Agentic AI workflows in financial reporting

Risk, governance, and control considerations

Agentic AI workflows transform the risk landscape and expand governance complexity

Have you considered how your entity's risk profile changes when using AI agents – and how it further evolves as those agents are orchestrated into agentic AI workflows? AI agents introduce new risks when they perform discrete tasks, and those risks increase as agents orchestrate steps in agentic AI workflows. This *increased autonomy elevates governance, control, and accountability considerations for financial reporting.*

Understanding the difference between agents and agentic AI workflows

While agentic AI workflows leverage individual AI agents to perform discrete tasks, the orchestration of those agents into an end-to-end workflow introduces expanded risks. Understanding which types of agents are deployed—and how they are orchestrated—is foundational to designing appropriate governance, controls, and oversight.

Topic	Agent	Agentic AI workflow
Capability	Performs discrete tasks. A single system that acts on input to achieve a specific goal by executing tasks.	An AI system architecture used to orchestrate complex, multi-step tasks in a process via task decomposition and execution, often leveraging multiple specialized agents.
Autonomy	Operates independently to complete a defined task.	Operates independently to pursue a high-level goal, capable of adapting its strategy, altering its course of action and orchestrating its own resources.
Components	An AI model (the 'brain'), memory, and a set of tools.	A collection of one or more agents (each with its own 'brain'), a suite of tools, and a shared memory or context, all orchestrated by a central AI model (i.e. an orchestrator agent).
Example	An agent performs a 3-way match using multiple formats of information guided by business rules provided as instructions from the user or orchestrator agent.	An orchestrator agent delegates tasks like 3-way matching, discrepancy research, and vendor communication to specialized agents, then routes the resolved issue for final approval. It is guided by business rules defined within the workflow's configuration and system integrations.

Vendor-provided ERP systems

As vendors embed agents into their systems, entities will need to actively scrutinize how vendors govern the related risks. Traditional safeguards, including standard SOC 1 reports, may not address all risks.

Enabling these agents goes beyond toggling a feature; it requires a thoughtful assessment of risk and clear establishment of responsibility for managing that risk. Entities may have limited ability to rely on vendor-provided controls to address reliability of AI generated outcomes. At the same time, key AI governance and security considerations over the model may not be visible or controllable by the entity. Entities will need to carefully evaluate whether the control objectives and related controls in SOC 1 reports adequately address agent-specific risks.

Novel risks

The deployment of agents – even for discrete tasks – introduces new risk considerations. Further, the risk landscape broadens as agentic AI workflows become increasingly autonomous by operating across systems and processes with limited human intervention, adapting decisions based on outcomes, and coordinating actions between agents.

Entities need to proactively identify and assess the new risks introduced by agents and agentic AI workflows, particularly those arising from increased autonomy and coordination across processes. Five potential novel risks are identified below.

Ungoverned agents

The decentralized creation of agents across the entity could complicate governance and make it difficult to establish a complete inventory of agent use with clear accountability for each.

Goal misalignment

By prioritizing the efficient completion of their targeted tasks, agents may exploit loopholes, miss the broader business context, and ultimately diverge from key financial reporting objectives – potentially compromising the integrity of financial data or overriding controls.

Cascading effects that impact reliability

Agents could rationalize and compound issues. The speed and coordination of agents can amplify errors, allowing issues originating in one task to cascade into other portions of the financial reporting process. As agents prioritize task completion, they may rationalize another agent's inaccurate output or unknowingly rely on a flawed output because validating other agent's work may be outside their task. Additionally, agents could lose critical knowledge through continuous learning cycles. This could affect reliability of results in financial reporting.

Systemic risks

Errors from a single agent could propagate across other interconnected IT systems, and agentic AI workflows could make issue tracking and resolution difficult. Using shared service accounts can further weaken the audit trail. As agents learn and adapt, security measures may have to be continuously adjusted.

SOD conflicts

Agentic AI workflows could compromise segregation of duties policies by collapsing traditionally separate roles into a single orchestrated process. This risk is amplified when agents use a shared model. Unlike a team of people, where diverse perspectives provide natural checks and balances, a single flaw in the shared 'brain' can instantly affect or bias all agents.

Managing Agentic AI workflow risks

Intentional **governance** across the full workflow, including clear ownership of risks, rigorous testing, defined points of human accountability, effective **controls**, and continuous monitoring helps to mitigate risks across agentic AI workflows. Treating agents as privileged system actors – with least-privilege access, traceable actions, and the ability for humans to intervene or shut down agent behavior – is critical to maintaining accountability, security and auditability.

Example workflow design considerations

Strategic orchestration: Define agent roles, sequencing, and permitted tools/APIs; prevent unapproved agent-to-agent behaviors.

Human-centric oversight: Embed approval points and exception routing; establish accountable owners to validate critical judgments.

Governance and accountability: Establish a central inventory/portfolio of agents that identifies each agent's ownership, risk classification, and policies for development and use.

Integrated security and compliance: Apply least-privilege access, identity controls, and continuous threat monitoring; align with regulatory and privacy requirements.

Rigorous testing and validation: Test expected and edge-case scenarios; validate outputs against known-good baselines; assess multi-agent interaction risk.

Adaptive monitoring and continuous improvement: Track performance, model drift, and anomalies; implement circuit-breakers/tripwires; retrain and refine based on observed behavior.

Key questions to guide governance and control considerations within your entity

Proactively engaging with vendor or service providers and internal stakeholders becomes critical to understanding how agents and agentic AI workflows are used and governed in practice. Early and ongoing dialogue enables an entity to identify potential gaps in controls over deployment, use, and monitoring and to avoid new or unmanaged risks that are not clearly owned, visible, or controlled within the financial reporting environment.

We outline several questions below that management can use before or during agent deployment to assess readiness, clarify ownership, and identify potential control gaps.

Governance

- Who has overall responsibility for the governance of agents used in financial reporting?
- Who can create, deploy, or modify agents, and how is that authority approved and monitored?
- Is there an inventory maintained of the solutions that allow agent creation?
- Who is responsible for understanding AI or agent use by vendors or service organizations?

Design of agents

- If an agent is granted privileged access, is that access periodically reviewed for ongoing appropriateness?
- What is the process for customizing and training agents on the organization's unique processes and data?
- What is the process for testing and validating the logic and output of AI agents before they are deployed?
- How are new models and agent updates reviewed, approved, and deployed?
- How is disruption to the financial reporting controls prevented when new models are deployed?
- To what extent, if any, are logs of key actions taken by AI agents available? How are agent actions retained and reviewed to support auditability and investigation?
- How is exception handling and performance monitoring handled?

Agentic AI workflows

- Are there clear boundaries for accountability, controls, and monitoring when multiple agents operate together as part of a process?
- Have potential SOD conflicts been evaluated before deploying an agentic AI workflow?
- Are agent-to-agent actions programmed to follow a logical sequencing or are decision paths orchestrated by an agent?
- How are escalation paths, human-in-the-loop approvals, and intervention points defined within the agentic AI workflow?
- Does each agent leverage a single, consistent model, or do they use different models? Further, does the orchestrator agent have the autonomy to select the most appropriate model for a given agent task?
- When an agentic AI workflow uses multiple models or agents that do not track changes made over time, what is the process for identifying and validating shifts in an agent's behavior, whether due to new training data or changes in orchestration and routing logic?

Control considerations, including where and when human review occurs, that address explainability, transparency, reliability and other AI risks are equally applicable to AI agents and agentic AI workflows. Refer to the AI and automation in financial reporting guide for more information.

Key questions to guide governance and control considerations within your entity (cont.)

These questions can be used by management to assess readiness, clarify ownership, and identify potential control gaps when using vendor-provided agents or service providers that offer agents or leverage agents in their processes.

Vendor or Service Providers

- How are agent identities created and authenticated?
- Are agents assigned specific user IDs with roles and/or permissions, or do they run with a shared service account?
- How do agents integrate into existing ERP models and what data governance is in place to determine proprietary data remains secure while being used by agents?
- When using a service provider for financial reporting or ICFR:
  - Have all relevant risks been mapped to control objectives in the SOC 1 report to identify potential unmanaged risks and responsibilities of the user entity?
  - Are AI agent-specific control objectives and related controls explicitly articulated in the SOC 1 report, or are they incorporated into the broader ERP system's control objectives?
  - If agent-related controls are embedded within broader ERP control objectives, how does the entity assess whether those controls adequately address agent-specific risks?

Considerations around vendor AI certifications and assurance

AI Certifications ≠ Assurance over AI risks

Vendor AI certifications may:

- Provide useful risk assessment to user entities
- Represent a third-party certification of aspects of governance and risk assessment
- Assist in vendor AI evaluation for governance requirements

Vendor AI certifications generally do not:

- Directly address process risk points or risks arising from IT relevant to ICFR

Items to evaluate when considering certifications in an entity's risk assessment or governance:

- Nature and scope
- Reputation of the framework and certifying party
- Certification type (e.g. examination under AICPA standards vs. ISO compliance)
- Any limitations included in the report

Assurance

SOC reports:

- SOC 1 reports are most common for addressing process control activities and GITCs relevant for ICFR. As AI is embedded into processes, mitigating process level risks may increasingly depend on complementary user entity controls (CUECs) executed by humans
- SOC 2 reports address security, availability, processing integrity, confidentiality or privacy, which may support certain GITCs and governance

Considerations when using SOC 1 reports:

- Verify that AI risks, including those around information, are explicitly addressed within control objectives in the report
- Consider whether sub-service organizations are significant and obtain those service organization reports
- Verify that CUECs are addressed at the entity

For further information

AI presents an incredible opportunity in today's rapidly evolving business landscape. Check out the firm's AI insights resource page, which includes resources related to responsible AI, AI events, and AI webcasts and replays.

Learn about us: **kpmg.com**