CYBERARK

An identity-centric approach to securing AI agents

A modern security strategy treats complex AI systems as privileged users with elements of both human and machine

Rahul Dubey | CyberArk

Identity security started decades ago with identity and access management and then moved into privileged access management. Along the way, the focus expanded from human identities to machine identities. Now with the introduction of AI agents, identity is beginning to play an even bigger role in human and machine authentication.

Identity is the gatekeeper for all the pillars of a zero trust architecture, which has been embraced at every level of government because of its effectiveness. When we think about bringing artificial intelligence into a network, we need to keep those zero trust principles in mind and consider risk management in terms of the external supply chain and the internal software development pipeline.

Buyers typically evaluate agentic AI technology in terms of speed and quality, but we can't overlook security. AI agents have access to an organization's most sensitive resources, which makes them privileged identities. And although they are machines, they share some of the same vulnerabilities as human users.

Determining proper authentication and access

Agentic AI programs help humans make decisions more quickly, but they also have reasoning and perception and are designed to proceed on their own with autonomous decision-making. Some complex AI agents manage multiple orchestrator agents and run complicated multi-task processes. When agencies deploy such complex AI systems, their risk level increases in terms of applications, data and access.

Therefore, AI agents must be properly authenticated and given an appropriate amount of access to network resources. That starts with deciding who defines an AI agent's level of access and how it is granted. In addition, it means deciding what applications the AI agent can use and for how long, just as we would restrict and monitor the access of a human user.

One of the most effective approaches is to give AI agents privileged rights temporarily and just-in-time for specific tasks, known as zero standing privileges. In other words, the AI agents will get only the access they need for a particular amount of time.

> “One of the most effective approaches is to give AI agents privileged rights temporarily and just-in-time for specific tasks, known as zero standing privileges.”

A new class of privileged machine identities

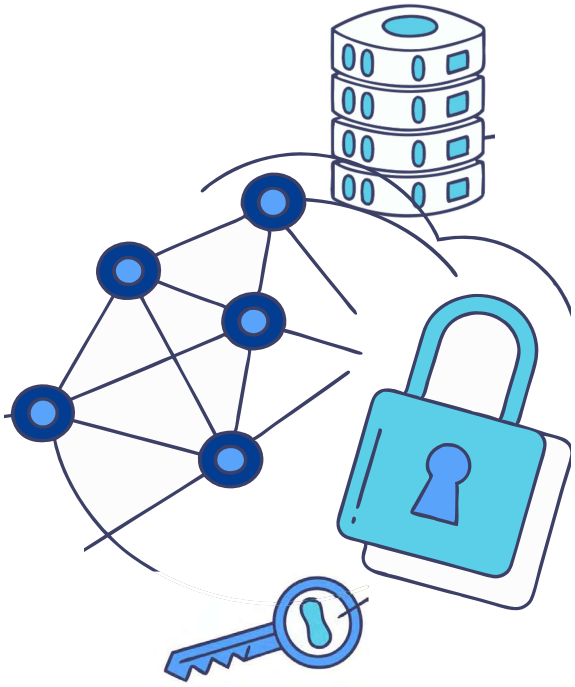
Knowing that AI agents are similar to privileged users, CyberArk recently introduced the Secure AI Agents solution, which allows us to apply our Identity Security Platform to AI agents. Together, those tools can enforce zero standing privileges and just-in-time access to minimize the attack surface.

The Secure AI Agents solution begins with discovery. It finds all the AI agents on a network—across cloud and developer environments—and gathers important context, such as who owns each agent, what it does and what it can access. Our AI Agent Gateway sits between AI agents and the tools they use to grant permissions only for a specific task and with the appropriate level of privileges. Those permissions are revoked automatically to enforce zero standing privileges for AI agents.

To strengthen authentication, CyberArk's technology does not store or hardcode credentials and instead supports secret rotation, which involves periodically changing static secrets such as passwords, and dynamic secrets, which involves generating unique credentials that automatically expire after use. Furthermore, the AI agents' privileged access activity is monitored and recorded, and it can be shared in our monitoring and logging tools as well.

That integration model leverages the power of our Identity Security Platform and treats AI agents as a new class of privileged machine identities. ■

Rahul Dubey *is vice president of global regulated markets solutions at CyberArk.*



ADOBE STOCK//HILMAWATI