

MARKET IMPACT REPORT

From AI to outcomes: closing the value gap in non-bank lending

Based on in-depth executive interviews across
the non-bank lending industry

Authors:

Divya Iyer, Practice Leader

Don Ryan, Research Fellow & Senior Advisor

Contents

	Page
Executive summary	03
The economics of people-heavy lending no longer add up; today's mortgage workforce is being built around AI	04
Cost-per-loan is the North Star, but AI's value expands beyond costs	07
Measure your AI outcomes in lending so you can manage your AI better	09
Your mortgage value chain bets lie in the AI investment matrix	10
The self-funded model is real, but only if you can see the savings	13
Closing the infrastructure gap can help scale AI	14
AI can't be scaled unless you're ready to govern it	16
Deciding whether to build, buy, or partner shapes how AI is delivered	17
Start here: A practical sequence for deploying agentic AI at scale	19
The Bottom Line	20
About Cognizant	22
About HFS	23

Executive summary

Artificial intelligence (AI) has stopped being a technology problem. The tools exist, the capability is real, and the deploying costs continue to fall. What enterprises still struggle with is showing that up in their numbers. That gap between what AI can do and why only few organizations see the impact reflected in their margins, cost structures, and productivity is where the real story lies. For CIOs, CTOs, and business unit heads in non-bank lending, an industry running on thin margins, labor-intensive processes, and cost structures built for a different rate cycle, closing that gap is an operating model problem. The institutions that solved this are showing results in their numbers. Those that haven't are still explaining why their investments haven't moved the needle.

Non-bank lenders built their businesses on a simple premise: move faster than banks, run leaner, and operate without the capital reserve requirements that constrain depository institutions. This just meant escaping the balance sheet burden, not regulatory oversight. That model is now under pressure. Since 2022, elevated rates, compressed volumes, and rising origination costs have exposed cost structures built for a different cycle. The challenge is no longer cyclical, but structural.

AI, when deployed correctly, is the most powerful lever available to address that, but its adoption across most lending institutions is fragmented. Lenders are not suffering from a lack of AI activity, but from AI investments that are scattered, ungoverned, and disconnected from measurable outcomes. In many cases, tools are being adopted ahead of a formal architecture, platforms are scaling ahead of governance, and reported savings often lack a clear baseline to validate them. Muted lending volumes and constrained capital have raised the bar for every investment decision. The combination is a difficult one: the urgency to act is high, the cost of getting it wrong is higher, and most lenders are yet to find a reliable path from AI deployment to profit and loss (P&L) impact.

The value of AI in non-bank lending is highly concentrated in institutions that have already done

the foundational work. This includes clean data, decomposed workflows, governance frameworks, and an operating model redesigned around AI rather than layered on top of existing ones. In those institutions, the economics are real and visible in P&L. For those who haven't been doing the foundational work, AI looks present but scattered, activity is high, and bottom-line impact is unclear. For most non-bank lenders, the issue is not that AI has failed them. They just haven't built the conditions for it to succeed.

HFS Research, in partnership with Cognizant, held in-depth discussions with senior executives from some of the largest originators and mid-sized regional players in the non-bank lending industry in Q1 2026. One theme ran through nearly every conversation: lenders can see the potential of AI but are unable to draw a direct line between deployment and a measurable improvement in their bottom line. That gap is holding most institutions back far more than funding, talent, and technology access.

For lenders navigating AI deployment under capital constraints, this report offers a self-funded AI approach to help them readily identify high-confidence, measurable AI projects, execute them to capture the savings they generate, and reinvest those savings into subsequent waves of deployment. Done well, this creates a flywheel: each cycle lowers the upfront capital burden while building a visible link between AI spend and business outcomes at each stage. Making it work requires a live measurement baseline established before a project begins and a full accounting of the true cost of AI. That means not just measuring headcount savings, but also the costs of compute, model hosting, governance, and ongoing human oversight.

This report also examines the forces, frictions, and strategic choices that will determine who comes out ahead. It is targeted at lending executives who know AI matters but haven't yet decided what to do about it or are not seeing value from their AI efforts.

The economics of people-heavy lending no longer add up; today's mortgage workforce is being built around AI

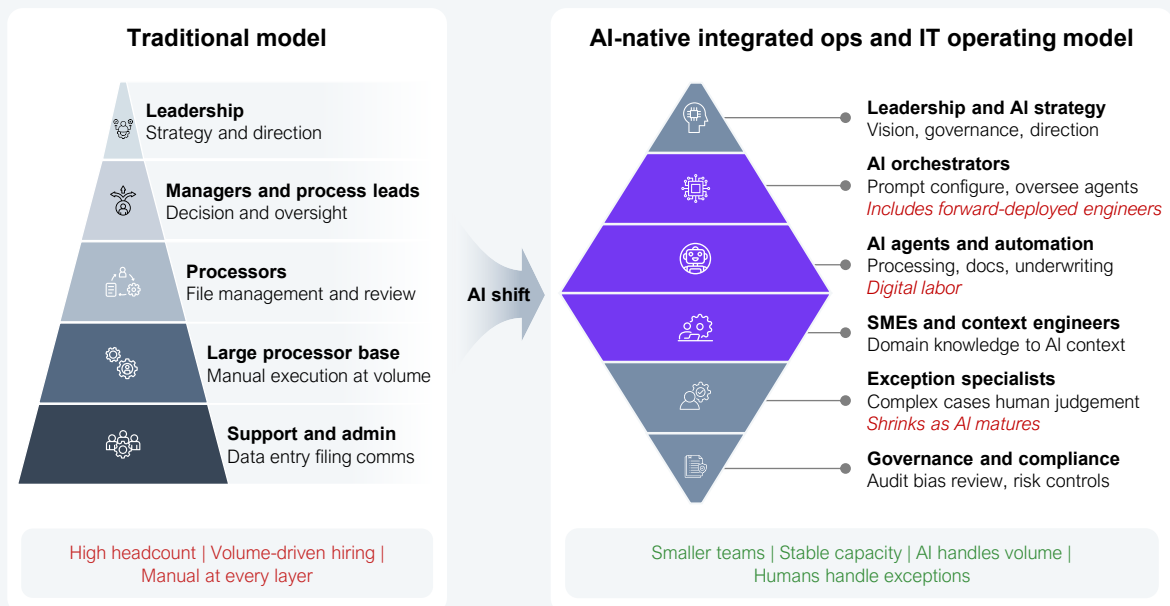
HFS Research has been tracking a structural shift accelerating across knowledge-intensive industries. The traditional workforce pyramid, organized by a hierarchy of labor with large volumes of entry-level workers at the base that's supervised by progressively thinner layers above, is collapsing. In its place, a diamond is forming: a smaller base of AI-fluent practitioners, a broad middle of orchestrators who design and manage human-AI workflows, and a focused apex of architects and governors who align intelligent delivery with business outcomes (see Exhibit 1). An increasingly important part of the middle layer is the forward-

deployed engineers (FDEs). These are practitioners who wire AI models into live systems, govern them in production, and ensure they continue performing once the technology vendors have moved on. Orchestrators can prompt, configure, and oversee agents only when they are embedded in real workflows with real data, permissions, and audit trails. FDEs make that embedding durable by building the pipelines, permissions, and monitoring to ensure agents keep working after go-live. That durability places them squarely in the diamond's broadest tier alongside orchestration.

Exhibit 1: AI is absorbing the wide base of the pyramid

The shift from volume-driven human execution to AI-orchestrated, exception-focused teams

- **Digital labor** — AI agents operating autonomously; no human in the loop for routine tasks
- **Human labor** — All other layers; roles shift from execution to oversight and context-setting



Note: Exception specialists layer will gradually shrink as AI/agent maturity increases
Source: HFS Research, 2026

Non-bank lending is a vivid early case study underscoring this shift. It has long been one of the most people-intensive corners of financial services, comprising loan officers, processors, underwriters, appraisal reviewers, and investigators. Each role performs specialized but largely manual tasks, adding time and cost to a transaction that customers increasingly want resolved in hours, not weeks. Most firms we interviewed had cost structures that reflected exactly this labor-intensive, process-driven structure.

To illustrate the underwriting cost alone, an executive described a firm with 20 underwriters earning approximately \$200,000 each, representing a \$4 million annual line item for a single function. At another institution where 300 processors support 2,000 loan officers, using AI to reduce that ratio is expected to yield \$700,000 in savings per processing pod, translating into tens of millions of dollars in potential savings across the layer. Add roughly 50 appraisal review analysts performing work described as almost entirely automatable, and the numbers compound quickly. Taken together with processors, appraisal reviewers, compliance staff, and post-closing roles, another executive estimated that roughly 700 people were performing important but repetitive tasks at an average salary of around \$60,000. The math is not subtle.

Beyond steady-state headcount, there is a structural cost that rarely features in the analysis: the surge-and-collapse hiring cycle. Every rate drop triggers a wave of hiring, and every contraction reverses it. AI does not just reduce steady-state headcount, it removes the cost of that cycle entirely.

This is not an argument for eliminating humans or pretending AI comes without cost. Building and running agentic systems, from simple automations

to complex, context-aware retrieval-augmented generation (RAG) agents and multi-agent architectures, requires real and ongoing investment in compute, storage, model refinement, governance, and human oversight. What changes is the shape of that cost, not its disappearance. A capable agent can be built once and deployed at scale, absorbing the marginal cost of additional volume without the headcount increase. That is the structural shift: breaking the link between volume and workforce size that has defined lending economics for decades.

But lenders should be clear about where the economics stands today. Per-unit AI costs, including compute, model hosting, and inference, are falling and will keep falling. That trend is widely cited but tells only one part of the story. Total cost does not necessarily fall just because unit cost does. As lenders take on more complex, multi-step agentic work, the volume of compute unit consumed per task increases, and the people needed to govern, monitor, and refine that work do not go away. Falling unit price and rising task complexity are moving at the same time but in opposite directions, making the net effect on a lender's actual AI bill unclear. That is precisely why the foundation matters more than the price curve.

Clean, well-governed data must be the bedrock, along with executable governance. For example, personally identifiable information (PII) should be blocked from a model automatically rather than flagged by policy. Building the infrastructure to scale safely, together with a workforce and culture fluent enough to act on the output, becomes the foundation that takes years to build and can't be bought once the economics look obvious. Lenders building that now are not betting that AI will be cheap. They are making sure they are ready before anyone can prove it will be.

A transformation of this scale is not only a technology investment. It brings a change management challenge that most lenders are not equipped to carry. Retraining teams, redesigning workflows, and reconstitution governance structures for AI require organizational muscle that takes time to build. A partner like Cognizant, which has made significant investments in an AI-enabled workforce purpose-built for this kind of transition, can absorb that burden and work alongside lenders as they build the internal capability to sustain it.

For institutions with a global offshore workforce, the model is different, but the pressure is the same. A CIO at a US mortgage servicing company, noted that in lower-cost geographies, the cost of running an AI agent can today exceed the cost of the human it replaces. But that gap is closing. While AI infrastructure costs are falling, offshore labor costs have moved in the opposite direction as wage expectations rise and the talent competition intensifies. The crossover point is a matter of when, not if. Based on our discussion with a CIO of a US mortgage servicing company, the timeline is estimated at three to five years. Firms with a global delivery footprint that aren't using this window to prepare risk continuing to lean on manual workarounds when it closes, exactly when competitors have already automated theirs.

"The offshore model puts itself against the wave of modernization. You're going back to manual workarounds as opposed to being nimbler. Their cost of actually doing that will be increased from what they are used to."

— CIO at a US mortgage servicing company

The challenge is not ignorance of this reality, it's inaction. Across our interviews, several senior executives demonstrated a clear understanding of the direction of travel but are yet to pull the trigger. Budget constraints, unclear ROI, a saturated vendor landscape, legacy infrastructure, and the distraction of navigating a difficult rate environment have all contributed to hesitation.

"We know which way we want to go. Taking the first step has been the issue."

— Vice President at a top three non-bank mortgage lender

Cost-per-loan is the North Star, but AI's value expands beyond costs

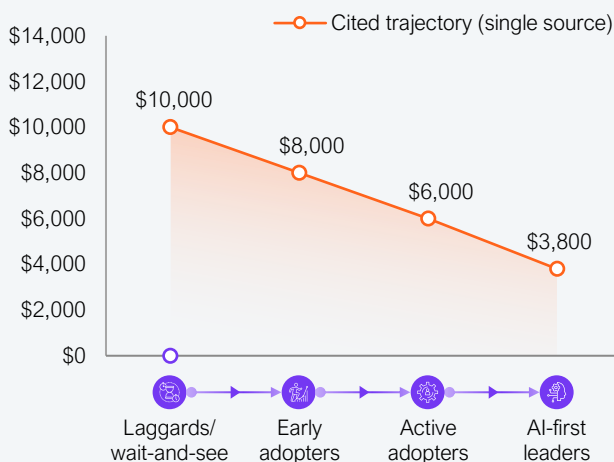
Across the 11 institutions we spoke with, cost-per-loan emerged as the “North Star,” a central measure of AI's impact. The industry average stands at \$11,094 per loan, according to the MBA's 2025 Annual Mortgage Bankers Performance Report. The leaders are already significantly below that. An executive told us about a COO who drove down costs from \$10,000 to \$3,800 and is targeting \$1,500 through agentic AI (see Exhibit 2). He built an internal agentic platform in two years, absorbing \$200,000 in annual reject costs while generating roughly \$10 million in overall savings. This proves what's possible when AI is treated as an operating model redesign.

But the path to those numbers is not uniform across the lending operation. And the interviews revealed clearly where AI is mature today, where it is developing, and where it remains aspirational.

The value of AI lands differently across origination and servicing, each with distinct maturity curves, different transformation timelines, and different economic levers. On origination, the maturity is advancing now. Document intelligence and condition clearing are in production. Conforming loan underwriting automation is the next step. The prerequisite is workflow rework, decomposing the process end-to-end before agentic can take hold. Institutions that have done that work are already seeing results. It's this prework, including process redesign, and data uplift, that determines readiness. For these institutions, the consensus across interviews points to end-to-end origination automation landing within 18–24 months.

Exhibit 2: Cost-per-loan compression is reshaping lending economics

Cost-per-loan compression



The source described this compression as spread across origination sub-functions (sales, processing, underwriting), not driven by one area alone.

Drivers cited in this account

Agentic condition-clearing in underwriting
\$100/agent vs. \$200K/underwriter per file

Direct underwriter headcount replacement
20 underwriters (\$4M) → AI vendor (~\$1M)

AI voice agents for origination/lead intake
Replaces loan officer assistants, processors

Blended AI + offshore “skeleton crew” model
AI handles volume; offshore covers the rest

In-house agentic rebuild replacing legacy LOS
Displaces vendor contracts over time

Note: This curve and all drivers above trace to a single practitioner account, describing underwriting and origination labor cost specifically, not a researched, blended cost-to-originate figure.

Source: HFS research in partnership with Cognizant, 2026

Servicing is on a different timeline. Call center automation is already the most mature AI use case delivering measurable returns. Beyond that, the picture gets more varied. Several executives noted that complex servicing interactions, where resolution requires judgment, negotiation, or regulatory care, remain difficult to hand to agentic AI. The servicing transformation takes longer than origination, and the re-work needed to get there is no less significant.

Executives did highlight that AI is creating value well beyond cost reduction. This includes processing times cut by 25% or more, underwriters moving from a hundred manual tasks a day to exception-only oversight, customer retention climbing above 80%, regulatory rule-lookup accuracy reaching 98% against a roughly 80% human baseline, and lead conversion improving 12% to 15% through better timed outreach. That 98% figure though is worth pausing on: institutions are keeping it under heavy human oversight before letting it run unsupervised, even though it already beats the

human benchmark. One pattern worth noting is that on the underwriting side, AI clears most conditions on a file, but the remaining items are deliberately left for a person to make the final call. Strong accuracy numbers have not translated into full autonomy. Institutions are choosing to keep a person in the loop and keep the cost of being wrong contained, even as they trust the AI to do most of the work.

Moreover, the value is not evenly distributed. Copilot tools, proposal generators, and call summarization deliver real productivity gains but sit at the surface of the operation. They make people work faster without changing the underlying lending economics: what it costs to originate, process, or close a loan. The largest gains belong to institutions that have done the foundational work first: decomposing workflows, cleaning data, and building the governance frameworks that let agents operate reliably at scale (see Exhibit 3). That's what separates the two columns.

Exhibit 3: AI creates measurable value across the lending value chain such as cost-to-originate declining by roughly 62% and customer retention over 80%

Without AI	Outcome	Directional possibility with AI embedded
\$10,000–\$12,000 industry average 	Cost per loan	 \$10K → \$3,800 today, trending toward \$1,500
30–45 day origination cycle 	Lender-side processing time	 Internal decisioning toward single-digit days; total time-to-close still gated by appraisal, verification, and title
Near 100% manual 	Human touch in document processing	 Meaningfully reduced; roughly 95% less manual interaction in parts of the workflow
Surge hiring with every rate drop 	Underwriter capacity	 Shift toward a stable, exception-focused team
Industry average reference point 	Customer retention	 Over 80% at one AI-powered servicer
Approximately 80% human accuracy in identifying current regulatory requirements 	Regulatory rule-lookup accuracy	 98% in internal testing; still under human review before full deployment
Undifferentiated outreach 	Lead conversion/marketing ROI	 12%–15% ROI improvement from AI-driven best-action models
Hire-and-fire with every rate cycle 	Workforce scalability	 Greater stability as agentic tools absorb volume swings

Sample size: Based on 11 US senior non-bank lender executives

Note: Figures are indicative, drawn from individual executive accounts in HFS Research interviews, not definitive or industry-wide benchmarks.

Source: HFS research in partnership with Cognizant

Measure your AI outcomes in lending so you can manage your AI better

None of the outcomes discussed earlier matter if you can't measure AI and agentic value. Several executives who we talked to in this research had significant AI investments underway but no reliable way to know what they were delivering. Before any project begins, lenders need live baselines such as cost-per-loan, time-to-close, processing accuracy, retention rates, and the infrastructure to track them consistently. This measurement architecture is more than just a reporting exercise; it's what makes the investment legible. Without it, the savings might be real but not measurable, the next investment cannot be justified, and the gap with the leaders keeps widening.

Additionally, what gets complicated is not measurement but attribution. When a loan officer's productivity increases and more volumes get processed, isolating how much of that revenue growth is AI or agentic driven versus market conditions, rate movements, or sales execution is difficult. A CIO at a US mortgage servicing company observed that there are simply too many moving parts to cleanly isolate AI's contribution.

"How much it contributes to the revenue goal is very difficult to ascertain. There are a bunch of variables and not enough equations to solve for all of them."

— A CIO at a US mortgage servicing company

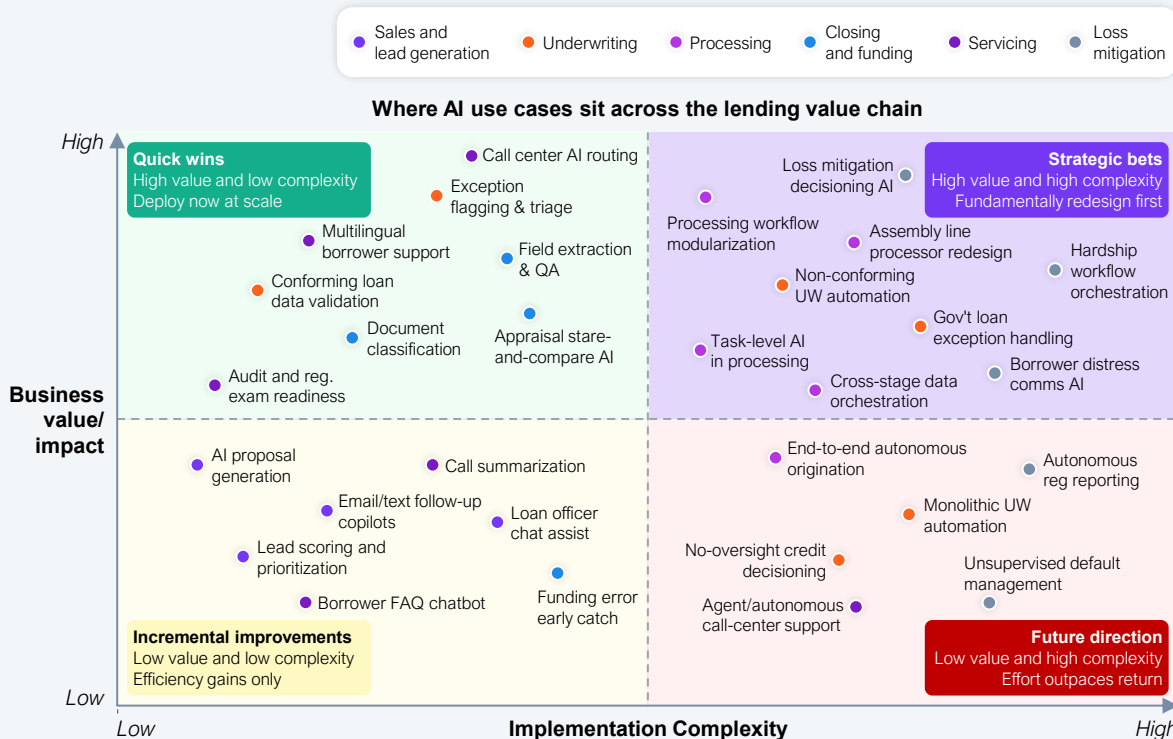
Your mortgage value chain bets lie in the AI investment matrix

Most lenders are investing where AI is easiest to deploy, not where it fundamentally changes lending economics. We mapped every meaningful AI use case discussed across our interviews against business value and the complexity of implementation in Exhibit 4. The position of each use case is a decision framework: the X axis tells you how much foundational work is required before AI can deliver, and the Y axis tells you what you get in return. That trade-off, i.e., where to place your bets, how much effort the foundational work requires, and what the risk and reward looks like, is what separates the institutions moving deliberately from those still spreading investment thin across the chart.

The pattern was consistent across four quadrants:

- **Incremental improvements:** Lenders starting with copilots, chatbots, and proposal generators often get stuck here, mistaking productivity gains for operating model transformation.
- **Quick wins:** Lenders here are scaling in call center AI, document intelligence, and conforming loan underwriting, high-value use cases that deliver real returns without requiring deep organizational transformation.

Exhibit 4: Scaling AI requires choosing where to accelerate now and where to fundamentally rethink the lending life cycle



Sample size: Based on 11 US senior non-bank lender executives
Source: HFS research in partnership with Cognizant

- **Strategic bets:** Processing modularization, loss mitigation, and non-conforming underwriting address the structural economics of lending, and most institutions remain underinvested here.
- **Future direction:** Lenders seduced by this category are letting ambition get ahead of reality. End-to-end origination automation, unsupervised default management, autonomous regulatory reporting, and agentic hardship call handling represent the direction of travel, not capabilities that are deployable today.

Each quadrant carries a different imperative and a different consequence for institutions:

1. **Incremental improvements are useful but not transformative:** Sales tools, copilots, call summarization, and FAQ chatbots should be deployed as a baseline. But they do not change the economics of lending like what it costs to process, underwrite, or close a loan. The risk is not deploying them; it is mistaking the productivity gains for operating model transformation while deferring the harder work that would significantly improve the economics of lending.

“For the sales guys, we’ve implemented an AI feature where they can put together proposals, type in a couple of simple parameters, and come back with an entire proposal. But nothing beyond that has been implemented that I know of, and I’m pretty much in all these meetings.”

— Vice President at a top three non-bank mortgage lender

2. **Deploying quick wins immediately:** Call center AI routing, multilingual borrower support, conforming loan data validation, document classification, and field extraction are the use cases here. Workflows are structured, data is clean, the human-in-the-loop model is understood, and ROI is already measurable in early deployments. There is no hard organizational problem to solve first. Each quarter that lenders delay is another quarter of cost savings and cycle time foregone.

“If we can take out an hour of human time looking at an appraisal report, it’s time the loan officer can do more loans, but the customer also gets their stuff back faster.”

— Security Architect at a mid-market mortgage servicer

3. **Strategic bets are transformative:** Processing modularization, loss mitigation decisioning, non-conforming and government loan underwriting are part of the most consequential and most underserved territory and where the future winners of non-bank lending will be decided. The common feature across every use case here is that deploying AI effectively requires doing something hard first: modularizing workflows, redesigning roles, cleaning and structuring data, and building governance frameworks. The technology is not the limiting factor. The organizational preparedness to do that foundational work before asking AI to perform is.

4. **Future direction, not yet within reach:**

Autonomous underwriting, unsupervised default management, autonomous regulatory reporting are the direction of travel and not deployable today. End-to-end origination automation takes 18–24 months for those that have done the workflow decomposition work. Fuller servicing transformation is a two-to-three year journey, contingent on process maturity and AI costs falling relative to offshore labor. There is also a human constraint: in hardship conversations and distress calls, borrowers want to speak to a person. Deploying autonomous AI in these moments without a clear escalation model is a customer relationship risk, not just a governance and compliance one. The lenders best positioned to this quadrant are building the foundations now, not rushing toward it prematurely.

The lenders pulling ahead are not those with the most vendors or the most pilots. They are the ones that chose, with deliberate discipline, which quadrant to prioritize and then did the architectural work that quadrant requires before asking AI to do the work.

Knowing which use cases to prioritize is just the first step. Equally important is building the framework to measure whether those use cases are delivering the outcomes AI was meant to achieve. Without that measurement in place, it's easy to work backward, picking up the hammer first and searching for a nail to justify it rather than confirming that AI solved the problem it was deployed against.

The self-funded model is real, but only if you can see the savings

While the gap between investment and attributable outcome is what holds most institutions back, it's also where the self-funded transformation model becomes possible. Use the savings generated by initial AI projects to fund subsequent waves of investment. This creates a flywheel that reduces the upfront capital burden and aligns incentives between the lender and whoever is helping them build. Executives described it as logical and well-suited to how private equity-backed mid-market lenders think about capital allocation. Without baselines in place before a project begins, it's impossible to attribute savings with enough confidence to ring-fence, achieve the set-out outcomes, and reinvest them. But even with baselines, measuring the ROI of AI-enabled workflows is more layered than a straight headcount substitution. Human labor costs are replaced not by a single line item but by a distributed set of technical costs across tokens, compute, network, and storage, each of which scales differently, sits in different budget lines, and requires a different kind of financial literacy to track and manage.

Closing that gap starts with the basics: establishing live baselines before any project begins, building the measurement architecture that makes outcomes visible across the new cost structure, and connecting deployment decisions to business results with enough precision to justify the next wave of investment. For lenders who can baseline costs, investments, and outcomes, the self-funded model becomes a practical path rather than a theoretical one. An AI builder partner like Cognizant, with the domain expertise, lending-specific regulatory knowledge, and implementation experience to navigate that journey, can help lenders build the measurement foundation required.

"There needs to be a direct line between adopting agentic AI and clear, measurable, quantifiable improvements to the bottom line or to the top line. That direct line is what most lenders are missing, and without it, even the best AI investments stay trapped at the edges of the operation."

— Ajai Nair, Head of Lending Portfolio, Cognizant

Closing the infrastructure gap can help scale AI

Like every industry that has gone through an AI-led transformation, mortgage lending requires the same prerequisite: the infrastructure to support it. That means the underlying data architecture, the governance discipline that make AI deployable at scale, and the security controls that protect what is being built. Without these in place, AI investment

produces point solutions that do not connect to each other, cannot be measured, and deliver far less than their potential (see Exhibit 5). Every disconnected AI deployment increases enterprise complexity unless the underlying architecture is unified.

Exhibit 5: Data, governance, and security gaps determine AI's ceiling in lending

Where infrastructure is mature

- Unified data foundation with standardized APIs
- Data treated as products, not just feeds
- Real-time ingestion across the customer lifecycle
- Customer 360, Mortgage 360, Transaction 360 products
- ML model automation with continuous drift monitoring
- End-to-end audit logging for AI decisions
- AI risk boards with defined KPIs
- Contractual data sovereignty controls across all vendors

Where infrastructure is still catching up

- Siloed data by division and LOS platform
- Manual data quality management
- No real-time data pipelines
- Inconsistent document formats with no master schema
- Limited API access to core systems
- No governance framework for AI data use
- AI governance in infancy, risk boards not yet operational
- Security controls running behind AI adoption pace

Infrastructure gap

Sample size: Based on 11 US senior non-bank lender executives
Source: HFS research in partnership with Cognizant

A leader described the need for data foundation in plain terms: there is no AI without data. Their architecture spans more than 30 petabytes with real-time ingestion, producing standardized data products, like Customer 360, Mortgage 360, and Transaction 360, available via consistent APIs across the business. Another took a different but equally deliberate path: consolidating fragmented technology teams across multiple divisions into a single organization and then investing close to a billion dollars in a new internal operating system as the platform for everything that followed. The investment in this case was not made toward AI, but toward the infrastructure that made AI deployable.

Unfortunately, most lenders are nowhere close to any of this. Data sits in silos by division, locked in legacy loan origination systems that do not expose APIs, with no consistent schema across document types. A CISO described the challenge of getting real-time logs out of AI applications: shiny and pretty on the front, but very ugly in the back. Another described turning on enterprise co-pilot only to find it indexing everything and surfacing document titles to users who should never have seen them. The reality is that data and infrastructure gap compounds with every AI investment made on top of an unstable data structure.

“It’s not just about bolting something on. You really must understand how the system works. You have to set up your system to be more standardized and to have that governance.”

— Head of Engineering at a non-bank
lending platform

AI can't be scaled unless you're ready to govern it

In almost every interview, the same story emerged: business leadership is pushing AI adoption aggressively, while security, governance, and compliance teams are trying to build guardrails around a moving target. It is a structural problem created by the pace of adoption.

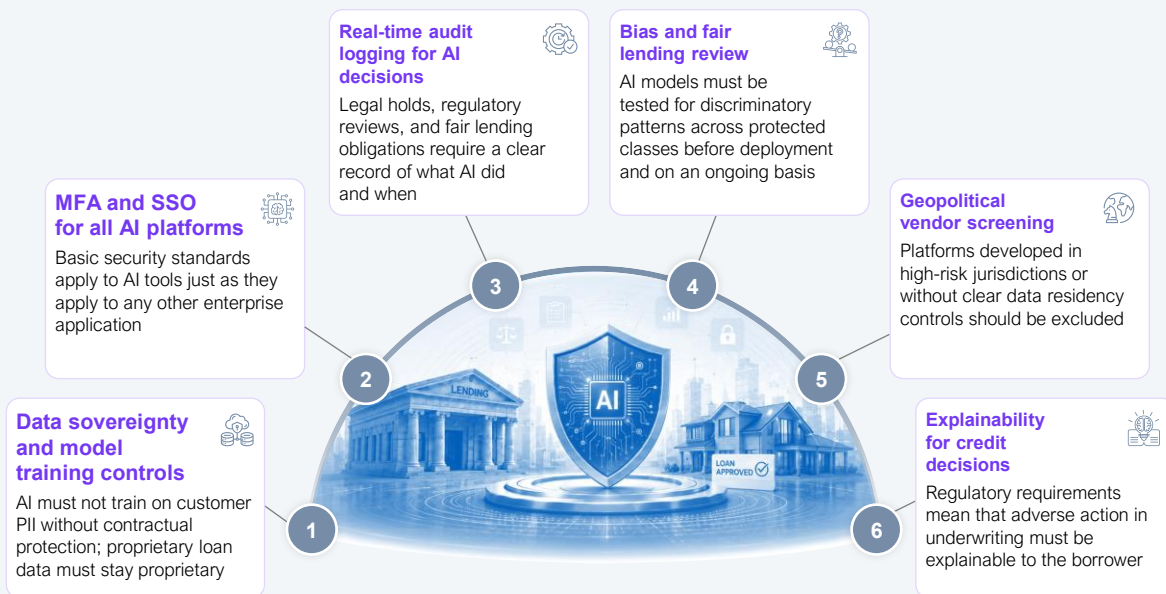
Many organizations described AI governance as being in its infancy. AI risk boards are being stood up and governance hires are being made. But formal metrics, defined KPIs for AI performance, and systematic oversight of what AI systems are doing in production are largely absent. A CISO put the risk plainly: at some point, an AI will make a mistake that leads to a legal hold.

"The business ask at the rate of adoption is higher than what security and developers can currently service. They want AI everywhere. Right now, we're running behind it."

— Security Architect at a mid-market mortgage lender

Several non-negotiables emerged consistently across interviews with firms that are scaling AI fast while embedding governance early enough to move safely at speed (see Exhibit 6). These are the baseline controls they have in place before scaling AI, regardless of where they are on the maturity curve.

Exhibit 6: Governance is not a constraint on AI, but a prerequisite to make AI deployable



Sample size: Based on 11 US senior non-bank lender executives
Source: HFS research in partnership with Cognizant

Deciding whether to build, buy, or partner shapes how AI is delivered

No question generated more divergent responses in this research than this one: who builds the technology that delivers AI? The answers fall across three distinct positions: build it yourself, buy it from a platform or vendor, or bring in a partner who has done it before. Where a lender lands depends on their scale, their technical capability, willingness to invest, and their appetite for risk.

Build the platform, own the advantage

The self-build thesis holds that the cost of building AI platforms has fallen so substantially that the traditional case for outsourcing is weakening. For lenders with the right talent and governance, self-build is becoming a more credible path than it was even a few years ago, and they understand their own processes better than any third party. The most cited example is a major servicer rebuilding its platform in-house using agentic AI, positioning itself to eventually offer that platform to other lenders and displace the traditional integrators entirely.

Buy from the people building mortgage-specific tools

The buy thesis comes into play when not every lender has the technical talent, capital, or appetite to build. Several executives expressed a clear preference for consuming AI through platforms and third parties. As one put it: this is a lending business, not a technology company, go to the

people building the best tools. But the buy decision is not simply about who has the best tooling and platform, it is about who has built it for mortgage context. Training AI agents on lending-specific workflows, regulatory requirements, and loan data is what separates a capable general platform from one that can operate reliably inside a lending institution.

For some, the buy decision is even more passive, waiting for AI to arrive embedded in core enterprise application platforms they already run. The catch is that these platforms come with a significant cost premium driven by the underlying investment in models, tokens, and supporting infrastructure the provider has already made. For budget-constrained lenders in a difficult rate environment, that premium can make the economics harder to justify than building or partnering directly.

“Our job was to close loans, not to be a technology company. Go to the people building the best tools and bring them in house.”

— Former Chief Operating Officer at a top five non-bank lender

Partner with domain expertise with skin in the game

The partner model is where domain knowledge and implementation experience create value that neither a platform vendor nor an internal team can easily replicate. Decomposing mortgage workflows, connecting siloed data, building governance frameworks, measuring outcomes; these are not problems that resolve through a software license, nor that every lender has the internal capabilities to solve alone. They require a partner that has done this before and can move from strategy to production without the lender carrying the full implementation risk. Global technology and AI builders like Cognizant bring lending-specific regulatory expertise and proven implementation

best practices that accelerate deployment and reduce compliance risk. A 2025 survey by Cognizant and HFS Research of [257 non-bank lenders and ecosystem partners](#) found that 30% of lenders work with full-service partners like Cognizant today, a figure projected to rise to 42% in two years. Outcome-based commercial models are becoming the preferred structure for exactly this reason: the partner co-funds the upfront investment, carries deployment risk by committing to deliver defined outcomes, and shares in the savings that successful delivery generates.

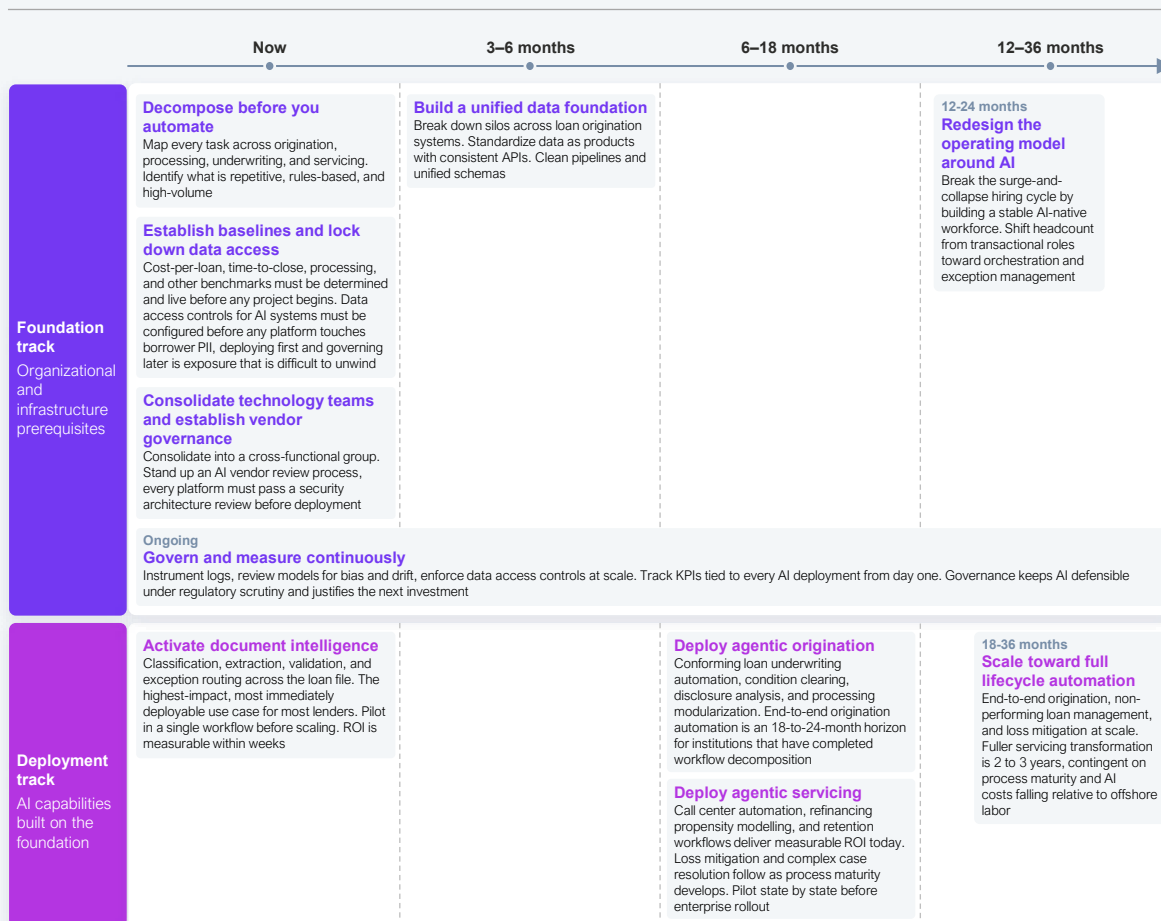
Start here: A practical sequence for deploying agentic AI at scale

The lenders pulling ahead aren't moving faster because they have better technology. They're moving faster because they made better decisions earlier about data, governance, and where AI gets embedded and where it doesn't. What follows is drawn from what the leaders in this research did, in roughly the order they did it in (see Exhibit 7).

Don't wait for perfect conditions to start. The foundation does not need to be complete before the first deployment begins but should be strong enough to support it. The firms that got this right

treated foundation-building and deployment as parallel tracks, not sequential ones. They started with the highest-impact, lowest-risk use cases while continuing to invest in data infrastructure, governance, and operating model redesign. What they never did was skip the foundation entirely or mistake early pilots as a substitute for it. Every firm that tried to shortcut the foundational work paid for it later: in failed deployments, in compliance exposure, and in AI that sat at the edges of the operation rather than inside it.

Exhibit 7: Sequencing is the strategy for deploying AI and agentic at scale



Sample size: Based on 11 US senior non-bank lender executives
 Source: HFS research in partnership with Cognizant

The Bottom Line: AI deployment alone won't decide the winners. The advantage goes to those who rethink how they operate. That foundational work is unglamorous, but it builds the readiness that lenders can't buy their way into.

Non-bank lenders scaled in a market where the economics worked in their favor. That market is gone. Elevated rates, compressed margins, and rising origination costs have exposed operating models built for a different era. For CIO, CTOs, and business unit heads, the answer is not more technology; it is a fundamental rethink of how technology, data, and automation are used to protect margins and a re-building of operating models that are resilient enough to perform across rate cycles.

The challenge is that AI maturity across non-bank lending is deeply uneven, and that unevenness is itself a strategic liability. Some lenders have done the hard work: redesigning workflows around AI, treating data as infrastructure, building governance before they scaled, and measuring outcomes with enough precision to know what is working and fund what comes next. But there are others who sit somewhere between fragmented experimentation and cautious observation, running AI in pockets, measuring it loosely, and waiting for a cleaner signal before committing to something bigger. That posture made sense two years ago. It does not make sense now. The gap between those two groups is already measurable in cost-per-loan, cycle time, and workforce stability. And it grows every quarter that fragmented AI remains the default rather than the problem being solved.

The foundational work that closes this gap is not glamorous, but it determines whether AI becomes a competitive advantage for your organization. The question is whether lenders use it to rethink how they operate or wait until the market forces the answer for them.

HFS Research authors



Divya Iyer
Practice Leader

Divya Iyer is a practice leader for HFS Research, leading industry analysis of banking and financial services and covering the intersection of technology, business, and financial services. She works closely with service providers, financial services enterprise clients, and the broader supporting ecosystem to actualize their goal of articulating the business impact and unlocking value in these relationships. In her two decades of experience, Divya has gained expertise in IT and business services and researched emerging and established digital business models, technologies, startups, and business solutions suppliers.



Don Ryan
Research Fellow &
Senior Advisor

Don Ryan is a research fellow & senior advisor for HFS Research. He joined the firm in August 2020 as senior vice president of research and consulting. Prior to HFS, Don was a director for global research and thought leadership for business and outsourcing services, digital technology, and AI deployment at KPMG LLC. While at KPMG, he collaborated with HFS on the highly publicized white paper, Enterprise Reboot: Scaling digital technologies in the new reality, a major study to delineate the impact of COVID-19 on technology implementation.



About Cognizant

Cognizant (NASDAQ: CTSI) is an AI builder and technology services provider, building the bridge between AI investment and enterprise value by building full-stack AI solutions for our clients. Our deep industry, process and engineering expertise enables us to build an organization's unique context into technology systems that amplify human potential, realize tangible returns and keep global enterprises ahead in a fast-changing world.

See how at www.cognizant.ai or @cognizant.

About HFS

- **INNOVATIVE**
- **INTREPID**
- **BOLD**

HFS Research is a leading global research and advisory firm helping Fortune 500 companies through IT and business transformation with bold insights and actionable strategies.

With an unmatched platform to reach, advise, and influence Global 2000 executives, we empower organizations to make decisive technology and service choices. Backed by fearless research and an impartial outside perspective, our insights give you the edge to stay ahead.



www.hfsresearch.com



[hfsresearch](https://www.linkedin.com/company/hfsresearch)



www.horsesforsources.com



www.horsesmouthpodcast.com