Taktile

A guide to responsibly deploying AI agents in financial institutions.

Dustin J. Eaton
CFE CAMS CFCI CFCS CAFP
CGSS CAMS-RM CAFS,
Taktile - Principal AML & Fraud

Jill Zucker Sheckman,
Varo Bank - EVP Credit Strategy
& Portfolio Management

Dr. Maximilian Eber,
Taktile - Co-Founder & CPTO

Contents

Starting point:

The gap between AI capability and responsible deployment at scale

AI sophistication is evolving rapidly. With the latest models, AI agents can now interpret unstructured documents, follow complex policies and processes, and make nuanced judgments as a human would – sometimes with even greater accuracy.

Today, an agent could verify customer identity and run adverse media searches during onboarding, then automate certain steps of credit risk assessment to accelerate underwriting. It could also triage fraud alerts based on deep research into customer history and detect complex AML patterns by connecting data across systems.

Most financial institutions no longer doubt what AI is capable of. Now the question is: **what does it look like to deploy AI agents responsibly – at a scale that delivers meaningful value?**

To answer this question, it is helpful to begin by looking at how we arrived at this turning point for AI in financial services. What exactly do the latest advancements make possible, and what does an informed yet proactive AI roadmap look like in this highly regulated landscape?

The breakthrough:
Why financial institutions can now automate high-stakes decisions with AI agents

Financial services teams are under pressure to show real value from AI within core operations. While the first wave of AI adoption saw individuals using generative AI as a personal productivity tool, now the mandate is to deploy agents systemically across entire lines of business.

This shift could transform how institutions make high-stakes decisions at every step of the customer journey – from deciding if a business passes compliance checks, to assessing credit risk and analyzing transactions for fraud. While the challenge is immense, leaders can feel encouraged. **The path to achieving safe, impactful AI transformation in financial institutions is clearer than it has ever been.**

In recent months, AI has reached a level where even regulators promote its potential to accelerate and optimize regulated workflows. Meanwhile, a benchmark from Taktile Labs found that agents can now surpass human accuracy in credit underwriting tasks like financial spreading: 96.5% accuracy vs. a 89% human baseline.

Collectively, the authors of this guide have spent over 20 years helping financial institutions build intelligent decision systems. Along the way, we've navigated constant regulatory complexity and countless AI hype cycles. We can say with confidence: AI is ready for financial services.

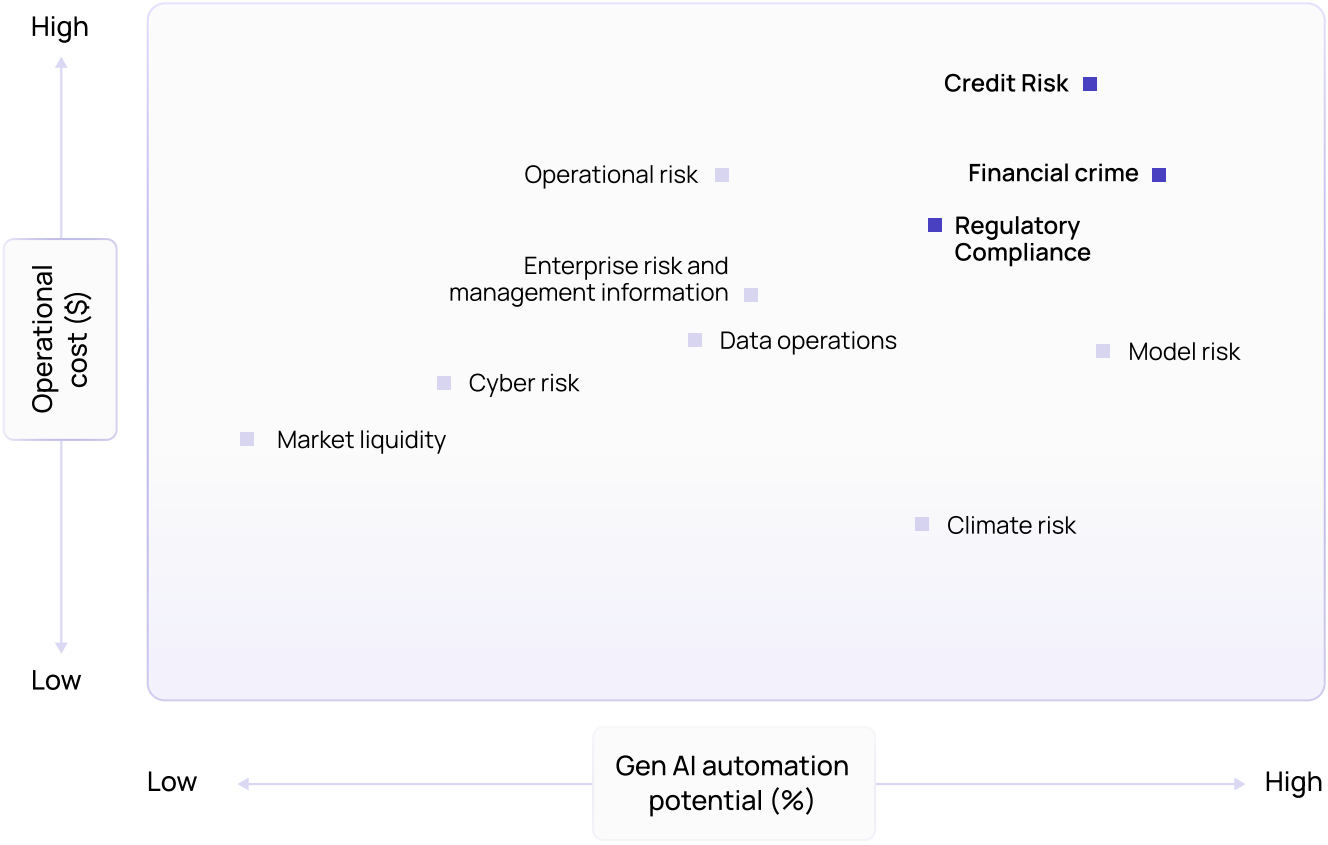
> Research from Taktile Labs shows that AI models crossed a critical threshold in December 2025. Agents built on the latest models can now **match or exceed human performance in tasks like financial spreading and KYB research**.

However, the path to reliable deployment at scale is complex, especially within larger institutions. While it is now relatively easy for a single engineer to build an agent that can reliably complete a task on a single device, the final 30% of the process is where many teams hit a wall. From connecting agents to data sources to orchestrating them end-to-end within existing systems, scaled AI deployment adds a whole new layer of intricacy.

In this guide, we offer an actionable roadmap to responsibly adopting agentic AI in credit, AML and compliance, and fraud. Our goal is simple: if you want to realize the benefits of AI within these processes but aren't sure where to start, you'll walk away knowing how to move forward with confidence.

Where to focus: the lines of business with the most to gain from agentic AI

McKinsey highlights three areas of financial services as having both the highest operational costs and the highest potential for AI-driven automation: credit risk, financial crime, and regulatory compliance.



Source: McKinsey & Company. "How Agentic AI Can Change the Way Banks Fight Financial Crime." Risk & Resilience, August 7, 2025.

These complex lines of business rely on large teams to gather and distill data from multiple sources, manually review customer cases and applications, and make nuanced judgments about risk.

The same McKinsey study reports that banks allocate 10 to 15% of their full-time workforce to AML and KYC alone. There's clear potential to elevate these teams with agents that automate investigative legwork, surfacing only the nuanced cases that require human judgment.

In credit teams, agents can help alleviate manual bottlenecks during initial risk assessment, as well as streamline reviews when applications escalate to final approvers.

What's stopped automation from going further, until now

Credit risk, financial crime, and compliance functions carry such a high operational burden because much of the work is manual by necessity. Until recently, it's only been possible to automate certain parts of high-stakes decisions such as whether a business should qualify for a loan, or if a transaction is fraudulent.

That's because:

- **Traditional rules engines can only automate steps in the decision journey where the outcome is pre-determined,** such as flagging transactions above a fixed threshold.
- **These systems can only make decisions based on highly structured data.** They can interpret income and credit score arranged in a tabular format, but they can't autonomously decide if a customer has submitted all necessary documentation for assessment.

> Financial institutions are already using AI to some extent in these decision use cases. For example, rules engines are often connected to ML models that generate risk scores. However, outside of consumer credit underwriting, we haven't yet seen financial institutions using AI to automate decisions end-to-end.
>
> Agentic AI is enabling a new operating model.

How agentic AI changes what's possible

The breakthrough of agentic AI is that it can handle the unstructured data and unpredictable situations that rules engines are can't process. **Financial services teams who have battled the limits of automation for years now have a powerful new tool.**

In its simplest terms, an agent is an LLM that doesn't just interpret data, it takes action. And it doesn't just act, it learns and improves through each new scenario or piece of human feedback.

Agents can both digest and reason over unstructured inputs like websites and customer documents. This means teams can task them with much more ambiguous questions than typical rules engine commands, like "triage based on age". You could prompt an agent to:

- "analyze financial statements to determine credit risk score" or
- "cross-check identity documents for fraud indicators"

As a result of these new capabilities, it is now possible to automate certain decisions end-to-end. Companies like Brex are already deploying agentic systems that fully automate decisions for straightforward fraud or KYC cases, while surfacing more nuanced cases to human analysts.

An agent tasked with gathering KYB data from a website could run an initial search, notice that some information is missing, conduct a further search, and raise an alert to a human analyst if it spots any red flags.

Critically, every override and correction becomes training data for the system, creating a human feedback loop that improves accuracy over time. For edge cases, human-agent collaboration could operate as follows:

A customer that sits right at the edge of an institution's risk threshold applies for a business loan.

An agent could analyze data from multiple sources and suggest the appropriate pricing based on organizational policy.

Identifying this as a nuanced case, the agent would surface it to a human underwriter for final judgement.

Beyond automation: agentic AI enables more than efficiency

Because agents enable far more than simple automation, the benefits can extend beyond efficiency and cost reduction. McKinsey predicts a 15-20% decrease in banks' overall costs as they adopt AI, but there are also opportunities for value creation.

Commercial lending teams are turning smoother customer experiences into a competitive advantage by using AI to reduce loan approval wait times. Additionally, by applying credit policies consistently every time, agents can help strengthen internal controls without sacrificing speed.

In financial crime use cases, Federal Reserve research demonstrates that AI-driven AML systems can achieve up to 92% fewer false positives compared to conventional methods. Fraud teams are using agents to analyze alerts and recommend rule updates to reduce system noise.

These examples point to a common theme. The most meaningful gains from agentic AI come not just from accelerating existing processes, but from achieving outcomes that were not previously within reach.

From theory to practice:
How this guide can help you put agentic AI into action responsibly and at scale

The opportunity is clear. However, the path from concept to deployment remains opaque for many teams. This guide is designed to bridge that gap. Together, we'll navigate through three phases of learning that will help you move from theory to practical agent deployments:

1. **Identify your opportunities:** Develop your understanding of the problems agents can solve, and zero in on the highest-impact areas to deploy them within your teams and systems.
2. **Understand the infrastructure:** Learn how to orchestrate agents for responsible governance and effective performance in production.
3. **Activate your roadmap:** Establish the step-by-step pathway to deploying your first agents, scaling them across your organization, and maintaining them over time.

Our goal is to help financial services teams responsibly activate agentic AI where it's expected to deliver the most value. Given that, we won't cover use cases that exist in all industries, such as code generation, internal chat, or customer support. Instead, we will focus on agent deployments within:

- Commercial credit underwriting
- AML and compliance
- Fraud prevention and investigation

To start, we'll walk through some key steps in each use case where agents can improve efficiency, enhance decision quality, and create better user experiences. By breaking down the process in this way, you'll get a clearer idea of exactly how agentic AI solves common problems.

It is unlikely that we'll cover the potential value add of AI in every step of your specific process. Our goal is to give you enough understanding that you can apply the patterns and benefits you learn to other steps within your operations. As you progress, you'll see how agents connect across the journey to create an intelligent system.

By the end of this guide, you'll be fluent enough in how agents work to identify problems in your own workflows and imagine how agents could solve them.

Mapping your impact areas:
1. Agents in commercial credit underwriting

Consumer loan underwriting has been relatively simple to automate for years. Since lenders typically assess consumer credit risk based on fixed knockout criteria and credit scores, rules engines can process these standardized inputs and quickly generate a decision.

Commercial underwriting is a different case altogether. The complexity of business documentation and financial statements mean that most B2B underwriting teams still rely on highly manual processes.

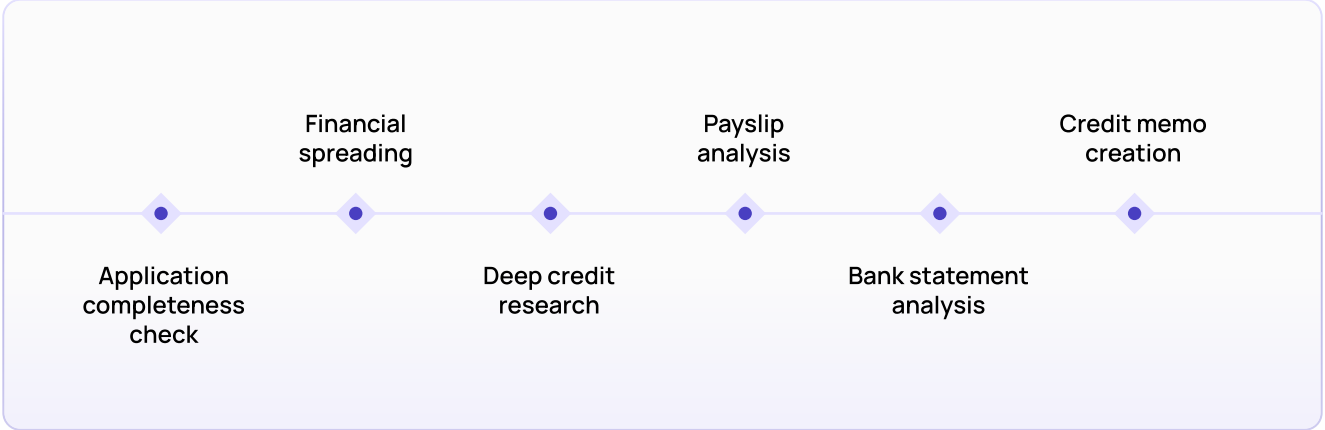
Now, agentic AI is changing what's possible.

Agentic AI can also deliver significant value in consumer credit use cases like mortgages, which require proof of income verification via pay stubs, financial statements, and tax returns. Teams can use agents to automate this analysis and give customers decisions faster.

Credit teams can use agentic AI to expand automation in four main areas of the underwriting workflow:

1. **Document intake and application completeness checks:** Agents can review a customer's documents and immediately prompt them to upload additional supporting documents if needed, reducing back-and-forth time for underwriters.
2. **Financial spreading and analysis:** Analysts can use agents to extract, normalize, and validate financial statement data and calculate key ratios. An agent could also analyze the financial spread to craft an initial risk summary.
3. **Case management and human review:** Even as teams begin using agents to fully automate lower-risk loan approvals, human oversight will still be necessary in complex cases. Agents can be trained to flag these cases to human reviewers with all relevant context attached.
4. **Credit memo creation:** As applications reach the final stage, an agent could summarize all the information collected into a draft credit memo. Teams would house memos within a review dashboard where reviewers can easily see the data and decisions behind them.

Let's take a closer look at two steps in the underwriting journey to understand exactly how agents can help to reduce manual workload and accelerate decisions.



Snapshot of moments in the underwriting process where agents can help improve decision speed and accuracy.

Agent use case: Application completeness check

Business value: Enhanced customer experience, lower cost per decision

Without agents, underwriters manually review submitted loan applications to verify that all required documents and data are present before beginning their credit analysis. Incomplete applications create friction for customers, who often wait days to learn if additional documentation is required.

With agents, completeness checks can be automated as underwriters focus on more complex steps in parallel:

- Agents validate application completeness by checking for missing or inconsistent information before files reach an underwriter's queue.
- Since an agent can autonomously review information and decide the next step, it could identify any missing or inconsistent documentation and prompt customers to resubmit.
- Using OCR and natural language processing, agents verify not only whether required documents are present, but also whether the information within those documents aligns with public records or third-party sources.

In an agentic system, underwriters receive files that have already been checked for completeness. This allows them to **focus their time on credit analysis and decision-making rather than coordinating back-and-forth on missing information**. Reduced processing times for customers can also help to reduce drop-off.

Agent use case: Financial spreading

Business value: Better decision quality, lower cost per decision

Financial spreading involves extracting financial data from income statements and balance sheets, and normalizing that data into a standardized format for analysis.

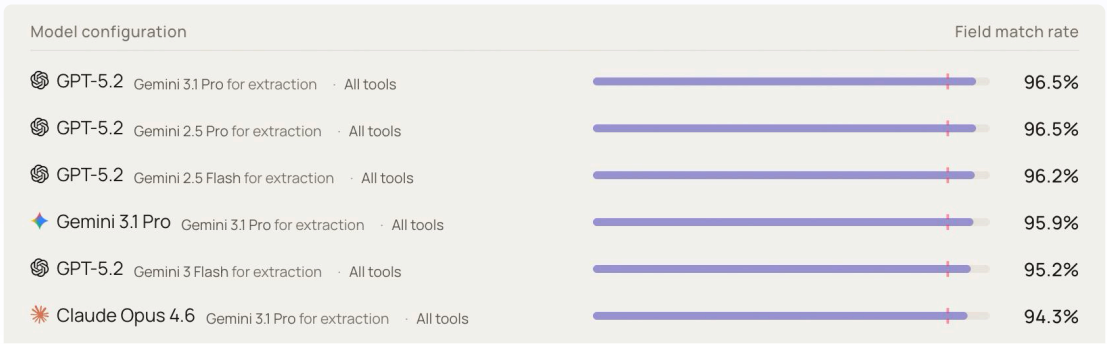
Without agents, underwriters have to manually input data from financial statements into spreadsheets to conduct their analysis. Businesses present financial information in varying formats, making automated extraction impossible. The manual nature of financial spreading creates bottlenecks in the underwriting process, and extends decision time for customers.

With agents, financial spreading is faster, and often more accurate:

- Agents digest financial statements and extract relevant data points accurately, regardless of how the information is structured or presented.
- Initial data validation is automated as agents cross-check figures for consistency—such as verifying that line items add up correctly—and flag discrepancies for human review.
- Over time, as agents process more financial statements across different industries, they learn common formats and improve extraction accuracy.

Agents can now surpass human accuracy in financial spreading.

In our benchmark at Taktile Labs, we found that agents using the latest models surpass human performance in financial spreading, achieving **96.5% accuracy vs. an 89% human baseline.** We see this as a yardstick for AI's potential to automate complex tasks where regulatory and business requirements make the margin for error incredibly tight.

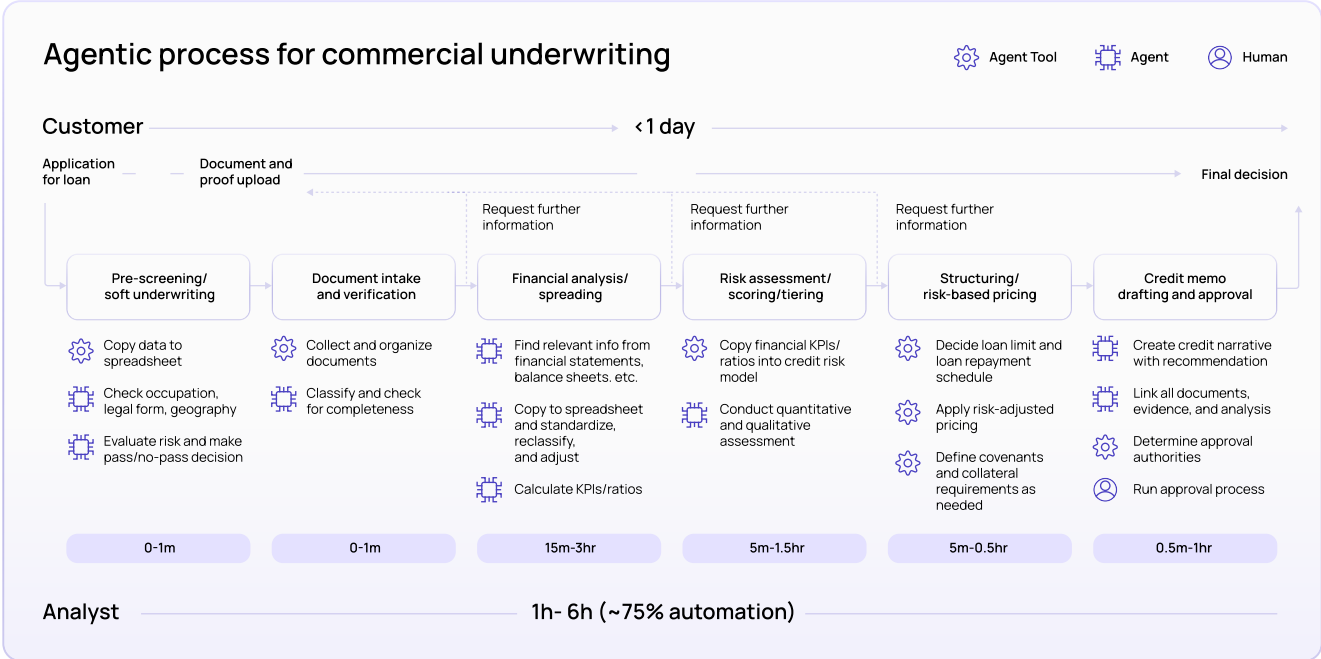


Putting it together: an agentic system for credit underwriting

Once agents are embedded across the credit journey, teams can route each application through a coordinated set of agentic capabilities to achieve faster underwriting decisions. These decisions are also often more consistent: agents aren't subject to human error, as they can be configured to apply credit policies with the same level of rigor in every case.

The agentic workflow enables teams to automatically verify customer inputs, convert unstructured files into structured signals, and assemble a credit memo for refinement and review. With the right front end, it's easy for humans to collaborate with agents, and for everyone involved in the review process to work from the same holistic source of truth.

The infographic below is one example of how specialized agents could be orchestrated to create an agentic underwriting process where decision time is reduced from days to hours.



An agent tool is a specific capability or function an agent can call to gather information. For example, querying a database, running a web search, or reading a document.

An agent is an AI system that autonomously reasons over a goal, decides which tools to use and in what order, interprets the results, and determines next steps—including when to escalate to a human. It orchestrates tools to complete multi-step tasks end-to-end.

In short: a tool does one thing, while an agent uses many tools to decide on the right action.

Mapping your impact areas:
2. Agents in compliance and AML

Anti-money laundering (AML) and compliance are among the most resource-intensive operations in financial services. During KYB, routine checks like adverse media and sanctions screening can consume hours of an analyst's work day. Further along the customer journey, transaction monitoring systems overwhelm teams with excessive false positives.

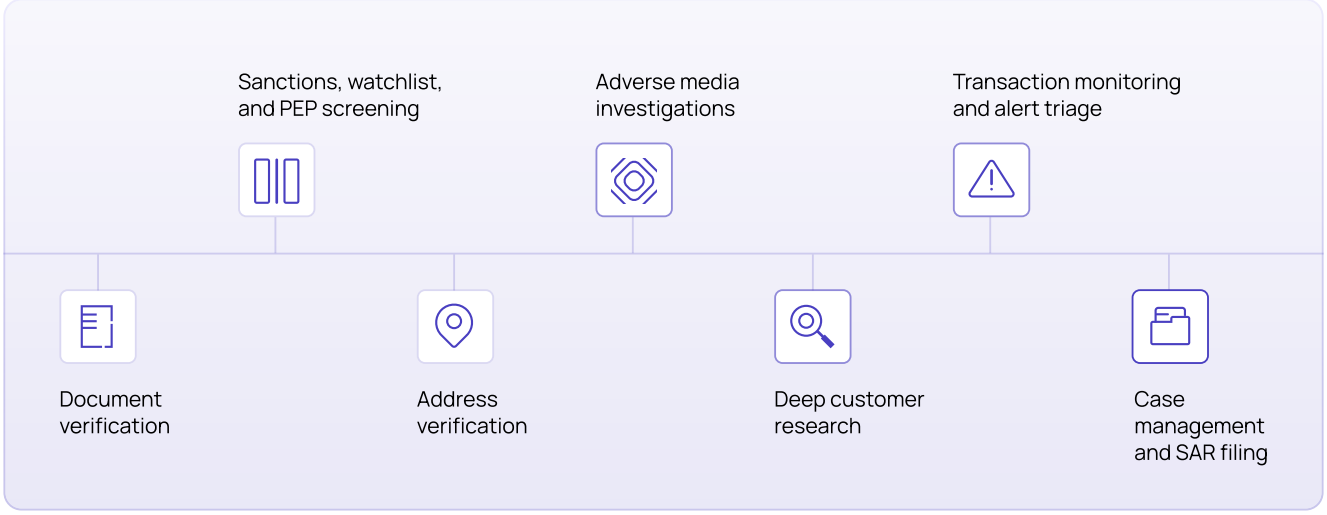
When analysts are consumed by false alerts, they have less time to focus on catching real crime. Additionally, with data scattered across systems, traditional rules engines can miss sophisticated AML tactics—especially those that involve multiple actors working together.

Transformation is long overdue, and AI is finally positioned to deliver the benefits it has long promised. In fact, regulators are actively encouraging AI adoption in AML use cases as the technology grows in sophistication. Deployed within the right guardrails, agents can help teams automate time-consuming investigative work and reduce the burden of false positives.

As we shift from theory to practice, we'll break down the AML and compliance journey into three connected layers to show where agents can have the biggest impact:

- Customer onboarding
- Ongoing transaction monitoring
- Case management and SAR filing

At each stage, teams can collaborate with agents to simplify their workflows, catch more crime, and resolve customer cases faster. Let's walk through these three phases of the customer lifecycle to see where and how agents make the most impact.



Layer 1: Customer onboarding

In this phase, agents support teams by reducing the manual burden of collecting, inputting, and verifying customer data. That's because if we zoom out, compliance decisions during onboarding come down to checking different types of customer information against internal policies and external data sources:

- Customer-provided information such as Business Ownership or Source of Funds
- Third-party data like sanctions watch-lists
- Web-based searches like address verification and adverse media

Agent use case: Document verification

Business value: Lower cost per decision, enhanced customer experience

In this task, the core value of agents is their ability to not only extract data but understand it. Agents can compile and cross-check customer information, autonomously resolving data gaps by coordinating additional information requests with the customer.

Without agents: Analysts use traditional Optical Character Recognition (OCR) to pull data from documents and manually cross-check it against registry entries and identify verification sources. Standard OCRs can't understand what the data means, identify errors, or decide what to do if information is missing.

With agents, manual workload is significantly reduced:

- An agent could learn to accurately read and pull data from customer documents, even if the layout is out of sync with standard formats.
- The agent automatically cross-references extracted data against registry entries and identity verification sources.
- It identifies inconsistencies, for example a CEO listed on a document who doesn't appear in public records, and flags them with a summary for the analyst.
- Rather than spend their time compiling data and going back and forth with customers, analysts focus instead on reviewing the case and making a decision.

Agent use case: Adverse media screening

Business value: Better decision quality, lower cost per decision

Without agents: While some institutions already automate the initial search itself, evaluating risk level, documenting findings, and deciding the next best action is still often a manual process.

With agents, analysts can reduce manual labor:

- Agents gather and read articles, link them to the right customer, and filter out low-risk alerts.
- When genuine signals are surfaced, agents can summarize findings and create a risk summary for analysts to review.
- As customer behavior evolves, agents can continuously monitor for adverse media signals, triggering human review as needed when new concerns arise.

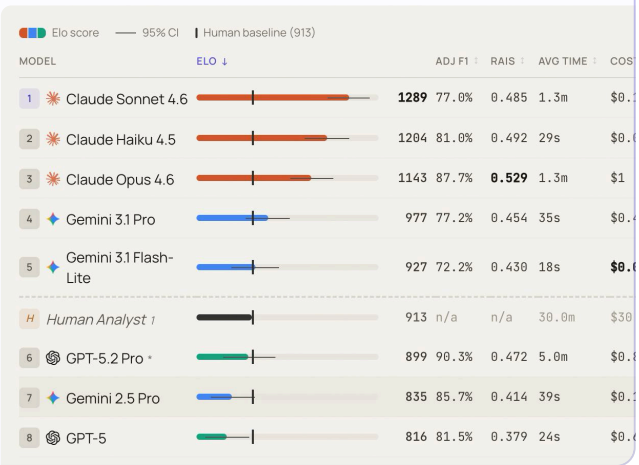
Watch-out: Sanctions and PEP screening errors can carry immediate regulatory penalties. When deploying agents for this use case, we recommend doing so with lower thresholds for human intervention, and mandatory senior-level review of documentation.

The case for a hybrid human-agent model for adverse media screening.

Even as institutions navigate regulatory concerns around using AI agents in due diligence, research shows that a hybrid human-agent approach can prove more reliable than a purely human system unsupported by AI tools.

At Taktile Labs, we tested an adverse media agent built on 7 frontier AI models across 47 real businesses. On a 20-point rubric spanning evidence and source quality, entity identification, and risk assessment, humans averaged 13.5/20 vs. 14.6/20 for AI agents.

A hybrid model where an agent screens first and a human reviews uncertain cases **can reduce analyst workload by 93%.**



Layer 2: Transaction monitoring

Once customers are onboarded, the next step in the AML lifecycle is ongoing transaction monitoring. Traditionally, transaction monitoring systems have been plagued by an impossible trade-off: cast a wide net and overwhelm analysts with false positives, or tighten the rules and risk missing genuine threats.

AML teams can use agentic AI to change the equation.

Agent use case: Intelligent alert triage

Business value: Better decision quality, lower cost per decision, enhanced customer experience

Without agents, investigators spend an excessive amount of time investigating false positives that should never have hit their caseload to begin with. Rules engines flag alerts based on fixed thresholds, but are unable to filter based on additional nuance or context.

With agents, the workflow changes fundamentally:

- When an alert fires, an agent could kick off the investigation by compiling transaction history, customer background, behavior patterns, and more.
- The agent would then classify the alert, resolving low-risk cases automatically while surfacing higher-risk cases to a human reviewer.
- For escalated cases, the agent delivers a pre-compiled case summary to the analyst, rather than a raw alert.
- As analysts approve, modify, or override agent decisions, the system refines its judgment over time.

Layer 3: Case management and SAR filing

This layer is where the value of the agentic system multiplies. Let's assume that our transaction monitoring agent has filtered out false positives and automatically resolved lower risk alerts.

Now, when higher-risk cases surface in an analyst's workflow, an agent can deliver them with all relevant context, plus their recommendation for how to proceed. As confidence in agent recommendations grows, AML teams can extend automation—allowing analysts to focus increasingly on more complex cases that don't fit a typical pattern.

Agent use case: SAR narrative generation

Business value: Better decision quality, lower cost per decision

Without agents, analysts piece together context from multiple systems before they can even begin to form an opinion. For cases that require a Suspicious Activity Report (SAR), they might spend hours synthesizing findings into compliant documentation.

With agents:

- Escalated cases surface to analysts with all relevant information, including customer history, transaction patterns, prior alerts, and risk scores.
- When a SAR is required, an agent can draft an initial report, flagging key risk indicators and structuring findings in the required format.
- Higher level reviewers are able to see both the SAR narrative and all relevant customer information in a single view, streamlining edits and refinements.

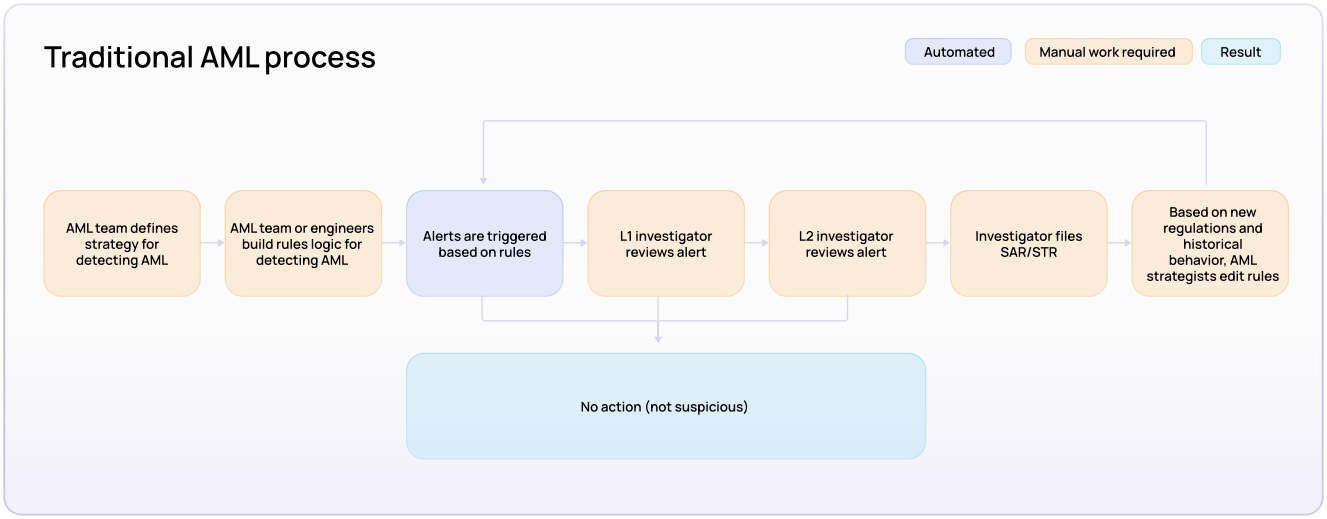
Putting it together: an agentic system for AML

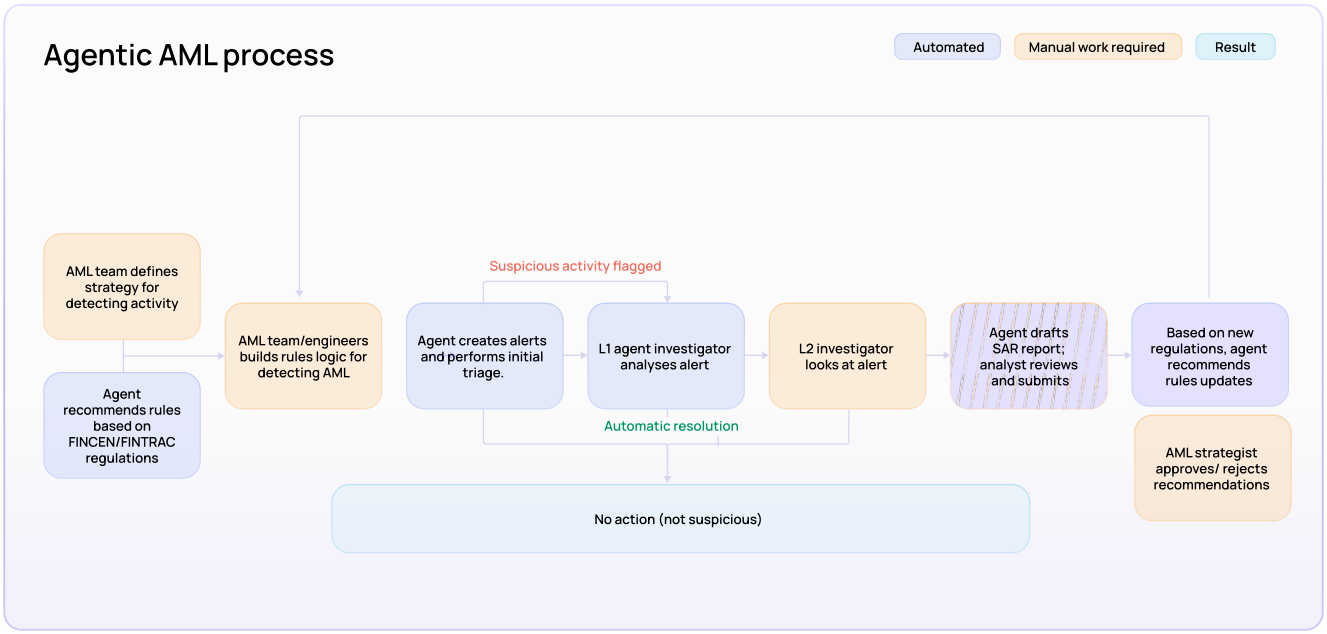
Now you have an overview of how agents can make an impact in each layer of the AML system, it's helpful to step back and see how they work together.

In the following infographics, you'll see one iteration of how your team could reimagine the traditional AML process, using agents to reduce manual workload, improve detection, and accelerate case resolution.

You will also notice that agents can also be used to make foundational system rules more intelligent, recommending adjustments as they notice common patterns. Additionally, you could feed agents regulatory guidelines, configuring them to automatically suggest rules updates based on the latest advisories.

Even if these infographics don't exactly replicate your current system, they can be a helpful blueprint for re-mapping your process. Some teams refer to this as translating their Standard Operating Procedure (SOP) into an Agent Operating Procedure (AOP).





Over time, this intelligent system creates a shared, evolving knowledge base for what accurate compliance and AML decision-making looks like in practice. Rather than analysts being split into individual high-flyers and lower-performers, the whole organization gets smarter, and customers benefit from more consistent experiences.

Mapping your impact areas:
3. Agents in fraud prevention and investigation

Fraud teams are under pressure to keep pace with crime as fraudsters adopt AI-driven techniques that are difficult to detect with traditional controls.

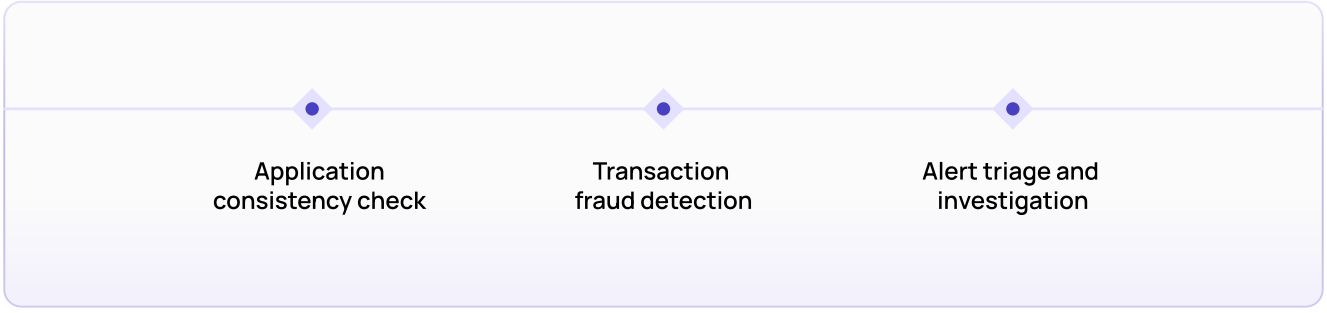
Deloitte's Center for Financial Services forecasts that AI could increase U.S. fraud losses to as much as $40 billion by 2027. This comes as fraud rings are using AI to create deepfakes and impersonate legitimate customers, develop highly convincing phishing campaigns across borders, and research institutions' controls to identify vulnerabilities.

The shift has created an imperative to innovate: the most effective way to fight AI is with equally sophisticated AI. However, the challenge is not only that threats are evolving—it's that fraud systems are often fragmented. Identity verification vendors, device and network intelligence tools, and transaction monitoring systems each generate their own outputs. Teams lack the holistic customer view required to separate false positives from real threats, while sophisticated fraudsters slip through the gaps between systems.

In light of these pressures, an agentic fraud system would ideally prioritize three things:

- Contextual investigation: investigating the context behind an alert to understand why a flag was created before routing it to a human reviewer.
- Cross-signal synthesis: enabling agents to reason holistically across data outputs from tools that weren't initially designed to talk to each other.
- Adaptive learning: building a system that gets harder to fool as it processes more cases, identifies patterns, and absorbs analyst feedback.

Deployed thoughtfully, agents can improve detection quality, reduce manual work, and increase investigation speed—enabling teams to catch more crime while creating less disruptive customer experiences. **Below, we explore three use cases that illustrate how agentic AI can strengthen and refine fraud defenses at key points in the customer journey.**



Agent use case: Application consistency check

Business value: Better decision quality, lower cost, enhanced customer experience

Fraud teams often check customer-supplied information against several siloed verification tools, each producing independent outputs. An application consistency check agent reasons across all of them to answer a question no single tool can: does everything the customer supplied add up?

Without agents, each verification tool produces its own pass or fail signal, and no system is responsible for comparing them. A utility statement with a different address than the driver's license might slip through because no tool is looking at the full picture.

With agents, it's easier to capture a holistic view:

- The agent reviews all customer-supplied information, checking for consistency across documents, application data, and third-party sources.
- It flags discrepancies and assesses whether they represent harmless formatting issues or genuine inconsistencies that warrant deeper review.
- Rather than routing every mismatch to a human, the agent classifies risk level and auto-resolves low-risk flags, surfacing the cases where something substantive doesn't add up.
- Over time, the agent learns which patterns of inconsistency tend to signal fraud versus genuine customer error, improving its judgement with each decision.

Agent use case: Transaction fraud detection

Business value: Better decision quality, lower cost, enhanced customer experience

Without agents, when a fraud flag fires, an investigator has to manually gather context from multiple systems before they can even begin their analysis. The alert tells them something may be wrong, but it doesn't give them enough information to determine the true level of risk at-a-glance.

With agents, the investigative workflow changes fundamentally:

- When a flag is raised, the agent pulls all relevant context, from transaction history to prior disputes and any other risk signals that might inform the current case.
- It assesses the alert in light of that full picture. Two consecutive charge-backs look very different if this customer has a clean history, versus a profile with warning signs.
- For low-risk cases, the agent can auto-disposition the alert without human involvement. It may also add the customer to a gray list: a watch category for accounts that haven't crossed a hard threshold but have shown behavior worth monitoring.
- For escalated cases, the agent delivers a case summary to the analyst. The analyst can make a better decision based on clear evidence, and resolve the case faster.

Putting it together: an agentic system for fraud prevention and investigation

In an agentic system, clear false positives can be resolved automatically while ambiguous cases are escalated with full context. Then, rather than investigate alerts across siloed tools, agents can pull onboarding data, transaction signals, and behavioral context into a single view for analysts.

As a result, teams are often able to resolve cases faster with more accurate outcomes—while analyst feedback helps improve agent reliability over time. **Let's look at how the different parts of the agentic system fit together in one specific example: a charge-back dispute.**

An agentic fraud system handling a charge-back could:

- Pull the customer transaction history, onboarding profile and previous dispute record
- Cross-reference the disputed transaction against behavioural signals
- Assess whether the pattern is consistent with legitimate use or indicative of fraud
- Auto-resolve if confidence is high or escalate to an analyst with a pre-complied summary
- Add the customer to a grey list if the behaviour is borderline, and monitor for further signals
- Digest analyst feedback so the next similar case is handled with sharper judgment

Agent orchestration: how to configure agents within existing operations for reliable outcomes

You may now have a identified an agent that you want to explore deploying within your operations. However, there are several layers of orchestration between a promising agent demo and an agent that can process thousands of decisions within a complex, regulated workflow. Critically, institutions that achieve real transformation go beyond equipping teams with agents as standalone tools; instead, they focus on integrating agents into end-to-end workflows.

Here's what we mean by that distinction. If an AML analyst uses an agent in isolation to synthesize customer data during an investigation, the core workflow itself doesn't change. Your compliance team would still have to manually review 100% of cases, copying agent outputs between systems and making approval decisions one by one. The same would be true if an underwriter used a standalone agent to assess credit risk more quickly based on compiled data.

The meaningful value comes when agents are embedded within your system. Instead of assisting humans case-by-case, an integrated agent could autonomously handle up to 95% of decisions—surfacing only the 5% of cases that genuinely require human judgment. Plus, your whole team would work from a single source of truth that gets smarter over time.

Achieving this agentic system requires connecting agents with a surrounding infrastructure that has four main components:

- **Context layer:** Clear connections to the internal and external data sources agents need to reason and act reliably.
- **Deterministic layer:** Policies and guardrails that govern agent behavior and ensure reliable, compliant outcomes.
- **Human review layer:** Some kind of case management interface that surfaces escalated decisions to analysts with full context so they can make informed decisions.
- **Monitoring layer:** Transparent audit trails and monitoring dashboards where teams can track agent decisions for regulatory review—as well as optimize agent performance.

Think of these four components as the difference between a standalone agent and an agentic system. In the sections below, we'll walk through each one in more detail.

1. Context layer: Giving agents the right data to make the right decisions

For AI to add value in complex financial services use cases like credit and AML, it can't just reduce manual work and speed up processes. Ideally, agents help teams use data more intelligently, enabling decisions that support growth while managing risk.

But today, even human teams don't always have all the context they need to make the most optimal decision in the moment. Customer data is often stored across multiple systems in inconsistent formats. Critical knowledge might live in one senior team member's head, while newer employees make do with the information readily accessible to them. Inconsistent context often produces inconsistent decisions.

When agents enter the picture, there's an opportunity to enhance both decision quality and consistency—but to achieve this, having a clear picture of the data that guides your decision-making is essential. To ensure agents make accurate, compliant decisions in practice, we recommend resolving disparate sources into a single context layer that weaves together:

- Customer-supplied data
- Third-party sources like cash-flow data or watchlists
- Internal databases and systems

As teams construct this context layer, there are three steps not to miss:

- **Map the internal and external data needed for your desired output:** Start by documenting every data source a human or rules engine would need in order to make the right decision within your current workflow. Your agents will need access to the same sources, so check these connections are up-to-date and have clear owners assigned for ongoing maintenance.
- **Compile historical data for "sandbox" testing:** In production, an agent may call external providers like credit bureaus. But during development, this is too costly and can negatively impact customers, for example deteriorating their credit score. Plan to compile historical data that includes edge cases so you can stress-test your agent's response to nuance.
- **Evaluate latency when assessing external data integrations:** When your agent is embedded in a real-time workflow, every third-party data call can slow down the time from prompt to output. For agents powered by multiple third-party calls, the provider response time can determine whether an agentic workflow is viable at scale.

2. Deterministic layer: Keeping agents under control

One of the biggest hesitations around deploying AI in regulated financial institutions is the potential for erroneous decisions that could cause compliance issues, monetary losses, or wider system breakdowns. This concern is natural. Agents are an emerging technology, and many teams are just beginning to learn what responsible deployment looks like in practice.

However, it is important to note that there is risk in any system. Even humans can apply policies inconsistently or overlook risk signals in their research. They key is to balance the benefits of agents' interpretive capabilities with the continued need for governance and guardrails. People work within governed frameworks today, and so should agents.

Remember that the main difference between agent-driven and rule-driven actions is that agents are probabilistic: the same input can produce slightly different outputs across runs, whereas rules execute the same way every time.

Given that, **we recommend fencing in agents with guardrails that check their work and limit the actions they are permitted to take.** We also suggest a hybrid rules-based and agentic approach to control both costs and risk. You can build this deterministic layer on three pillars:

1. Guide agents with guardrails for consistency, safety, and control

There are controls you can put in place to limit what agents have permission to do within your organization, and to ensure the reliability and consistency of agent outputs:

- **Safety guardrails** control what an agent can do, for example limiting what kind of data they can access or when they can make a fully autonomous decision vs. escalating to a human for final review.
- **Reliability guardrails** validate the agent's work before a final decision is made.
  - For example, you could have a financial spreading agent extract data and calculate total revenue from raw customer documents — while including a rule that cross-checks certain metrics to make sure everything adds up.

2. Configure agents with your internal policies built-in

Credit, fraud, and compliance teams typically make decisions based on fixed policies, some of which are required by regulators.

As you translate decision-making into an agentic system, it's important to ensure that non-negotiable policies continue to govern agent behavior—even as agents add an investigative layer that enhances decision-making beyond straightforward rules.

You can encode these policies in two places:

- **In the decision system:** rule nodes in your decision workflow can enforce fixed requirements and thresholds; for example, always flagging transactions over a specific amount for fraud analysis.
- **In the agent:** prompts and tools can be configured to follows your standard operating procedure, including the steps to take, the data to evaluate, and how to structure outputs for review.

3. Get clear on where agents add value, and where rules make more sense

Not every step in your decision process will benefit from an agent. With so much industry hype around AI, it can feel like you need to solve every problem with an agent. But this approach can actually slow down decision-making, and increase costs unnecessarily.

Unlike rules engines, agents require running an LLM, which can be costly in terms of both data usage and model subscription costs. Calling agents as opposed to running a rule-based decision can also increase latency. We recommend the following approach:

- **Continue to use rules for deterministic decision points**, for example eligibility thresholds like age limits, or other fixed policy constraints.
- **Use agents when a decision requires multi-step reasoning and investigation**, such as pulling an entity name from a customer document, checking this against a corporate registry, and asking the customer to submit additional information.

3. Human review layer: Designing for human-agent collaboration

In a sophisticated agentic workflow, agents would autonomously handle the majority of cases end-to-end, surfacing only nuanced cases for human judgment. We recommend beginning with high-touch human review, expanding automation as your confidence grows. Regardless, it's always important to design agentic systems for easy collaboration between agents and people.

We suggest building your human-agent collaboration framework with the components below:

1: A review interface that gives teams holistic context for informed decisions

Agents need clear, comprehensive context to make good decisions—and so do people. Using agents to route a complex fraud alert or high-risk credit application to a human reviewer only adds value if that reviewer can also make a good decision quickly. To facilitate this, humans and agents would collaborate within a single interface that:

- Consolidates full customer history, data, and risk signals in a single view—so the analyst or underwriter doesn't have to toggle between systems
- Surfaces the agent's reasoning and the evidence it used, not just its conclusion
- Shows a clear recommended next action, while giving reviewers an easy way to override or share feedback on this recommendation
- Supports configurable workflows and escalation patterns that can mirror and support a team's existing processes

2: A co-pilot that acts as a research assistant inside the case manager

The most powerful review interfaces include an AI assistant or co-pilot. This feature allows analysts and underwriters to ask questions about the current case in natural language. Rather than sift through documents and databases, they can quickly surface the data most pertinent to the decision in front of them, and other context that could help to make a decision faster.

For example, an effective AI co-pilot could enable analysts to:

- Instantly access additional information on a customer or transaction (such as querying transaction history or KYC information)
- Request a short, bulleted summary of a complex credit history
- Compare the current case against similar historical cases and use these insights to make a more informed decision

3: A feedback loop that empowers teams to enhance agent performance

Every time an analyst approves, overrides, or adds context to an agent action, that feedback should run back into the system to inform future decisions. With diligent feedback, it wouldn't take long for an agent to meaningfully refine their judgment in your specific context. For example, in a fraud system, feedback could help agents hone their ability to separate genuine fraud alerts from false positives.

As your team's knowledge funnels into one central source of truth, institutional knowledge is codified rather than staying in individual contributors' heads. This can help to improve decision quality and consistency across an entire team, not just among the most experienced members.

4. Monitoring layer: Auditing decisions and optimizing performance

Once your agents are running in production, teams need a way to monitor every decision they make, identify issues if things start to go wrong, and keep improving performance over time. Keeping an accurate record of how decisions were made, both agent and human, is also essential for regulatory compliance.

An effective monitoring layer will include:

- **Explainable audit trails for every agent action**, including the data and tools the agent called to to reach a decision.
- **Real-time testing capabilities** that allow you to measure agent accuracy and effectiveness when using different configurations of AI models, risk thresholds, and tools.
- **Performance tracking dashboards** that give you a clear view of agent accuracy, especially any drift below the human baseline, so you can continuously improve and adjust over time.

By deploying your agents within an infrastructure that threads together context, control, collaboration, and monitoring, you'll be in a strong position to not only implement effective agents, but to scale them across your organization.

Activating your deployment plan:
How to successfully integrate an agent

To help you move to the next stage, we've set out a roadmap for deploying and scaling agents at speed while ensuring compliance, maintaining internal alignment, and working toward measurable impact.

Based on our experience in agent deployments within larger institutions, we've broken this down into a five-step framework. However, the below is helpful guidance for organizations of any size.

1. Define a clear end goal and metrics for success

This may sound obvious, but it is very common for teams to start with the agent they want to build, rather than the problem they want to solve. A successful agent project starts with a clearly defined objective and measurable outcomes:

- **Align on the problem you're solving and the outcome you want to achieve.** Your desired outcome might have several layers, since streamlined operations can impact customer experience, and better decision quality can benefit P&L.
- **Determine how you will measure this success in practice.** These measures of success can include both quantitative and qualitative metrics. For example, goal metrics for an AML alert triage agent could include:
  - X% false positive reduction
  - X% decrease in analyst workload
  - Improvement in analyst work life and engagement
- **Set rigorous guardrails for the project scope.** It can be tempting to try and solve everything at once. But it's more realistic to focus on a single use case within a single team to start, and expand from there as you build confidence in agent performance.

By prioritizing this first step, you'll avoid drifting into endless iteration, expanding scope, and debating what "good enough" means until the project stalls.

2. Map the current process, then translate into an agentic workflow

Before introducing an AI agent into a high-stakes workflow, it's important to first answer a simple question: how does the process run today and why? By mapping the process as it exists end-to-end, you'll be able to clearly define the steps an agent would need to carry out—and where human input is still required for safety or compliance.

As you map your Agentic Operating Procedure, consider questions such as:

- What triggers the workflow, and what data source is called at that moment?
- What are the deterministic steps that must run the same way every time? This could include eliminating credit applicants below a certain age threshold or flagging transactions above a specific amount.
- Where do teams have to do manual investigative work like extracting information from documents, running web searches, or writing case summaries?
- What are the scenarios where cases or applications are escalated for deeper research?
- What are the regulatory requirements or internal policies that govern your process?

3. Establish joint ownership between business and technical teams

Establishing joint ownership of agent projects across business owners and technical teams helps help to ensure your agents are set up to:

- Solve a real business problem
- Generate accurate outputs and decisions
- Operate effectively and compliantly within your existing systems

While technical buy-in is essential, it's equally important to keep business owners in the loop at every step. These experts understand the existing workflow, including the main issues and bottlenecks they want to address with AI, and the regulatory requirements that guide where human review is non-negotiable. They also have the topic-specific knowledge needed to judge whether agent outputs are correct.

Consider establishing an organizational structure that keeps both sides engaged and accountable before, during, and after agent deployment.

4. Audit your data and build data integration time into your roadmap

Many teams build agent prototypes using a single set of historic data designed for testing purposes. But if you want an agent to process real decisions for real customers, it needs to be integrated with multiple active data and static sources. Some of these will be supplied through third-parties that demand lengthy contracting processes.

We suggest that teams begin by establishing a clear view of the data the agent will use to make decisions. Auditing data structure and quality will help you to ensure that your agent can learn from realistic inputs, and produce outputs you can trust in production.

The next step is building your API connections to database systems, as well as third-party data sources such as:

- Web search providers for adverse media search agents
- Identity verification services for loan applications
- Company registries for KYB checks

Finding the right data mix will help you manage project costs, as there is often a balance to be established between cost and decision accuracy.

5. Identify a strong internal champion, or nominate yourself

In every AI transformation we've managed or participated in, we've noticed a common trend: the most successful projects are led by an internal champion who refuses to get sidetracked by negative opinions. Every organization will have certain stakeholders who lean toward risk mitigation over innovation. An internal champion will often guide these voices to acknowledge a harder truth: there is always risk in any process, even without AI.

Manual reviews can be inconsistent and hard to audit. People make mistakes, apply policies differently, and can't always explain why they made a certain judgment after the fact.

The job of an internal champion isn't to sell AI as risk-free. It is to address concerns proactively and make sure governance is built in from day one. We recommend starting with narrow a scope, operating shadow runs to build confidence, and raising automation levels as you gather hard data on agent performance.

Conclusion:

The new agentic model for financial services' highest-stakes decisions

As our understanding of AI evolves beyond theory, a tangible map of the agentic financial institution begins to emerge.

This is an institution where talented credit, compliance, and financial crime teams can automate repetitive tasks and focus their efforts on higher value work. A reality where businesses no longer have to wait for weeks to access life-changing capital. Where fighting financial crime is proactive and data-informed at every step.

The future imagined by leaders at the intersection of AI and financial services is no longer a future state. It is an imminent new operating model that will redefine how banks, fintechs, and insurance companies generate profit, manage risk, and deliver value to their customers.