotics to defense, construction, and energy — defines the actual scale of the opportunity.⁶

The industrial deployment is already measurable: Siemens reports a 40 percent reduction in production downtime from AI-driven predictive maintenance at its Amberg plant; Figure AI's humanoid robots ran ten-hour daily shifts for more than five months on BMW's Spartanburg line, contributing to over 30,000 vehicles across roughly 1,250 operating hours.

Capital architecture is the report's core

Funding a buildout of this scale requires the full capital stack — equity, public and private debt, sovereign capital, and new joint-venture structures, some of which have yet to be invented. Goldman describes a "complexity premium": where the railroad era turned on a single instrument (long-tenor infrastructure bonds) and the internet era on two (venture and growth equity, technology M&A), the AI economy requires six or more simultaneously.

The mechanics matter to any institution positioning as an intermediary:

- **Index concentration is the binding constraint.** AI-related debt issuance from hyperscalers and AI infrastructure companies reached roughly \$121 billion in 2025 — about 6.6 percent of the \$1.83 trillion US investment-grade market — and is projected to rise to as much as 20 percent in 2026. Allocator position limits, sector caps, and single-name thresholds were not calibrated for a digital-infrastructure share approaching one-fifth of the index. The constraint binds on both sides and pushes issuance into EUR, GBP, and CHF, and into private markets.
- **Private markets are absorbing the overflow.** Private credit has grown from roughly \$500 billion in AUM a decade ago to over \$2.1 trillion. Infrastructure funds raised a record \$221 billion last year and may reach \$3 trillion in assets by 2030.
- **Public and private technique are blending.** Meta's \$27 billion "Beignet" financing for its Hyperion data center brought amortizing, cash-flow-matched project debt into the public investment-grade market at scale, with comparable deals following in 2026.
- **New collateral classes are being underwritten.** Private credit is structuring facilities against GPU fleets — short-duration, rapidly depreciating assets — using shorter tenors, embedded refresh provisions, and hyperscaler offtake as the credit anchor. Sale-leasebacks let data center owners monetize built assets while retaining control. Robot-as-a-Service is shifting robotics procurement from capex to opex.

- **Deal scale is resetting.** Goldman acted as lead left bookrunner on Alphabet's \$90 billion equity raise in June 2026, the largest equity offering on record. NextEra Energy's announced all-stock acquisition of Dominion Energy carries an enterprise value above \$120 billion.

The report also tracks sector-specific AI deployment across defense, energy, and industrial verticals, with notable capital events in European dual-use AI. Helsing, founded in 2021, reached a €12 billion valuation in a €600 million Series D in June 2025.⁷ PhysicsX raised a \$300 million Series C in June 2026 at a roughly \$2.4 billion valuation, its AI-native physics simulation compressing aerospace, semiconductor, automotive, and energy engineering cycles from days to seconds.⁸

NEXTFI READ

Goldman's contribution is macro-structural, and its sequencing claim runs opposite to the other two sources. McKinsey and BCG describe a phase that is *ending*: **pilots**. Goldman describes a phase that is *beginning*: **financing**.

Institutions that read only the first pair will optimize internally while missing the intermediation opportunity; institutions that read only Goldman will chase the mandate without the operating model to service it.

Section II: Cross-Document Themes and Tensions

Where the Three Sources Align

The convergence across McKinsey, BCG, and Goldman Sachs is substantial, and the three sources reinforce one another at multiple levels.

Theme 1: The pilot era is over — for internal transformation

McKinsey finds nearly 90 percent of organizations short of reinvention and the majority reporting no meaningful enterprise value. BCG frames its 2026 guidance around two years of underwhelming pilot ROI and warns that half-measures lead to pilot purgatory. Both point the same direction: experimentation as a strategy has run its course.

Goldman should not be enlisted for this claim. Its position on sequencing is explicitly that most capital, infrastructure, and M&A *has not yet* been deployed. The internal transformation phase is ending; the external financing phase is only beginning. Collapsing the two produces a false sense that the window has closed.

Theme 2: Value capture requires organizational redesign, not tool deployment

McKinsey is most explicit and now most quantified: organizational readiness accounts for nearly twice the value-capture difference that personal readiness does, and redesigned workflows deliver a 5.3x multiplier. BCG's end-to-end agentic model carries the same message from the operational side — its instruction is to pick two or three workflows and redesign them end-to-end with named P&L targets, not to distribute chatbots.

Theme 3: Governance and trust are bottlenecks as significant as technology

McKinsey identifies trust in the organization as the critical readiness factor across all three horizons, with middle-manager anxiety at one in four. BCG makes governance step five of seven and warns that retrofitting is exponentially harder, specifying audit-grade transparency, human-in-the-loop boundaries, immutable audit trail, and multi-jurisdiction data residency. Goldman's capital markets framing requires regulator-ready structures for financing instruments that do not yet have settled precedent.

Theme 4: Financial institutions occupy a dual role

Banks, asset managers, and funds must transform their own operating models using AI while simultaneously serving as the capital formation and structuring engine for the broader AI economy. Goldman makes the external role explicit; BCG and McKinsey frame the internal imperative. No single source states the dual mandate — but it follows directly from reading them together, and it is the most consequential synthesis available here.

Where the Sources Create Tension

Tension 1: Urgency versus sequencing

Goldman's macro framing implies a narrow positioning window. BCG's pilot-ROI diagnosis suggests that speed without structure has already failed for most institutions. McKinsey's horizon model implies that sustainable progress requires sequential capability building rather than acceleration at all costs. Institutions face pressure to move fast from one direction and evidence that undisciplined speed has not worked from another.

Tension 2: Scope of transformation

BCG is asset-management-specific and operationally granular. McKinsey is cross-industry and organizational. Goldman is macroeconomic and capital-markets-centric. Each offers a different resolution to the same pressure — what transformation actually looks like at the institutional level. There is no single answer across these sources, which is itself a signal that institutions should resist importing generic frameworks wholesale.

Tension 3: Optimism about the window

BCG's stance is that the window is open *because* most competitors have not moved beyond pilots, and that committing now compounds advantage. McKinsey's data is more sobering: 44 percent of organizations that have reached reinvention still describe themselves as not ready for the required people and culture shifts. Both can be true, and the reconciliation is uncomfortable — the advantage is available, and arriving there does not feel like arrival.

NEXTFI ASSESSMENT

The tension between urgency and sequencing is the most operationally significant finding in this synthesis. It maps directly to the failure mode we observe most often in client engagements: organizations that accelerate deployment without first resolving data architecture, governance, and organizational alignment.

Fast pilots become expensive liabilities. McKinsey's 5.3x workflow-redesign multiplier, read alongside BCG's instruction to pick two or three workflows and go all-in, points to a specific resolution: depth over breadth.

Institutions should concentrate effort on a small number of end-to-end workflow redesigns with named P&L targets through Q1 2027, not broadening pilot coverage.

Section III: Alignment with NextFi Intelligence

These findings extend and, in several areas, sharpen prior NextFi Advisors published intelligence.

- **"The Convergence Economy: Nine Strategic Imperatives"** established that AI compresses cognitive labor costs while programmable settlement infrastructure compresses transaction coordination costs. McKinsey and BCG confirm the cognitive labor compression dynamic while demonstrating that organizations are not yet capturing the surplus that compression should generate.

McKinsey quantifies the shortfall at the organizational readiness layer specifically. The failure is organizational, not technological, consistent with NextFi's thesis that execution, governance, and alignment determine outcomes.

- **"Multi-Agent Design Models for Financial Institutions"** anticipated the BCG agentic architecture for asset management. BCG's hub-and-spoke orchestrator model, specialized agent taxonomy, MCP-based data connectivity, and per-agent knowledge graph projections correspond closely to the multi-agent workflow architectures NextFi has modeled for institutional distribution, research, and client management. The BCG publication confirms that these designs are now being presented at C-suite level as production-ready rather than experimental.
- **"Three Trends, One Infrastructure Stack"** identified the convergence of AI workflow automation, digital asset rails, and institutional operating model redesign as a single structural force. Goldman's capital stack analysis reinforces the third leg: the financial institution of the next decade is simultaneously a transformation target and a capital intermediary for the AI buildout.
- **Context Engineering.** McKinsey's finding that organizational readiness outweighs personal readiness by nearly two to one is the empirical form of the argument advanced in our *"Context Engineering: How Enterprise Operating Models Impact AI Performance & Output."* Model capability is not the constraint; the operating model surrounding the model is.

The gap worth naming

None of the three sources integrates the digital asset infrastructure dimension that NextFi views as structurally co-present with AI transformation. Institutions reading only these three documents will underweight the settlement layer.

The concrete version of that claim is more useful than the general one. Goldman devotes an entire section to financing structures that are candidates for programmable settlement and tokenized collateral:

- Project debt with cash flows matched to contracted facility revenue
- Sale-leasebacks recycling capital out of built assets
- GPU-backed asset-based finance with embedded refresh provisions
- Credit wrapping that blends investment-grade anchors with non-investment-grade counterparties
- Robot-as-a-Service subscription economics

Each of these is an instrument whose administration — collateral tracking, waterfall execution, cash flow matching, counterparty settlement — is precisely the work programmable rails are built to compress.

Goldman identifies the demand. It does not identify the settlement infrastructure that will service it. That is the seam.

KEY INSIGHT

The central finding across all three source documents is not that AI is transforming financial institutions, it is that most financial institutions have not yet structured themselves to capture the transformation that AI makes possible.

McKinsey's readiness gap (70 percent personal against 27 percent organizational), BCG's unclaimed 300 to 500 basis points, and Goldman's \$7 trillion of capital not yet deployed each describe a different dimension of the same gap: the distance between deployment and value.

For NextFi clients, this gap is a program design problem, not a technology selection problem. **The institutions that close it fastest will do so by resolving governance, data architecture, and organizational sequencing before scaling, not after.**

Section IV: Implementation Framework for Institutional Leaders

BCG's seven steps describe *what* to do. McKinsey's horizon data describes *why it works*. Neither resolves sequencing under governance constraint, which is where institutional programs most often fail.

The framework below is NextFi's sequencing and governance overlay on that source material.

Phase 1: Diagnostic and Prioritization (Weeks 1–6)

- Place current AI activity honestly against McKinsey's horizons — enablement, automation, reinvention — recognizing that nearly 90 percent of organizations sit in the first two and that McKinsey's own sample skews toward the advanced end.
- Score organizational readiness separately from personal readiness, and assign an owner to the former. This is the variable that accounts for 48 percent of the value-capture difference, and in most institutions no one owns it.
- Select two or three workflows for end-to-end redesign, following BCG's instruction to concentrate rather than distribute. Distribution, investment research, and middle/back office operations carry the highest capacity-unlock estimates (up to 45, 50, and 60 percent respectively).
- Identify the data architecture and governance gaps that have prevented pilot results from compounding. These are almost always the rate-limiting factor.

Phase 1 exit criteria: named workflows with baseline economics; an organizational readiness score with a named owner; a documented data gap inventory.

Phase 2: Workflow Redesign and Governance Architecture (Weeks 6–16)

- Redesign target workflows around what AI makes possible, not around existing process maps. This is where McKinsey's 5.3x multiplier is earned, and it is available without first reaching a higher horizon.
- Build model risk governance in from day one. BCG's warning is unambiguous: retrofitting is exponentially harder. Both middle/back office automation and agentic client engagement require approval workflows, audit trails, and human-in-the-loop escalation paths defined before deployment, not after.
- Invest in leadership AI fluency as a governance control, not a training line item. McKinsey's 3.9x multiplier attaches to leadership fluency and its 3.3x multiplier to skill-building support.
- Address leadership trust and employee communication explicitly. Middle-manager anxiety runs at one in four; low trust raises anxiety by 1.5x. Treat it as a program risk with an owner, not as change management overhead.

Phase 2 exit criteria: two or three workflows redesigned with named P&L targets; a signed human-in-the-loop boundary map; an MRM protocol covering agentic systems.

Phase 3: Production Deployment and Capital Positioning (Months 4–12)

- Move redesigned workflows into production with defined performance metrics. Measure at the workflow level, not the tool level, and tie outcomes to scorecards and compensation — BCG's sixth step, and the one most often skipped.
- For banking and capital markets institutions, engage Goldman's framing directly. The index concentration constraint, the migration of AI financing into private credit, and the emergence of GPU and project-debt collateral structures are simultaneously a client service opportunity and a strategic planning input.
- Begin integrating digital asset settlement capabilities into AI workflow architecture. The convergence of programmable rails and AI-automated workflows is an active design consideration for institutions building durable operating models now.

Phase 3 exit criteria: workflow-level P&L attribution; AI outcomes present in performance reviews; a documented position on settlement infrastructure.

Horizon Map: Where Most Institutions Are vs. Where Value Lives

Horizon (McKinsey)	Description	Share of organizations	Leaders reporting enterprise value
Enablement	Individual adoption, tool deployment, pilots	Majority; ~90% sit in Horizons 1–2	13%
Automation	End-to-end automation of existing cross-functional workflows	Balance of the ~90%	24%
Reinvention	Blank-slate redesign of roles, workflows, operating models	11%	48%

Institutional Priority Matrix by Function

Function	BCG capacity unlock	NextFi priority	AI mechanism	Governance requirement
Operations (middle/back office)	Up to 60%	High	Reconciliation, breaks triage, NAV oversight, regulatory reporting	Exception handling, human escalation
Investment research	Up to 50%	High	Continuous scanning, signal discovery, summarization and modeling	Model risk framework, audit trail
Portfolio construction and risk	Up to 50%	High	Constraint checking, rebalance recommendations, scenario testing	PM/IC sign-off, independent validation
Marketing	Up to 55%	High	RFP and DDQ generation, personalization at scale	Content library approval, compliance review
Sales and client engagement	Up to 45%	High	Agentic CRM, prospecting, real-time signal scanning	Suitability review, communications supervision
Trading	Up to 40%	Medium	Pre-trade compliance checks, execution planning, surveillance	Best execution, audit trail
Enterprise data architecture	Enabler	Foundational	Unified knowledge graph, MCP connectivity, per-agent projections	Data governance, lineage, residency, access controls

Legal, compliance and risk	Up to 40%	Critical enabler	Continuous surveillance, AI risk monitoring	Independent validation, MRM policy, information barriers
AI and data governance	Up to 30%	Critical enabler	Agent evaluation harness, guardrails as a service	Senior committee, auditability by design

Capacity-unlock figures are BCG's three-to-five-year estimates for asset management and vary by asset class and operating model. Priority rankings and governance requirements are NextFi's.

Section V: Regulator-Ready Considerations

The governance dimension is consistent across all three sources, even where they approach it differently. For NextFi clients, regulator-ready means operationally integrated, not separately maintained.

Key considerations for institutions building toward the automation and reinvention horizons:

Model Risk Management (MRM) integration

- AI models deployed in investment decision support, client engagement, and operational workflows fall under existing model risk frameworks, and increasingly under AI-specific supervisory guidance. Institutions should not wait for final rule-making to establish validation standards. Build MRM protocols that accommodate both current SR 11-7 expectations and forward-looking AI governance requirements.
- For institutions operating in Europe or serving European clients, the EU AI Act's risk-tier obligations should be scoped into the same protocol rather than tracked separately — confirm current applicability and timing with counsel.

Audit trail and explainability

- Agentic systems that act on behalf of clients — generating proposals, executing outreach, flagging compliance exceptions — require documented decision logic. This is a design input, not a technology afterthought.
- BCG's architecture makes the same point structurally, specifying audit-grade transparency in agentic workflows and an immutable audit trail at the data layer. Systems that cannot produce explainable outputs at the workflow level will fail regulatory review and internal audit scrutiny regardless of their performance metrics.

Human-in-the-loop escalation architecture

- BCG names human-in-the-loop boundaries explicitly as part of day-one governance; McKinsey's workflow redesign recommendations require the same logic implicitly. Regulators across jurisdictions are converging on the expectation that consequential AI-driven decisions in financial services retain a defined human review point.
- That architecture must be documented, tested, and governed — not assumed. Note that BCG's cost-reduction estimates are themselves predicated on human sign-off remaining for regulated and client-facing output.

Data residency and model lineage

- BCG specifies multi-jurisdiction data residency, information barriers between mandates, immutable audit trail, and a unified catalog carrying lineage, quality scoring, and freshness as infrastructure requirements — not as compliance add-ons. Enterprise data architecture is simultaneously a performance prerequisite and a regulatory one, and the institutions that treat it as an IT project rather than as infrastructure will pay for it twice.

Third-party and vendor risk

- Most AI deployment in financial institutions runs through third-party models, cloud infrastructure, and fintech vendors, and BCG's own guidance is that a hybrid buy-and-build approach is inevitable. Regulator-ready governance requires that third-party AI risk be assessed, documented, and reviewed with the rigor applied to internal model development.
- Goldman's index-concentration analysis has an operational analogue here: vendor concentration in AI infrastructure is accumulating faster than most third-party risk frameworks were designed to measure. For EU-regulated institutions, DORA's ICT third-party and concentration risk provisions are the natural home for this work.

Board diagnostic

Five questions a board should be able to answer, adapted from McKinsey's framing:

1. Where will AI create value for us specifically, and what is that worth in basis points?
2. How must work change to capture it, and which two or three workflows are we redesigning first?
3. What skills do we need, and who is accountable for building them?
4. How do we govern a system of humans and agents, and where are the human review points?
5. What is our organizational readiness score, who owns it, and when is it next measured?

One dimension the three source documents do not address

The settlement layer is also moving. Goldman describes an emerging capital architecture of project-matched debt, sale-leasebacks, GPU-backed asset finance, credit wrapping, and subscription-based physical AI procurement, and identifies the demand for all of it without identifying the infrastructure that will administer it. Institutions designing AI operating models today without accounting for programmable settlement are, in our assessment, building toward a second redesign within 24 to 36 months.

The durable advantage belongs to institutions that architect for both compressions simultaneously — cognitive labor and transaction coordination — treating them not as parallel workstreams but as a single transformation surface.

STRATEGIC BOTTOM LINE

The gap between pilot-stage AI investment and enterprise-level value capture is real, measurable, and now documented across McKinsey, BCG, and Goldman Sachs simultaneously. That convergence signals not incremental urgency but a structural inflection.

The institutions that emerge as category leaders in AI-enabled financial services will not be distinguished by which tools they deployed first. They will be distinguished by how completely they restructured the organizations, workflows, and governance architectures around those tools before their competitors did.

The evidence is now specific enough to act on: organizational readiness is nearly twice as decisive as individual readiness, redesigned workflows carry a 5.3x value-capture multiplier, and the

available prize in asset management is 300 to 500 basis points of margin.

The firms winning this transition are not running more pilots. They are running fewer, better-designed production systems, with governance built in from the start, data architecture treated as infrastructure rather than as an IT project, and workforce transition managed as a program risk rather than as a change management afterthought.

About NextFi Advisors

NextFi Advisors is an independent advisory and consulting firm helping financial institutions move from experimentation to execution across AI and digital asset transformation. Our work is commercially viable, regulator-ready, and operationally durable.

Contact: barry.eisenberg@nextfiadvisors.com | **Web:** www.nextfiadvisors.com

Sources: (1) *From Adoption to Impact: Three Horizons of AI Transformation*, McKinsey & Company, People & Organizational Performance Practice and QuantumBlack, AI by McKinsey, July 2026. Authors: Aaron De Smet, Drew Goldstein, Holly Price, Tanguy Catlin, Laura Pineault, Nicolette Rainone, and Susann Almasi. (2) *AI-First Companies Win the Future: Executive Perspectives on Asset Management*, Boston Consulting Group, July 2026. (3) *Harnessing AI for the Real Economy*, Goldman Sachs Global Banking and Markets, 2026; opening letter by Dan Dees, Co-Head of Global Banking and Markets. Additional context drawn from NextFi Advisors published intelligence, including *The Convergence Economy: Nine Strategic Imperatives*, *The Convergence Economy: Repricing of Intelligence and Transactions*, *Multi-Agent Design Models for Financial Institutions*, *Three Trends, One Infrastructure Stack*, and *Context Engineering: How Enterprise Operating Models Impact AI Performance & Output*, available at intelligence.nextfiadvisors.com.

Disclaimer: This Intelligence Brief is produced by NextFi Advisors for informational and strategic planning purposes only. It does not constitute investment advice, legal advice, regulatory guidance, or an offer to buy or sell any security or financial instrument. The views expressed represent the analytical perspective of NextFi Advisors and do not necessarily reflect the positions of any source institution cited. Source documents are summarized and interpreted; readers should consult original sources directly. Regulatory references are indicative and should be confirmed with counsel. NextFi Advisors makes no representation as to the completeness or continued accuracy of third-party data referenced herein.

Sources and Citations

Numbered references correspond to the superscript markers in the body text.

¹ McKinsey surveyed 750 employees and leaders across North America, Europe, Asia-Pacific, Latin America, and Eastern Europe, the Middle East and Africa, with data collected February to April 2026. Organizational readiness and enterprise value questions were answered by a subset of 608 executives, senior leaders, and middle managers; 717 respondents were reliably classified into a single horizon. McKinsey notes that recruitment targeted respondents from organizations in more advanced horizons and that estimates may not reflect overall market prevalence. Source: *From Adoption to Impact: Three Horizons of AI Transformation*, McKinsey & Company, July 2026.

² Source: *AI-First Companies Win the Future: Executive Perspectives on Asset Management*, Boston Consulting Group, July 2026. Value estimates are illustrative three-to-five-year firm-level figures and assume human sign-off remains for regulated and client-facing output.

³ Goldman Sachs projects hyperscalers alone will invest more than \$6 trillion in AI through 2030, with total AI infrastructure investment estimated at \$7 trillion or more between 2026 and 2031 across compute, power, and data centers; baseline aggregate capex estimates total approximately \$7.6 trillion over that period. Source: *Harnessing AI for the Real Economy*, Goldman Sachs, 2026.

⁴ Global hyperscaler capex projected above \$760 billion in 2026; global data center supply grew from 30 GW (2019) to 57 GW (2024), with 65 GW further projected online by 2030. Source: *Harnessing AI for the Real Economy*, Goldman Sachs, 2026.

⁵ Goldman Sachs estimates US AI investment at approximately 1.2% of today's annual GDP, noting that even optimistic scenarios remain well below the 3% to 4.5% of GDP represented by the railroad buildouts of the 1800s. Source: *Harnessing AI for the Real Economy*, Goldman Sachs, 2026.

⁶ SaaS accounts for less than 0.5% of global GDP; the technology sector overall is approximately 5% of global GDP against global GDP above \$100 trillion. Source: *Harnessing AI for the Real Economy*, Goldman Sachs, 2026.

⁷ Helsing, founded 2021, reached a €12 billion valuation in a €600 million Series D in June 2025. Source: *Harnessing AI for the Real Economy*, Goldman Sachs, 2026, citing Helsing press release, 17 June 2025.

⁸ PhysicsX raised a \$300 million Series C in June 2026 at a roughly \$2.4 billion valuation. Source: *Harnessing AI for the Real Economy*, Goldman Sachs, 2026, citing PhysicsX press release, 8 June 2026.