Shumaker

JANUARY 21, 2026 | PUBLICATION

"Analysis of New Cyber Threats: Artificial Intelligence (AI) Driven Risks Accelerating in 2026"

AI is no longer an emerging risk; it is now a central driver of offensive and defensive cyber capabilities. As organizations adopt AI tools to improve efficiency, adversaries are leveraging the same technology to automate attacks, exploit non‑human identities, and probe complex systems faster than security teams can respond. For clients operating in an increasingly volatile digital environment, understanding how AI is reshaping cyber risk is essential to maintaining a defensible security posture.

INDUSTRY SECTOR

Technology

SERVICE LINE

Technology, Data Privacy, Cybersecurity & AI

RELATED PROFESSIONALS

Nick Carr

MEDIA CONTACT

Wendy M. Byrne
wbyrne@shumaker.com

This article highlights the most consequential AI‑accelerated threats for 2026 and outlines the steps businesses can take to prepare.

AI Supercharges Traditional Attack Methods

One defining trend of cybercrimes is AI-assisted velocity, not new attack methods created by AI. AI makes conventional attacks such as phishing, credential stuffing, and vulnerability exploitation, faster, cheaper, and more scalable than ever before. Threat actors use AI to write highly convincing messages, mimic voices, scan networks for misconfigurations, and automate repeated attack attempts.

AI Agents as High-Risk Identities

Many organizations assign AI agents their own identities (a user account just like a human's), complete with application programming interface (API) keys, delegated permissions, and autonomous action workflows. While this enables automation and efficiency, it also introduces a new class of identity risk that traditional cybersecurity controls were never designed to manage.

1. Misconfigured AI Agents Can Become High-Privilege Backdoors

An AI agent with its own identity is, functionally, a non-human employee, but one that works at machine speed, never sleeps, and can be manipulated far more easily.

A misconfigured agent can read or move sensitive data, trigger automated workflows, execute privileged actions in cloud systems, and interact with clients or staff. Because these identities often bypass multi-factor authentication (MFA), operate continuously, and may not be audited or rotated, AI identities provide attackers with highly attractive targets.

2. AI Agents Can Be Compromised Through Prompt Injection

Prompt-injection attacks can manipulate an AI agent into performing unauthorized actions using its own credentials, effectively turning the AI agent into an unwitting insider threat. If an attacker manipulates the model's inputs, through emails, client documents, shared data, or even User Interface (UI) instructions, they can force the agent to:

- Expose sensitive information it has permission to access
- Execute actions "as the AI agent" using the agent's identity
- Bypass policy constraints
- Make alterations to systems, workflows, or datasets
- Execute transactions

This is especially dangerous when the agent's identity holds privileged or system-level access.

3. Non-Human Identities are Rarely Managed Like Humans

Unlike human employees, AI identities:

- Don't rotate passwords unless expressly configured to do so
- Don't leave the company (meaning their credentials rarely expire)
- Are embedded in automations with no clear owner
- Often lack MFA or behavioral monitoring

This makes AI identities ideal persistent footholds for attackers.

4. Attackers Focus on the Vulnerabilities of AI Agents

Because AI agents follow deterministic rules and scripts, attackers can know:

- Exactly where to poke
- Which workflows are automated
- Which identities lack MFA
- Which logs nobody checks

In short, AI agents are easier to exploit than humans, and they cause damage faster.

5. AI Agents in Browsers

Browser-based AI agents represent a uniquely exposed attack surface. Because they operate within the same environment where users access internal dashboards, Software as a Service (SaaS) platforms, and

sensitive web applications, they inherit the user's identity, session tokens, and access rights.

AI agents in browsers can be manipulated by malicious websites, scripts, or embedded UI elements through prompt‑injection techniques, allowing attackers to trigger unauthorized actions "as the user." Many browser agents also have the ability to click, submit, navigate, or execute workflows, which means a compromised agent can modify settings, access restricted data, or carry out transactions at machine speed. With limited logging, oversight, or permission controls, browser‑based AI agents have quickly become one of the fastest‑growing and least‑monitored cybersecurity risks.

Mitigation Strategies: Defending at Machine Speed

To respond effectively, organizations must adopt blended strategies combining governance, technical defenses, and cultural readiness. Here are some practical risk-mitigation strategies for 2026.

1. Establish AI Governance Frameworks

- Incorporate AI‑specific risk assessments, model oversight, and audit trails.

2. Implement Identity-Centric Security

- Treat AI agents as first‑class identities with their own access controls and lifecycle management.
- Regularly conduct a full inventory of non-human identities.
- Enforce least privilege access and automated key rotation.
- Implement continuous monitoring for anomalous machine to machine interactions.

3. AI Firewals

An AI Firewall is a security layer that filters inputs, actions, and outputs of AI systems. Instead of blocking network traffic like a traditional firewall, AI Firewalls analyze the intent of prompts, the behaviors of AI agents, and the sensitivity of the data they handle or expose. By enforcing policy guardrails, an AI Firewall can stop malicious inputs, unsafe agent behavior, and unintended data leakage before harm occurs.

4. Use AI to Defend Against AI

- Deploy AI‑enabled threat detection capable of spotting behavioral anomalies, synthetic media, and cloud‑based impersonation.

5. Browser AI Agent Protections

- Run AI agents only in sandboxed browser profiles with no stored credentials.
- Restrict the agent's ability to click, submit, or navigate without human approval.
- Block high risk scripts, iframes, and unknown domains when agents are active.
- Log all agent actions for visibility and investigation.

6. Enhance Incident Response with Autonomous Capabilities

- Detect anomalous agent behavior (mass downloads, rapid clicks, off hours actions).
- Prebuild playbooks for prompt injection, token theft, or compromised agents.
- Practice red team and tabletop exercises targeting AI agent failure modes.

7. Training & Culture

- Educate staff on AI‑agent capabilities, shadow‑AI risks, and hidden prompt‑injection vectors.
- Require disclosure and approval for any browser extensions using AI.

Conclusion

AI is not simply another category of cyber threat, it is the infrastructure enabling a new generation of hyper‑efficient, hyper‑scalable attacks. Businesses depend on trust, accuracy, and confidentiality, and must urgently adapt to this changing landscape. The businesses that succeed in the AI era will integrate AI‑informed governance, deploy machine‑speed defenses, and maintain vigilant human oversight.

If you would like more information on managing AI-accelerated cyber risk—including governance of non-human identities, prompt-injection exposure, and browser-based AI agent controls—please contact Shumaker Partner Nick Carr.

Whether you are building an AI governance framework, tightening identity-centric access controls and key rotation for AI agents, or deploying machine-speed safeguards such as AI firewalls and enhanced monitoring, Shumaker's Technology, Data Privacy, Cybersecurity & AI Service Line provides forward-looking, practical guidance to help organizations stay secure, resilient, and compliant as AI-driven threats accelerate in 2026.

Shumaker