March 9, 2026

via Electronic Mail

Alicia Chambers
Executive Secretariat
National Institute for Standards and Technology
U.S. Department of Commerce
100 Bureau Drive
Gaithersburg, MD 20899

RE: Request for Information Regarding Security Considerations for Artificial Intelligence Agents

Dear Ms. Chambers:

The Independent Community Bankers of America (ICBA) [1] appreciates the opportunity to respond to the National Institute of Standards and Technology (NIST) Center for AI Standards and Innovation's Request for Information (RFI) regarding security considerations for artificial intelligence (AI) agent systems.

As the only national trade association dedicated exclusively to serving the interests of community banks, ICBA represents thousands of community banks of varying asset sizes across the United States. Our member banks play a critical role in supporting small businesses, local communities, and rural economies, and as such, we recognize both the opportunities and challenges AI presents within the financial services sector. Community banks are increasingly evaluating AI-enabled capabilities for fraud detection, customer service, operational efficiency, cybersecurity operations, and third-party risk management. As AI capabilities evolve from informational systems to agent systems capable of taking actions that affect outputs, the security expectations, control requirements, and risk ownership questions become substantially more complex—particularly for highly regulated critical infrastructure sectors such as financial services.

ICBA supports NIST's leadership in developing voluntary, practical guidance that enables secure innovation while recognizing real-world constraints for smaller institutions and their technology providers. We recommend NIST's future guidance on AI agent systems reflect the realities of community banking, where even small operational errors can quickly become customer harm, financial loss, compliance issues, or safety-and-soundness concerns. Agent systems should be treated as a

[1] *The Independent Community Bankers of America ® has one mission: the create and promote an environment where community banks flourish. We power the potential of the nation's community banks through effective advocacy, education, and innovation. As local and trusted sources of credit, America's community banks leverage their relationship-based business model and innovation offerings to channel deposits into the neighborhoods they serve, creating jobs, fostering economic prosperity, and fueling their customers' financial goals and dreams. For more information, visit ICBA's website at icba.org.*

866-843-4222
icba.org
1615 L Street NW
Suite 900
Washington, DC 20036
518 Lincoln Road
P.O. Box 267
Sauk Centre, MN 56378

higher-risk category of AI because they can take actions—not just generate text—that impact sensitive data, customer communications, and business operations. NIST should encourage strong identity orchestration to support a "Know Your Agent" (KYA) approach, so institutions can identify which agent is acting, what it is authorized to do, and who approved that authority. Guidance should also emphasize constrained deployments by default in high-risk environments, clear accountability across the vendor ecosystem, and auditability that enables banks to investigate incidents, protect customers, and demonstrate appropriate oversight.

1. Security threats, risks, and vulnerabilities affecting AI agent systems

(a) What are the unique security threats, risks, or vulnerabilities currently affecting AI agent systems, distinct from those affecting traditional software systems?

From a community bank perspective, the key distinction is simple: AI agents do not just advise, they can do. That ability to take action introduces risk that looks less like a typical technology feature and more like operational risk and third-party risk. In a banking environment, a wrong step can affect customers quickly, disrupt critical workflows, and create compliance concerns even when no funds move.

One of the most significant risks is unintended action that creates real-world harm. An agent could send a message to the wrong customer, mis-handle sensitive information, change a record, or trigger a workflow that is difficult to reverse. Even seemingly small errors can drive complaints, reputational damage, and increased workload for already lean teams. These risks grow when agents are embedded into day-to-day systems where speed and scale are the point of adoption.

Community banks are also concerned about new forms of manipulation that resemble social engineering but aimed at systems rather than people. Agents can be influenced by what they read in emails, documents, tickets, or online content. A malicious or misleading message can steer the agent toward unsafe actions. In practical terms, this can look like an attacker tricking an agent into sharing information, escalating access, or taking an action that appears legitimate to the surrounding systems.

Finally, agent systems can create governance challenges when access and accountability are not clear. If an agent has broad access to bank systems or vendor tools, a small mistake, or a compromise, can have outsized impact. When something goes wrong, the institution needs to answer basic questions quickly: what happened, who authorized it, what changed, and what customers or systems are affected. Agent systems can make these questions harder to answer if actions are not clearly attributable and consistently logged—especially when capabilities are delivered through vendors and layered across multiple technologies.

(c) To what extent are security threats, risks, or vulnerabilities affecting AI agent systems creating barriers to wider adoption or use of AI agent systems?

Many AI systems used in financial services are primarily advisory: they help summarize information, flag suspicious activity, or support internal decision-making. These tools can still create meaningful risk, particularly around privacy, data governance, and accuracy, but the harm is often indirect and can be reduced through review before action is taken.

AI agents raise the stakes because they can execute steps within workflows and across connected tools. In a typical community bank environment, operations already rely on interconnected systems—core processing, customer relationship management, email, ticketing, fraud tooling, cloud services, and vendor portals. Agent systems are often designed to span these systems. That means a mistake or manipulation can spread faster and farther than in more bounded AI use cases.

Banking also imposes a higher bar for accountability and documentation than many other sectors. For community banks, it is not enough to know that an AI system generated a response; institutions must be able to demonstrate appropriate oversight and retain records that support auditing, customer protection, and incident response. This is why ICBA encourages NIST to frame agent security in regulated environments around KYA. Agents should be treated like operational actors that require clear identity, clear permissions, and clear records, similar to how banks manage users, service accounts, and third parties.

(d) How have these threats, risks, or vulnerabilities changed over time? How are they likely to evolve in the future?

The threat landscape is moving rapidly as AI capabilities shift from chat-based assistance to action-oriented systems. Early generative AI deployments focused heavily on content risks, such as inaccurate responses, privacy leakage, or harmful outputs. Agent systems introduce a new category of risk: wrong outcomes created by real actions. In banking, this shift matters because consequences can be immediate, measurable, and tied to customer trust and regulatory expectations.

At the same time, agents are becoming more connected to enterprise workflows and vendor platforms. As integration expands, so does the potential blast radius. The more tools an agent can access, and the more permissions it holds, the greater the operational impact of a poor decision, a misconfiguration, or manipulation through untrusted inputs. Community banks are particularly sensitive to this dynamic because many will adopt agent capabilities through core providers and other third parties. As agent stacks become more complex, clarity about responsibilities for controls, monitoring, incident response, and updates becomes essential.

Looking ahead, ICBA expects increasing autonomy and speed to be the main drivers of both value and risk. Institutions will face pressure to adopt automation for efficiency and customer experience, but without strong guardrails that pressure may outpace risk management. For community banks, scalable guidance must emphasize practical measures that can be implemented through governance, vendor management, and oversight. NIST's future work should encourage identity orchestration and KYA controls, constrained permissions by default in high-risk contexts, clear accountability across the vendor ecosystem, and auditability sufficient to investigate incidents and protect customers.

2. Security practices for AI agent systems

(a) What technical controls, processes, and other practices could ensure or improve the security of AI agent systems in development and deployment? What is the maturity of these methods in research and in practice?

ICBA encourages NIST to focus guidance on practical controls that community banks can implement through governance, vendor management, and operational processes, without assuming a large internal AI security team. In banking, security practices for AI agents should prioritize preventing customer harm and operational disruption by ensuring that agent systems are tightly governed, carefully limited in what they can do, and continuously supervised in ways that align with safety-and-soundness expectations.

First, future guidance should emphasize clear KYA practices enabled by identity orchestration. In community banking, the ability to identify which agent is acting, what it is authorized to do, and who approved that authority is foundational. Agents should be treated like operational actors - similar to user accounts and service accounts - so permissions can be limited, actions can be attributed, and accountability is clear when something goes wrong. This includes requiring that banks and vendors can maintain clear records of agent identity (including versioning), authorization boundaries, and a complete audit trail of consequential actions.

Second, NIST guidance should encourage "constrained-by-default" deployment practices for high-risk environments like financial services. Community banks should be able to adopt agent technology in a staged way, starting with low-risk tasks and expanding capabilities once controls and oversight are proven. Practical examples include limiting which systems an agent can access, limiting which actions it can take, setting transaction thresholds and guardrails, and requiring human approval for high-impact actions. In ICBA's view, these are not advanced technical requirements; they are the digital equivalent of long-standing banking controls such as approvals, segregation of duties, and transaction limits.

Third, NIST should elevate vendor accountability and third-party risk management as a core part of agent security practices. Community banks will often receive agent capabilities through core processors, fintech partners, and managed service providers. When the "agent" is delivered through a layered ecosystem, responsibilities can become unclear, particularly for monitoring, updates, incident response, and audit support. NIST guidance would be especially valuable if it helps standardize what community banks should expect vendors to provide (e.g., clear documentation of what the agent can do, how permissions are managed, what logs exist, and how incidents will be handled).

Finally, ICBA recommends that security practices explicitly address oversight and operational readiness. Community banks need confidence that they can detect abnormal behavior, stop it quickly, and investigate it afterward. Guidance should emphasize routine testing before deployment - including testing for manipulation through emails, documents, and web content - monitoring of agent actions in production, and incident response playbooks tailored to agent behavior. These expectations align with the banking sector's emphasis on strong controls, change management, and documented oversight.

(c) How might technical controls, processes, and other practices need to change, in response to the likely future evolution of AI agent system capabilities or of the threats, risks, or vulnerabilities facing them?

ICBA expects that as agent systems become more capable, the primary risk shift will be from "a system that helps staff complete work" to "a system that completes work on behalf of the institution." As that evolution continues, security practices will need to shift from one-time pre-deployment checks toward continuous governance because capabilities, integrations, and risks may change rapidly through model updates, new tools, or expanded access to bank workflows. NIST's RFI correctly anticipates that

practices must change as agent capabilities and threats evolve, and the community banking sector will need guidance that supports that lifecycle view.

From the community banking lens, the most important evolution will be the move toward dynamic control rather than static permissioning. As agents become more widely integrated, it will not be sufficient to grant broad access and rely on policy statements. Instead, agent authority should become more context-driven—tightening automatically for higher-risk actions (e.g., anything affecting customer accounts, sensitive data, or outward-facing communications) and requiring higher levels of review when risk signals increase. In plain terms, as agents get more powerful, banks will need stronger stop signs, clearer escalation paths, and more automated ways to restrict what an agent can do when conditions change.

(e) Which cybersecurity guidelines, frameworks, and best practices are most relevant to the security of AI agent systems?

ICBA recommends that NIST's future guidance for AI agent systems be grounded in the supervisory risk-management frameworks that community banks already use to manage safety-and-soundness risk. Community banks are unlikely to treat AI agents as a standalone new risk category; instead, they will manage agent-enabled capabilities through established governance practices for model risk and third-party risk. Anchoring NIST guidance to these existing approaches will make the guidance more actionable, more scalable, and easier to incorporate into routine risk oversight without creating unnecessary burden.

In particular, NIST should recognize that many of the most important security questions for AI agents in banking align closely with well-established model risk management principles. The Federal Reserve's SR 11-7 guidance emphasizes risk-based governance and oversight that is appropriate to an institution's size, nature, complexity, and extent of model use.[2] For community banks adopting agent systems—often through third parties—these principles translate into clear expectations around governance, appropriate controls, and the ability to understand and manage model-driven behavior in the context of the bank's operations and risk exposure.

NIST should also reference the OCC's recent Model Risk Management: Clarification for Community Banks bulletin, which highlights that community banks should tailor model risk management practices to be commensurate with the bank's risk exposures, business activities, and the complexity and extent of model use.[3] This approach is particularly relevant for AI agent systems because institutions may adopt agent functionality in stages, starting with lower-risk uses and expanding only when controls, oversight, and outcomes support broader use. NIST guidance that mirrors this approach commensurate with risk principle will help community banks implement agent security expectations in a practical way.

Finally, because community banks commonly obtain agent capabilities through vendors (i.e., core processors, fintech partners, cloud providers, and managed service providers), NIST's guidance should align closely with the banking agencies' Interagency Guidance on Third-Party Relationships: Risk Management, which sets out sound risk management principles across the full third-party relationship

[2] Bd. of Governors of the Fed. Reserve Sys., *SR 11-7, Guidance on Model Risk Management* (Apr. 4, 2011).
[3] Office of the Comptroller of the Currency, OCC Bulletin 2025-26, *Model Risk Management: Clarification for Community Banks* (Oct. 6, 2025).

life cycle.[4] Agent security in community banking will often be achieved through a combination of internal governance and third-party oversight. NIST can support secure adoption by encouraging clear accountability between banks and vendors, and by recommending standardized information that third parties should provide to enable bank oversight (e.g., what the agent is designed to do, what it can access, how changes are managed, what monitoring exists, and what incident support is available) in a way that fits within established third-party risk management programs.

3. Assessing the security of AI agent systems

(a) What methods could be used during AI agent systems development to anticipate, identify, and assess security threats, risks, or vulnerabilities?

ICBA recommends that NIST emphasize assessment approaches that are practical for community banks and scalable across third-party provider environments. In banking, security assessments should not be limited to whether an agent produces accurate or safe responses, it must also address whether the agent can be safely deployed in operational workflows without creating unacceptable risk to customers, sensitive data, or the bank's control environment. Community banks need assessment methods that help them answer straightforward questions: What can this agent do? What can it access? What could go wrong? How would we know? And how do we stop it?

From the community banking lens, effective assessment starts with a clear understanding of the agent's intended use and boundaries. Before deployment, banks and vendors should be able to document the agent's purpose, the systems and data it can interact with, the actions it can take, and the conditions under which those actions are permitted. This is where KYA framing becomes important. Assessments should confirm that the agent has a clear identity, a clear permission model, and an auditable record of authorization.

ICBA also encourages NIST to emphasize testing that reflects how community banks will actually encounter risk in practice. In banking operations, many of the most meaningful security failures will not look like traditional hacking events, they may look like an agent being manipulated by untrusted inputs (i.e., emails, documents, or web content) or taking an unintended action based on incomplete information. Accordingly, assessment methods should include scenario-based testing for misleading instructions, manipulated content, and workflows where the agent is asked to complete multi-step tasks. A practical test should confirm the system behaves appropriately when it encounters ambiguity, conflicting instructions, or suspicious inputs, and that it reliably pauses or escalates instead of pushing through to action.

Finally, ICBA recommends that NIST's assessment guidance focus on outcomes that matter to regulated institutions: whether the agent's actions are attributable, whether activity can be monitored in a meaningful way, and whether the institution can quickly contain or reverse harm. Community banks must be able to detect anomalous activity, halt the agent when necessary, investigate incidents, and provide documentation suitable for compliance reviews and customer remediation. Assessment,

[4] *Interagency Guidance on Third-Party Relationships: Risk Management*, 88 Fed. Reg. 37,920 (June 9, 2023).

therefore, should include evaluation of monitoring, alerting, audit logs, and the institution's ability to disable agent functions quickly across integrated systems.

(b) Not all security threats, risks, or vulnerabilities are necessarily applicable to every AI agent system; how could the security of a particular AI agent system be assessed and what types of information could help with that assessment?

ICBA believes the most significant assessment challenges for community banks stem from how agent systems are built, delivered, and updated, often through third parties, and how quickly their capabilities can change. Community banks frequently do not deploy standalone agent systems in a controlled lab environment. Instead, they receive agent capabilities as features embedded within vendor platforms, where the bank may have limited visibility into the underlying model, tooling, or orchestration logic. This can make it difficult for community banks to independently validate security claims or understand how the system will behave across real-world workflows.

A second challenge is that agent risks are often driven by context and connectivity, not by a single technical flaw. The security of an agent depends heavily on what it can access, what tools it can use, and how broadly it is integrated into bank operations. Two banks could deploy the same agent technology but face very different risk depending on permissions, connected systems, transaction authority, and governance. This makes one-size-fits-all testing less useful and increases the importance of clear scoping and boundaries, especially for higher-risk functions like customer communications, account servicing workflows, fraud operations, or privileged IT actions.

Third, assessing agents is challenging because harmful outcomes may resemble legitimate activity. An agent that takes an unsafe action may still appear to be working as designed if logs only show an approved API call or automated workflow step. Community banks need assessment approaches that focus on traceability and auditability. For example, the ability to reconstruct what the agent was trying to do, what inputs it relied on, what permissions applied at the time, and what changed as a result. Without this, incident investigation and regulatory examination become difficult, especially for institutions with small teams and limited capacity to interpret complex technical evidence.

Finally, the pace of change introduces a practical governance problem. Agent systems may evolve through frequent updates to models, tool integrations, and vendor features. A security assessment performed at one point in time can become stale if updates expand capabilities, alter behavior, or change what the agent can access. For community banks, this makes change management and third-party transparency essential. NIST guidance can be particularly helpful if it encourages clear vendor practices around change notifications, documentation updates, and mechanisms to pause, roll back, or constrain agent functionality when risk concerns arise.

In addition, ICBA expects that agent security practices must evolve to handle growing dependency on third parties and rapid product iteration. Community banks may not control model updates, tool integrations, or vendor feature releases, yet they remain accountable for managing operational and compliance risk. NIST guidance should encourage practices that ensure updates do not silently expand agent authority or change behavior without visibility, such as strong change management expectations, clear vendor notification standards, and easy-to-use rollback or disablement capabilities when problems arise.

4. Limiting, Modifying, and Monitoring Deployment Environments

(*a) AI agent systems may be deployed in a variety of environments, i.e., locations where the system's actions take place. In what manner and by what technical means could the access to or extent of an AI agent system's deployment environment be constrained?*

From a community banking perspective, the most practical approach to AI agents is to constrain the deployment environment by default, both in terms of where the agent is allowed to operate and the scope of actions it can take. Community banks operate in a highly regulated environment where systems are interconnected and where customer trust, safety-and-soundness, and operational resilience are critical. NIST guidance should therefore encourage bounded deployments that limit the agent's reach to only what is necessary for a specific business purpose, and that prevent agents from freely navigating across tools, systems, or data sources.

In practical terms, community banks will constrain agent deployment primarily through governance and access controls that are familiar to banking leaders. For example, limiting access to systems and data based on least privilege, limiting actions to predefined workflows, and requiring additional approval for higher-risk steps. KYA should be foundational here - the agent should have a clear identity, clearly documented authority, and clearly documented boundaries. Constraining the deployment environment should include limitations on which internal systems the agent can connect to, what external resources it can interact with, and what types of actions it is permitted to initiate. This is the digital equivalent of strong internal controls, making sure the agent cannot do more simply because it is technically capable.

ICBA also encourages NIST to acknowledge the vendor reality for community banks. Many banks will rely on core providers and third-party platforms to implement these constraints. NIST guidance will be especially helpful if it identifies common ways that banks can require bounded deployment through contracts and oversight, such as requiring vendors to provide configurable limits on tools and actions, clear documentation of the agent's operating boundaries, and straightforward mechanisms to restrict or disable specific functions when risk concerns arise.

(b) How could virtual or physical environments be modified to mitigate security threats, risks, or vulnerabilities affecting AI agent systems? What is the state of applied use in implementing undoes, rollbacks, or negations for unwanted actions or trajectories (sequences of actions) of a deployed AI agent system?

ICBA recommends that NIST emphasize deployment designs that make agent-driven actions easier to correct, contain, and reverse. In banking, mistakes are not hypothetical—they must be handled through well established processes for error resolution, customer remediation, record correction, and incident response. Agent systems should be deployed in ways that support these processes.

In practice, modifying the deployment environment to mitigate risk often means creating safe lanes for agent activity. Community banks are likely to prefer agent deployments where actions are staged, confirmed, and recorded before becoming final, particularly in higher-risk contexts such as customer communications, account servicing workflows, fraud operations, and privileged IT activities. For example, rather than allowing an agent to finalize an external action automatically, the environment can be designed so the agent prepares a recommended action and routes it for review when predefined risk thresholds are met. This approach is consistent with traditional banking controls like approvals, segregation of duties, and transaction limits.

With respect to undo, rollback, or negation, ICBA believes these concepts should be treated as essential operational safeguards for agent systems. The most mature rollback practices in community banking are likely to exist where they align with existing operational controls—such as the ability to reverse a workflow step, correct an internal record, retract or correct a customer communication, disable an automation, or restore a prior configuration state. Community banks will benefit from NIST guidance that encourages agents to operate in environments where such reversals are feasible and clearly defined, and where incident response teams can quickly contain harm by pausing the agent, restricting permissions, or reverting changes. While not every action will be perfectly reversible, guidance should encourage design choices that reduce irreversibility and support rapid containment and recovery.

(d) What methods could be used to monitor deployment environments for security threats, risks, or vulnerabilities?

ICBA recommends that NIST highlight monitoring practices that support early detection, rapid containment, and effective investigation of agent-related incidents, without assuming that community banks can build specialized monitoring capabilities from scratch. Community banks already invest in monitoring for fraud, cybersecurity threats, and operational disruptions, often with the support of managed service providers. Agent systems should fit into this existing monitoring ecosystem and not create blind spots where the institution cannot see what occurred or respond quickly.

For community banks, monitoring should focus on observable outcomes that matter, such as unusual activity, unusual access patterns, unexpected changes to records or configurations, abnormal volumes of transactions or communications, and deviations from expected workflows. Monitoring must also support accountability. Banks must be able to trace actions back to a specific agent identity and configuration, understand what the agent was attempting to do, and determine whether the action was authorized under the bank's policies and risk thresholds. In short, monitoring should help banks answer what happened? What changed? Who or what initiated it? And what should be done next?

ICBA also encourages NIST to recognize challenges in applying traditional monitoring to agent systems. When agents operate through legitimate tools and workflows, unsafe actions can appear similar to normal automation activity, making detection more difficult. Additionally, when agent capabilities are delivered through vendors, community banks may have limited direct visibility into the underlying system behavior. NIST guidance can help by encouraging clear vendor practices around logging, alerting, and incident support, and by recommending that banks have the ability to disable or constrain agent functions quickly when abnormal behavior is detected.

Finally, ICBA notes that monitoring must be implemented in a way that is consistent with legal and privacy expectations, particularly where monitoring could involve customer data, employee communications, or sensitive internal records. Community banks will need guidance that supports security and oversight while reinforcing the importance of appropriate data handling, access controls, and purpose limitation. Practical guardrails - such as limiting monitoring access to those with a need to know and ensuring monitoring data is retained and used appropriately - will help institutions deploy agent oversight responsibly.

5. Additional Considerations

(a) What methods, guidelines, resources, information, or tools would aid the AI ecosystem in the rapid adoption of security practices affecting AI agent systems and promoting the ecosystem of AI agent system security innovation?

From a community banking perspective, the fastest way to drive secure adoption is to give banks and their vendors a practical starter kit that fits into existing governance, third-party oversight, and examination expectations. Community banks do not need a separate, technical security program solely for AI agents; they need clear guidance they can operationalize through policies, vendor contracts, and routine control testing. NIST can accelerate adoption by publishing a short set of implementation-oriented resources that translate agent security into familiar banking control concepts, including sample governance language, staged deployment approaches - starting with low-risk uses - and clear examples of what bounded agent deployments look like in practice.

ICBA also recommends that NIST encourage standardized documentation and assurance artifacts that vendors can provide to customers. Many community banks will consume agent capabilities through third-party platforms (e.g. core processors, fintech partners, managed service providers). A common barrier to adoption is not willingness, but visibility—banks need a consistent way to understand what the agent can do, what it can access, what controls exist, what logs are available, how changes are handled, and what incident support will be provided. A standardized agent assurance packet, designed to plug into existing third-party risk management, would materially reduce friction and improve security outcomes.

Finally, NIST could promote security innovation by publishing shared testing resources that are accessible to smaller institutions and vendors, such as scenario libraries for agent misuse, sample risk assessment templates, and plain-language guidance for implementing KYA identity, permissions, and audit trails. These tools would help normalize baseline expectations across the ecosystem and reduce uneven security practices across institutions of different sizes.

(b) In which policy or practice areas is government collaboration with the AI ecosystem most urgent or most likely to lead to improvements in the state of security of AI agent systems today and into the future?

ICBA believes government collaboration is most urgent where it can reduce fragmentation and create a common baseline across sectors, especially for critical infrastructure like financial services. Community banks operate under supervisory expectations that emphasize governance, risk-based controls, vendor oversight, and resilience. NIST can help most by convening government and industry to align on practical, scalable expectations for agent systems that reinforce existing risk management approaches, rather than adding parallel requirements that are difficult for smaller institutions to implement.

Collaboration is also urgently needed on the shared problem of third-party dependency. As agent capabilities are embedded into vendor platforms, security outcomes increasingly depend on transparency, change management, and incident support across multiple providers. Government can help by encouraging common practices for vendor disclosures, update notifications, and shared responsibility for monitoring and incident response. Community banks benefit when oversight expectations are clear and consistent across the ecosystem, so institutions can set enforceable requirements through procurement and contracts, and vendors can design to a known standard.

Finally, ICBA encourages continued public-private collaboration on threat information sharing and response readiness for agent-related incidents. As agent systems become more widely adopted, community banks will need timely insights into emerging patterns of misuse - including manipulation through untrusted content - and practical response playbooks that support containment, investigation, and remediation.

(c) In which critical areas should research be focused to improve the current state of security practices affecting AI agent systems?

From the community banking lens, the highest-value research is research that improves the reliability and controllability of agent actions in real operational settings. Banks need stronger methods to prevent agents from being steered into unsafe behavior by misleading or malicious inputs, particularly when agents consume emails, documents, or web content as part of a workflow. Research that produces practical mitigations, especially those that can be embedded into vendor platforms, would directly reduce operational risk and customer harm.

ICBA also sees a pressing need for research and development on KYA controls that are both robust and usable: strong identity binding for agents, permissioning models that reflect real-world workflows, and mechanisms for step-up approval for high-impact actions. In banking, secure adoption depends on being able to show that actions were authorized, attributable, and auditable. Research that improves auditability—without requiring banks to retain overly sensitive data or introduce privacy risk—would materially strengthen incident response and oversight.

Finally, ICBA recommends prioritizing applied research on containment and recovery, including practical approaches to undo, rollback, or otherwise mitigate unwanted sequences of actions. Banks are accustomed to operating with controls that prevent, detect, and correct errors. Agent systems should be deployed in a way that supports rapid containment and structured remediation. Research that improves reversibility, safe staging of actions, and real-time detection of abnormal agent behavior will unlock broader adoption while protecting customers and maintaining operational resilience.

In closing, ICBA appreciates NIST's leadership in convening stakeholders on security considerations for AI agent systems and looks forward to continued engagement as NIST develops practical guidance. We welcome the opportunity to discuss these recommendations in greater detail, and to serve as a resource on the unique operational realities of community banks, including the safety-and-soundness, customer protection, and third-party risk considerations that shape AI adoption. Please contact Anjelica Dortch, Vice President, Operational Risk & Cybersecurity Policy, at Anjelica.Dortch@icba.org with any questions or to arrange a discussion.

Sincerely,

/s/

Anjelica Dortch
Vice President, Operational Risk & Cybersecurity Policy