NITIC News Article
Blog Post Published 11.03.2026
https://www.nitic.org/auditing-the-hidden-supply-chain-behind-enterprise-ai-ag

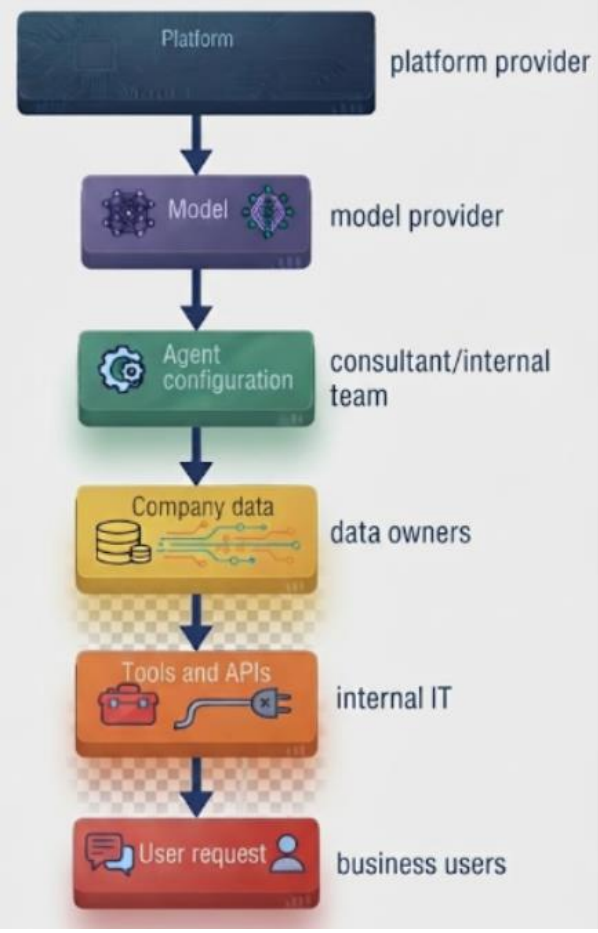
Auditing the Hidden Supply Chain Behind Enterprise AI Agents

Jiri Jirik
NITIC Tech Team Member
Moraine Valley Community College

When an enterprise AI agent makes a mistake, auditors first need to establish accountability. While most discussions about auditing AI agents focus on the model or the platform, much of the risk lies in the supply chain behind them.

An enterprise AI agent is rarely a single product. It sits on top of a supply chain of vendors, models, configurations, data sources, tools, and permissions, each possibly managed by a different party. That hidden supply chain is where many audit and governance risks exist.

What sits behind the agent



A useful way to think of an enterprise AI agent is as a stack of layers, each with its own owner, as shown in Figure 1. Each layer is configured, connected, and approved before an employee types a single request. The final interaction is simple. The journey to it is not.

An agent that summarizes public policies carries limited risk. An agent that can retrieve employee records, provision accounts, or communicate externally does not. The question is never if the agent sounds intelligent. The question is what data it can access, what tools it can invoke, and what actions it can trigger.

Shadow AI vs. AI sprawl

Shadow AI refers to employees or departments using AI tools outside approved processes, such as pasting internal material into a public generative AI service. ISACA notes that shadow AI can bypass existing IT, privacy, and compliance controls (Mathew, 2025; Rajasekharan, 2025) .

AI sprawl is broader: it can occur even when every platform has been formally approved. IBM describes agent sprawl as the uncontrolled growth of AI agents without a unified strategy and governance structure (IBM, 2026) . For example, one department deploys a knowledge agent, another uses a different platform for human resources (HR) support, and a consultant builds customer service agents. Each project may make sense on its own. Taken together, the organization may quickly lose track of what agents exist and who is accountable for them. Shadow AI is a question of approval. AI sprawl is a question of visibility and accountability.

Platform approval is not agent approval

An approved AI platform does not mean that every agent built on it has been reviewed. An agent that summarizes published policies and an agent that updates records in the enterprise resource planning (ERP) system are not equal and should not receive the same level of review.

The governing principle is simple: approve at the agent and use-case level, not only at the platform level. Before an agent enters production, the organization should be able to identify:

- The business purpose and the business, technical, and data owners.
- The platform, model, version, and any third parties involved, including suppliers of plugins and connectors.
- The data sources, tools, and application programming interfaces (APIs) the agent can access.
- The permissions the agent uses, and whether it acts under its own identity or a user's.
- The actions the agent may take, and where human approval is required.
- The monitoring, logging, and change-management processes.

Stored in a controlled inventory, this information works as an AI bill of materials (AIBOM) that traces the supply chain behind each agent.

The accountability gap

The hidden supply chain behind an AI agent spreads responsibility across many parties, yet compliance requires clear accountability. While each layer has its own owner the organization still owns the result. Consequently, when an agent gives incorrect information

or takes an unauthorized action, the organization must be able to determine what happened and why. That requires an end-to-end record of:

1. The initiating event and the agent identity, owner, version, model, and configuration.
2. The data retrieved and the tools, APIs, or other agents invoked.
3. The policies evaluated and the approval, override, or escalation decision.
4. The action taken and its result.

A generic chat transcript is not enough. Auditors need evidence that the agent stayed within its approved role.

What auditors should examine

NIST's AI Risk Management Framework treats AI risk as a lifecycle and system-context problem (National Institute of Standards and Technology, 2023) . Traditional IT controls still apply, but they must now account for dynamic agent behavior and chains of delegated authority.

A practical audit program should address six areas.

1. **Agent inventory and ownership**

Maintain an inventory of agents in development, pilot, and production, including consultant-built agents, embedded software-as-a-service (SaaS) copilots, and known shadow AI tools.

2. **Data connections and retrieval**

Auditors should confirm that the data owner approved each connection and that sensitive information is masked. Retrieved content also opens a new attack path: a document or email can carry hidden instructions that the agent may follow, a risk known as indirect prompt injection (OWASP Foundation, 2024) .

3. **Tools, APIs, and authority**

An agent's risk increases when it can take action. PwC recommends assigning each AI agent a unique digital identity, defining clear capability boundaries, and applying least privilege (de Boer, 2026) . Auditors should also determine whether each agent acts under its own identity or with the delegated permissions of the user who invoked it.

4. **Change management**

AI systems can change without a code release. A model update, new prompt, or changed permission can alter behavior. Auditors should verify that model versions are pinned or that updates trigger retesting, and that prompt changes follow the same approval process as code.

5. **Testing and evaluation**

Agent behavior is probabilistic, so a test passed at go-live may not hold after a change. Auditors should look for pre-deployment testing, adversarial testing (red-teaming), and retesting after material changes.

6. **Logging and continuous monitoring**

Logs should allow the organization to reconstruct the end-to-end record described earlier. That evidence supports both audit and incident response.

From AI sprawl to assured autonomy

Well-governed AI agents can produce stronger evidence than many manual processes.

However, that benefit is not automatic. Organizations need to treat agents as governed digital workers. Each one should have a defined purpose, a named owner, a unique identity, limited authority, and a clear process for retirement. That is what assured autonomy means: agents that act independently, within limits the organization can prove.

The central audit question should not be, “Did we approve an AI platform?” Instead, organizations should ask:

“Can we identify every agent acting on our behalf, trace the supply chain behind it, understand what it can access and do, and prove it remained within approved boundaries?”

Organizations that can answer that question are building a foundation for responsible AI adoption. Those that cannot may discover that AI sprawl has already become a security, compliance, and accountability issue.



This material Is based upon work supported by the U.S. National Science Foundation under Grant No. 2300188. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.