NITIC News Article
Blog Post Published 09.17.2026
https://www.nitic.org/cybersecuritys-ai-paradox-our-deadliest-adversary-or-our-last-ally/

Cybersecurity's AI Paradox: Our Deadliest Adversary or Our Last Ally?

Dr. Michael Greer
Associate Professor
Columbus State Community College

If you measure AI trends by simply watching commercials that appear during your favorite TV shows, you could clearly see that in 2025 the conversation shifted from simply mentioning AI to a more specific and newer concept: AI agents. Organizations have moved beyond experimenting with AI and saying they are thinking about incorporating it into their products, toward a tangible technology known as AI agents. This shift has come about not because agents generate better text or images, but because they represent an evolution of AI systems toward more autonomous, capable, and deeply integrated pipelines (Amazon Web Services [AWS], 2024; Google Cloud, 2024).

The presentation focused on Generative AI. A future blog post will cover Agentic AI security.

Artificial intelligence didn't just change what's possible in cybersecurity — it changed the math. The same models that let a security operations center triage a year's worth of alerts in an afternoon are also letting an adversary draft a thousand flawless spear-phishing emails before lunch. That's the paradox at the center of this talk: AI is simultaneously the fastest-growing risk in the industry and the most promising tool we have to fight back.

2026: The Era of Operational Dependency

The experimentation phase is over. AI is no longer a side project bolted onto security tooling — it's operational infrastructure, and the numbers back that up:

- **The Driver:** In a recent World Economic Forum survey, AI was named the single most significant driver of cybersecurity change for 2026.
- **The Risk:** Organizations are now identifying AI-related vulnerabilities as their fastest-growing category of cyber risk.
- **The Reality:** A meaningful share of global breaches now directly involve AI capabilities, according to IBM's 2025 breach data.

The baseline has moved. Whatever your organization's relationship with AI was a year ago, it isn't that anymore. Organizations must not dismiss AI, rather embrace it.

The Double-Edged Shield

Every capability AI gives a defender, it also gives an adversary — the difference is who adopts it faster.

- **Automation of Attack:** The scale tips toward adversaries when they weaponize AI to expand the attack surface faster than human teams can patch it.
- **Acceleration of Defense:** The scale tips back toward defenders when AI is used to compress detection, triage, and response times.

Organizations that intend to stay in the fight won't be the ones that avoid AI — they'll be the ones that secure AI while using AI to secure everything else.

The New Economics of Attack

Building on a framework from security expert Richard Bird, the talk makes the case that AI doesn't just add new attack techniques — it rebalances the *cost* of offense entirely:

Attack Stage	Traditional Cost	AI-Augmented Cost
Reconnaissance & Phishing	High human effort to research targets and craft context	Near-instant, localized, grammatically flawless mass spear-phishing
Payload Iteration	Manual drafting and recompilation	Automated, continuous mutation to evade detection
Lateral Movement	Sequential human decision-making	Scripted adaptation that outpaces human defenders

Work that used to require a skilled team and days of effort can now be generated, iterated, and deployed at machine speed. That asymmetry — cheap offense against expensive defense — is the paradox in its purest form.

The Death of Traditional Trust Signals

Drawing on current identity research, the talk argues that identity has stopped being a credential problem and become a *synthetic content* problem. Voice cloning, synthetic video, and fabricated executive instructions can now convincingly impersonate legitimate behavioral patterns inside a trusted corporate system. Trust signals based solely on how someone sounds or looks on a call have been, in the speaker's words, permanently compromised.

AppSec in the "Mythos" Era

The talk points to a quieter crisis unfolding in application security. Traditional perimeter defenses — firewalls, WAFs — were built to inspect structured traffic. But a language model is steered entirely through its inputs, in natural language. Traditional defenses simply can't parse a conversational payload designed to extract sensitive data or manipulate an underlying API. AI didn't just add a new attack surface; it broke assumptions AppSec has relied on for two decades.

AI as the Last Ally

Pivoting from threat to solution. If attacks are scaling at machine speed, defense has to as well.

The incident lifecycle gets compressed. Traditional response — alert, triage, correlation, threat hunting, containment — assumed humans had time to think between steps. That assumption doesn't hold anymore. The key insight: security teams can't out-scale AI-driven attacks by hiring more people. Machine-speed iteration has to be met with machine-speed detection.

Process-first SOC automation. Applying recent frameworks, the talk lays out a practical pipeline: AI ingests millions of raw logs and alerts, clusters them by attack chain to eliminate false positives, and generates investigation summaries and threat-hunting context — all before a human ever looks at it. A SOC analyst then reviews the AI-generated context and approves the containment action. The result: AI is strongest not as a replacement for analysts, but as a force multiplier that turns Tier 1 responders into Tier 3 threat hunters.

The business case is measurable. Organizations running legacy operations without AI integration see expanding breach lifecycles and compounding incident response costs. Organizations leveraging AI extensively see significantly reduced breach costs and drastically shortened containment timelines, per combined IBM and WEF 2025–2026 data. Defensive AI isn't a nice-to-have technical upgrade — it's a financial imperative.

The Governance Gap

Adoption is outrunning control. As AI adoption inside organizations climbs, security controls are lagging further and further behind — and a significant share of modern breaches trace back to AI pipelines the business deployed without IT oversight in the first place ("shadow AI").

The fix isn't simply fighting AI with more AI. It's locking the technology inside a mature, pragmatic governance framework.

A Framework That Actually Fits: NIST CSF 2.0

Rather than treating AI risk as an entirely new discipline, the talk argues it belongs inside existing cyber risk management — specifically through three pillars of action:

1. **Secure AI Systems** — govern the models you build.
2. **Defend with AI** — empower the SOC.
3. **Thwart AI Attacks** — harden the perimeter against synthetic threats.

Practical Safeguards, Straight from OWASP

For teams building or deploying LLM-powered tools, the talk offers three concrete engineering fixes for prompt injection and API bypass risks:

- **Validation & Isolation:** Treat every LLM output as untrusted user input. Sandbox model interactions entirely away from core backend databases.
- **Least Privilege for Agents:** Restrict model API access strictly to what's necessary — an LLM should never hold root access capable of orchestrating system-wide changes.
- **Adversarial Red Teaming:** Continuously test model boundaries with automated prompt-injection payloads *before* deployment, not after.

Building Resilience, Not Just Detection

The talk closes with a shift in mindset — from purely detecting AI-driven threats to being able to recover from them:

1. **Trust & Identity:** Cryptographically sign legitimate internal data so AI-generated fabrications can be detected.
2. **Data Provenance:** Keep immutable logs of exactly where AI models source their training data and what decisions they make.
3. **Automated Recovery:** Use AI not just to detect compromise, but to dynamically rebuild and restore clean environments afterward.

The bottom line: if an AI system fails or is compromised, you need to know exactly how to contain it and roll it back — before you need to.

The Takeaway

AI hasn't handed either side of the security fight a permanent advantage. It's handed both sides a faster one. Adversaries get cheaper reconnaissance, better-disguised phishing, and identity spoofing that breaks the trust signals we've relied on for years. Defenders get compressed incident lifecycles, SOC automation that turns junior analysts into threat hunters, and a measurable return on investment.

The organizations that come out ahead won't be the ones that pick a side of the paradox. They'll be the ones that govern AI rigorously, defend with it aggressively, and build the resilience to recover when — not if — something gets through.

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.