

From AI Risk to Governed Action

How Runtime Controls Help Organizations Deploy Autonomous AI Agents without Expanding Cyber, Data, and Operational Risk

Thank you for your interest in exploring this content.

Carahsoft is the **Trusted Government IT Solutions Provider®** supporting a broad portfolio of industry-leading technologies through OH MSCA, NASPO ValuePoint, Texas DIR-CPO-5683 and a wide range of other contract vehicles.

As the **Master Government Aggregator®**, Carahsoft connects government agencies, industry partners, and technology providers to deliver innovative, mission-focused solutions.

In partnership with Broadcom, we provide technology solutions that drive modernization, strengthen operations, and ensure compliance with evolving government standards.



To learn more about how Carahsoft can support your technology needs, please visit carahsoft.com



Explore More Resources:
carah.io/BroadcomResources



Join Events & Webinars:
carah.io/BroadcomEvents



Discover Technology Solutions:
carah.io/BroadcomSolutions



Learn About Procurement:
carah.io/BroadcomContracts



Connect With Our Team:
Broadcom@carahsoft.com
(571) 662-3260

Product Brief

Challenges

Increasingly capable AI systems expand the attack surface when they can access enterprise data, invoke tools, and act in real workflows.

Opportunities

Governed autonomous agents can accelerate high-value work without granting unconstrained access to sensitive systems, data, and business processes.

Business Outcomes

Runtime controls help organizations move agentic AI from proof of concept to governed production with stronger accountability, resilience, and trust.

From AI Risk to Governed Action

How Runtime Controls Help Organizations Deploy Autonomous AI Agents without Expanding Cyber, Data, and Operational Risk

Overview

Artificial intelligence (AI) has created a difficult [management problem](#) for business leaders; the risks are numerous, fast-moving, and unevenly distributed across sectors, stakeholders, and use cases. [MIT Sloan](#) reports that experts identified dangerous capabilities, competitive pressures, AI-enabled weapons and cyberattacks, concentrated power, and false information as the risk areas most likely to produce the most severe harm over the next five years.

For enterprises, the immediate challenge is not only what AI systems can generate, but what increasingly capable systems can do when they are allowed to retrieve data, invoke tools, update records, and trigger downstream actions. As organizations move from experimentation to autonomous and semi-autonomous AI agents, the problem shifts from model output quality alone to the governance of machine-executed authority, especially in sectors such as finance and national security that [MIT Sloan](#) identifies as especially vulnerable to AI risk.

That does not mean every AI risk is a runtime-agent governance problem. Risks such as market concentration, misinformation, model bias, or dangerous model-development capabilities require broader technical, organizational, and policy responses. Treating all AI risk as a single technical-control problem can obscure where targeted controls matter most. In situations where AI agents can access enterprise systems, interact with sensitive data, or take action in business processes, organizations need stronger controls over identity, authorization, delegation, tool access, and auditability to reduce exposure and govern high-value use cases with greater confidence.

Why AI Risk Is Changing

AI risk is changing because increasingly capable systems are being connected to real systems, business processes, and decision flows. [MIT Sloan](#) reports that experts see dangerous capabilities, AI-enabled cyberattacks, competitive pressures, concentrated power, and false information as the most likely risk areas to cause the greatest harm over the next five years.

For organizations, that shift matters because AI can exponentially expand these risks. [MIT Sloan](#) notes that AI is particularly well suited for coding, pattern recognition, and information synthesis. These qualities make it easier for attackers to identify vulnerabilities, generate code, and accelerate offensive cyber work. The same dynamic applies inside enterprises when increasingly capable systems are given access to internal data, connected tools, and operational workflows.

The challenge is compounded by competitive pressure. As [MIT Sloan](#) explains, organizations might feel pressure to deploy AI quickly, resist constraints, or underinvest in safety when they believe that AI will confer a meaningful economic or strategic advantage. In this case, competitive pressure is not a separate risk but a condition that can intensify other risks. For business leaders, the result is a growing gap between what AI systems can do and what governance practices are mature enough to control.

That gap becomes especially important when AI systems move beyond assisting humans and begin acting with limited autonomy.

Which Risks Require Runtime Governance

Runtime governance becomes most important when increasingly capable AI systems are allowed to act inside real workflows. Once an agent can retrieve sensitive data, invoke enterprise tools, update records, trigger transactions, or delegate work to other agents, the question is no longer only whether the model generates useful output. It is whether that system is operating within approved authority, using only permitted tools and data, and producing the evidence needed to investigate and contain misuse.

That makes a narrower but highly consequential set of risks especially relevant. MIT Sloan highlights AI-enabled cyberattacks and dangerous capabilities as major concerns, and both become more operationally important when AI is connected to software, infrastructure, and enterprise processes. In practice, that means organizations must focus on risks such as unauthorized tool use, excessive privilege, credential theft or replay, delegation abuse, data exposure, unverified service connections, and insufficient accountability for autonomous actions.

Competitive pressure makes these risks harder to manage. MIT Sloan notes that organizations may feel pressure to move quickly even when governance practices have not caught up, creating incentives to deploy increasingly capable systems before controls are mature. In that environment, runtime governance is not a replacement for broader AI governance, but it is a necessary control layer for any organization that wants to operate autonomous agents in real workflows without giving them unconstrained authority.

How AgentMinder Helps

AgentMinder helps organizations apply runtime controls to autonomous AI agents that access enterprise data, connected tools, and business workflows. It is not a solution for every AI risk. It is designed to govern the identity, authority, delegation, and actions of agents once they are permitted to operate in real environments.

Table 1: Defining the AgentMinder Role by Risk Category

Risk Category	AgentMinder Role	Examples
Risks that the solution directly addresses	Enforces and records runtime controls over agent actions	Unauthorized tool use, excessive privilege, credential theft or replay, delegation abuse, unapproved service connections, and incomplete action-level accountability.
Risks that the solution helps contain	Limits the operational impact of risks but is not the complete solution	Manipulated prompts or inputs that could trigger actions, AI-enabled cyber misuse, sensitive-data exposure, and risky actions in customer, payment, operational, or mission workflows.
Risks that are outside the primary scope of the solution	Requires complementary technical, organizational, or policy measures	Model bias, misinformation truthfulness, dangerous model-development capabilities, environmental harm, labor disruption, weapons development, and market concentration.

For risks that require runtime governance, AgentMinder provides a dedicated control plane for agent identity and authorization. It gives each agent a verifiable identity, binds its authority to an approved mission, evaluates each tool call against the relevant identity, purpose, target resource, and policy, and preserves the evidence needed to understand what occurred.

- **Verifiable agent identity:** Establish a distinct, accountable identity for every agent rather than relying on shared service accounts, unmanaged API clients, or opaque application credentials.
- **Mission-aware and intent-aware authorization:** Evaluate whether a requested action aligns with the agent’s declared purpose, permitted tools, target resource, and applicable policy (not merely whether a token or API scope is valid).
- **Independent runtime enforcement:** Enforce policy outside the model so manipulated prompts, compromised models, or untrusted inputs cannot directly produce unauthorized actions.
- **Controlled tool access and delegation:** Limit agents to approved tools and services, preserve parent-child authority boundaries, and prevent delegated agents from silently exceeding the mission or permissions originally granted.
- **Runtime observability and audit:** Correlate agent identity, mission, policy decisions, tool calls, and outcomes to support investigation, oversight, incident response, and ongoing policy refinement.

These controls do not replace model safety, enterprise AI governance, cybersecurity programs, or human accountability. They provide the runtime discipline needed to ensure that autonomous agents act within approved authority, use only permitted tools and data, and produce evidence that can be reconstructed during oversight or investigation.

Why Broadcom® AgentMinder

Broadcom approaches this challenge as more than a model-safety or point-product problem. Organizations need more than static guardrails or after-the-fact monitoring; they need a practical way to govern what autonomous agents are authorized to do once those agents are connected to real data, tools, and workflows.

That is where AgentMinder fits. Broadcom designed AgentMinder to serve as a dedicated control plane for agent identity and authorization, giving each autonomous agent a verifiable identity, binding its authority to an approved mission, evaluating each tool call against policy and context, and preserving the runtime evidence needed for investigation and oversight. Rather than claiming to solve every AI risk, this approach focuses on the narrower but critical problem of governing what autonomous agents are authorized to do in production environments.

Broadcom also brings the deployment flexibility and architectural discipline organizations need as they move from pilot project to governed adoption. AgentMinder can run alongside AI models in private clouds, public clouds, or on-premises environments, helping enterprises keep policy enforcement close to sensitive data, operational boundaries, and existing control frameworks. By extending existing identity and access investments with runtime controls for autonomous agents, Broadcom helps organizations move high-value AI use cases forward with stronger security, clearer accountability, and greater confidence that governance can keep pace with deployment.

Why This Matters Now

MIT Sloan's findings suggest that leaders do not need perfect forecasts or complete regulatory certainty before acting. The more immediate challenge is whether governance can keep pace with the speed, capability, and competitive pressure surrounding AI deployment. For organizations exploring autonomous AI, waiting until agents are deeply embedded in critical workflows before addressing runtime control, accountability, and oversight is likely to increase both risk and friction.

That is why the issue is not simply whether AI systems are powerful or useful. It is whether organizations can apply the correct controls at the point where those systems gain access to real data, enterprise tools, and operational authority. In that environment, runtime governance becomes a practical requirement for responsible adoption, helping organizations reduce exposure while building confidence in the safe use of autonomous AI.

The broader lesson is that AI risk should not be treated as a distant concern or a one-time compliance exercise. As MIT Sloan notes, leaders should make AI risk part of the same governance conversations they already have around cybersecurity, privacy, safety, and business continuity. For organizations deploying autonomous agents, that conversation increasingly depends on whether autonomous actions can be controlled with the same seriousness applied to human access, privileged actions, and high-consequence workflows.

Reference:

Brown, S., *These are the most urgent AI risks, according to 272 experts*, MIT Sloan School of Management, July 20, 2026, <https://mitsloan.mit.edu/ideas-made-to-matter/these-are-most-urgent-ai-risks-according-to-272-experts>.



For more information, visit our website at: www.broadcom.com

Copyright © 2026 Broadcom. All Rights Reserved. The term "Broadcom" refers to Broadcom Inc. and/or its subsidiaries. All trademarks, trade names, service marks, and logos referenced herein belong to their respective companies.

FAR-TGA-PB100 August 20, 2026