

Tank vs. Drone: Rethinking Cybersecurity Strategy in the Age of AI

Yaron R. Levi

A word before you begin

Most conversations about AI and cybersecurity start with fear: smarter attacks, faster criminals, tools we do not yet understand. Yaron Levi starts somewhere more useful. In *Tank vs. Drone*, he shows that AI rarely invents new weaknesses. It finds the ones we already have, the unpatched server, the forgotten admin account, the “temporary” rule that became permanent, and exploits them at machine speed.

His image is one that sticks: organisations driving heavily armoured, expensive tanks into a sky full of cheap, tireless drones. Yaron’s answer is not more armour. It is discipline: know what you have, simplify before you automate, and govern AI with the same rigour as any other critical system. For everyone working on AI literacy, it is a timely reminder that understanding AI also means understanding how it can be turned against us.

This paper is Yaron’s contribution to the Global AI Literacy Project 2026. His insights contribute to the Chapter 2 of the Global AI Literacy Study, which we will present on 22 October 2026 at the House of Lords in London. I hope this paper helps you look at your own organisation with fresh eyes.

Anastassia

Global AI Literacy Project 2026/ CEO and Founder of AI Edutainment GmbH

About the author



Yaron R. Levi is a Chief Information Security Officer with more than two decades of experience across technology, healthcare and entertainment. He has served as CISO of Dolby Laboratories and, before that, as CISO of Blue Cross and Blue Shield of Kansas City and Director of Information Security at Cerner Corporation. Earlier in his career he held security roles at Intuit, eBay and ANX. He is a Research Fellow of the Cloud Security Alliance, advises venture capital firms and security start-ups, co-founded the Kansas City CISO forum and serves on the board of InfraGard Kansas City. He is a 2018 graduate of the FBI CISO Academy.

Tank vs. Drone

Rethinking Cybersecurity Strategy in the Age of AI

*Why AI-enabled attackers are outmaneuvering armored, perimeter-first defenses —
and what security leaders should do about it*

Yaron R. Levi, CISO - August 2026

Executive Summary

For three decades, enterprise cybersecurity has largely followed a perimeter-first logic: build strong outer defenses, add layers of controls, and trust that scale and thickness will hold. Artificial intelligence is exposing the limits of that logic — not by inventing new categories of attack, but by letting adversaries operate cheaply, adaptively, and at machine speed against defenses that remain expensive, complex, and slow to change.

This paper frames the current threat environment as an asymmetry between tanks and drones. The traditional enterprise security stack — firewalls, VPNs, endpoint detection and response, data loss prevention, web application firewalls, multi-factor authentication, and security information and event management platforms — functions like a tank: heavily armored, expensive, and optimized for predictable threats moving along known paths. It increasingly faces swarms of cheap, AI-enabled “drones”: automated reconnaissance tools, generative-AI-assisted phishing operations, and, in at least one documented case, an AI agent that autonomously carried out most stages of a full cyberattack lifecycle against dozens of organizations.

Drawing on recent government, standards-body, and peer-reviewed research — including the G7 Cyber Expert Group, the EU Agency for Cybersecurity (ENISA), the World Economic Forum, the U.S. National Institute of Standards and Technology, MITRE, and the Verizon and IBM/Ponemon annual breach studies — this paper finds a consistent conclusion: AI does not rewrite the fundamentals of cybersecurity. It amplifies existing weaknesses — unpatched systems, misconfigurations, weak identity governance, and accumulated technical debt — and compresses the time available to find and exploit them. Most organizations remain more exposed by unmanaged complexity than by any AI-specific novelty.

The paper closes with five practical shifts for security leaders: treat technical debt as a visible, owned risk metric rather than a background cleanup task; invest in telemetry and data quality before adding more AI-branded tools; deploy AI first in the domains where response latency is most costly; secure AI systems themselves end-to-end, using emerging technical standards; and close the AI governance gap that current research shows most organizations still have wide open.

1. Introduction: From Trench Warfare to Drone Swarms

Many security leaders built their careers in an era when “defense in depth” meant bigger firewalls, more appliances, and a thicker stack of controls at the network edge. That posture increasingly resembles trench warfare: dug in, heavily fortified, and built to resist a slow-moving, symmetrical opponent. Today it can feel more like trying to move a main battle tank through a sky full of cheap, autonomous drones.

In its 2026 Global Cybersecurity Outlook, the World Economic Forum found that 94% of surveyed executives now identify AI as the most significant driver of change in cybersecurity for the coming year [1]. This paper examines what, specifically, is changing — and, just as importantly, what is not.

2. The End of the Perimeter-Only Model

For decades, enterprise security followed what is often called a “Maginot Line” mindset: build strong perimeters, central choke points, and a thick stack of controls at the network edge. Over the past decade, that model has evolved toward “identity is the new perimeter,” greater investment in detection and response, and an operating assumption of eventual breach. The shift has been directionally correct, but incomplete.

The G7 Cyber Expert Group—the technical working group that advises G7 finance ministries on cyber risk to the financial system—published a statement in October 2025 explicitly describing AI as a “double-edged sword” for cybersecurity: it strengthens detection and response for defenders, while also increasing the speed, scale, and impact of malicious cyber activity [2]. The European Union Agency for Cybersecurity (ENISA) went further in its own 2025 Threat Landscape report, describing 2025 as “the first year in which AI fundamentally reshaped the cyber threat landscape,” based on an analysis of nearly 4,900 tracked incidents [3].

In other words, many organizations are still driving tanks — heavy, expensive, slow-to-turn architectures — into a fight that increasingly rewards swarms of software-defined drones.

3. Tanks and Drones: An Asymmetric Fight

Picture the traditional enterprise stack as a tank: expensive, heavily armored with controls such as firewalls, VPNs, endpoint detection and response, data loss prevention, web application firewalls, multi-factor authentication, and SIEM platforms, and optimized for predictable threats moving through known paths, such as those catalogued in the MITRE ATT&CK® framework [4].

As in most forms of asymmetric conflict, attackers need to succeed once; defenders must succeed every time. The defender's structural advantage is that attackers rarely have complete knowledge or context about their target. But that advantage narrows sharply when reconnaissance itself can be automated, run continuously, and scaled at near-zero marginal cost.

That is no longer hypothetical. In 2025, Anthropic disclosed and disrupted what it described as the first documented large-scale cyberattack executed with substantial AI orchestration. According to the company's report, an AI coding agent was used to automate reconnaissance, harvest credentials, and penetrate victim networks; the AI system also made tactical decisions about which stolen data to exfiltrate and drafted psychologically targeted extortion demands, in some cases exceeding \$500,000 per victim, against at least 17 organizations spanning healthcare, emergency services, government, and religious institutions. Anthropic concluded that "AI has lowered the barriers to sophisticated cybercrime," enabling operators with limited technical skill to run operations that would previously have required years of training [5].

This is the pattern regulators and researchers describe more broadly. ENISA documented state-linked actors from China, Iran, and North Korea using commercial AI platforms for reconnaissance, code generation, and persona creation, alongside a shift toward purpose-built malicious AI tools designed to evade detection [3]. The World Economic Forum likewise found that criminals are using generative AI to automate phishing campaigns and produce deepfakes, with 29% of surveyed cybersecurity leaders naming the advancement of adversarial AI capability as a top concern [1].

The asymmetry, then, is economic as much as technical: defenders continue to invest in expensive, hard-to-reconfigure "tanks," while attackers increasingly field swarms of inexpensive, disposable, constantly-learning software drones.

4. Complexity and Technical Debt: The Real Attack Surface

Here is the uncomfortable part: a meaningful share of the modern security stack exists not because the world is inherently more dangerous, but because IT environments are inherently undisciplined. Every unmanaged SaaS application, policy exception, "temporary" firewall rule, legacy subnet, one-off VPN, and orphaned administrative account is an artifact of complexity and weak operational discipline — technical debt accumulated in the pursuit of speed. The common organizational response is to bolt on more tools rather than resolve the underlying debt: another monitoring platform, another gateway, another agent, hoping additional armor compensates for missing hygiene.

The data bears this out. In its 2026 Data Breach Investigations Report, Verizon found that exploitation of unpatched vulnerabilities overtook credential abuse as the single leading initial access vector, present in 31% of breaches — even as the median time

organizations take to patch known-exploited vulnerabilities rose from 32 to 43 days year over year, a 34% increase that runs directly counter to the pace attackers now operate at [6]. Credential theft, excessive permissions, and misconfigured assets round out the report's leading exposure categories, and a majority of breaches still involve some human element [6].

AI does not need to be sophisticated to exploit gaps like these; it simply needs to be fast, tireless, and cheap enough to run continuously against every exposed asset an organization has. The problem, in short, is not only that attackers have drones. It is that many organizations have turned their own infrastructure into a labyrinth of technical debt that those drones are very good at exploring.

5. What AI Is Actually Changing

Across government statements, regulator guidance, and industry research, a consistent picture emerges: AI is primarily an amplifier. It increases the speed, scale, and accessibility of both attack and defense [1][2][3]. On offense, generative AI is being used to produce fluent phishing in multiple languages, simulate voices and faces for social-engineering, and generate new malware variants intended to slip past legacy, signature-based detection. On defense, organizations are adopting AI across the security stack: the World Economic Forum found 77% of organizations have adopted AI for cybersecurity in some form, including 52% for phishing detection, 46% for intrusion and anomaly response, and 40% for user- and entity-behavior analytics [1].

The table below summarizes several of the most-cited, most-recent figures behind this shift.

Metric	Figure	Source
Organizations reporting some use of AI for cybersecurity defense	77%	WEF, 2026 [1]
Social-engineering campaigns assessed as AI-assisted by early 2025	>80%	ENISA, 2025 [3]
Global mean time to identify and contain a breach (nine-year low)	241 days	IBM/Ponemon, 2025 [10]
Average global cost of a data breach (down 9% year over year)	\$4.44M	IBM/Ponemon, 2025 [10]
Organizations with no formal AI governance policy	63%	IBM/Ponemon, 2025 [10]
Breaches where vulnerability exploitation was the initial access vector	31% (leading vector)	Verizon DBIR, 2026 [6]
Median time to patch known-exploited vulnerabilities (up from 32 days)	43 days	Verizon DBIR, 2026 [6]

AI is also introducing a genuinely new attack surface: the AI systems themselves. NIST's updated taxonomy of adversarial machine learning describes attack classes including prompt injection, data poisoning, model extraction (“model theft”), and adversarial examples engineered to fool AI-based detectors [7]. The OWASP Top 10 for Large Language Model Applications catalogues the same risk categories from an application-security perspective, aimed at practitioners building or deploying AI-enabled tools [8]. MITRE's ATLAS™ knowledge base — a counterpart to ATT&CK built specifically for AI systems — now documents dozens of adversary tactics and techniques targeting machine-learning pipelines [9].

Yet the same research shows that AI governance has not kept pace with AI adoption. IBM Security's 2025 Cost of a Data Breach Report, produced with the Ponemon Institute, found that 97% of organizations that experienced an AI-related security incident lacked proper access controls on their AI systems, that 63% had no formal AI governance policy at all, and that unsanctioned “shadow AI” added an average of \$670,000 to the cost of a breach where it was present [10]. The World Economic Forum similarly found that only 64% of organizations formally assess the security of the AI tools they use — an improvement from 37% a year earlier, but still short of universal practice [1].

AI, in short, does not create an entirely new class of cyber risk overnight. It accelerates the risks organizations already carry and layers on a distinct set of AI-specific failure modes that most organizations have not yet learned to govern.

6. What Hasn't Changed: The Core Principles

Strip away the hype, and the core security principles that have held since the 1990s remain intact:

- Know what you have.
- Minimize implicit trust.
- Reduce attack surface.
- Design for failure and resilience, rather than perfect prevention.

The G7 Cyber Expert Group, ENISA, and the World Economic Forum all converge on the same underlying point: AI mainly increases the frequency and impact of already-known attack classes rather than inventing wholly new ones [2][3][1]. And per Verizon's 2026 findings, most major incidents still trace back to familiar failings — exposed and unpatched services, mismanaged credentials, and insufficient segmentation — rather than to anything uniquely AI-driven [6].

AI does not remove the need for basic security engineering. It raises the cost of neglecting it.

7. The Real Shift: From Human Speed to Machine Speed

Where AI does fundamentally change the landscape is in timing and tempo.

On the defensive side, IBM's 2025 research found that the global mean time to identify and contain a breach fell to 241 days — the lowest figure in nine years — and attributed the improvement directly to faster breach containment driven by AI-powered defenses; average global breach costs fell 9%, from \$4.88 million to \$4.44 million, over the same period [10]. Independent, peer-reviewed literature reviews reach a similar conclusion from a different angle. A 2025 systematic review in ACM Computing Surveys and a companion 2025 IEEE literature review both synthesize a growing body of empirical studies showing that AI-assisted detection, behavioral analytics, and automated alert triage measurably improve security-operations-center outcomes — while cautioning that integration complexity and inconsistent data quality can just as easily erode those gains [11][12].

On the offensive side, the same tempo shift is visible. In the AI-orchestrated campaign Anthropic disrupted, a single operator used an AI agent to run reconnaissance, credential harvesting, exploitation, data-triage, and extortion-note generation across at least 17 targets with minimal direct human involvement in day-to-day execution [5]. Processes calibrated to human pace — manual ticket queues, monthly change-advisory boards, weekly patch windows — become liabilities once adversaries are operating at machine speed on the other side of the fight.

Humans remain essential for designing guardrails, handling ambiguity, setting risk appetite, and making complex tradeoffs. But the day-to-day control loop increasingly cannot afford to wait for a human to click “approve.”

8. Building Defensive AI Capabilities

If attackers are fielding AI-enabled “drones,” defenders need their own — autonomous or semi-autonomous capabilities that can observe, decide, and act across the environment, with humans supervising strategy rather than manually driving every maneuver. A few building blocks are effectively non-negotiable:

- **Rich, consistent telemetry.** AI models require clean, comprehensive data from endpoints, identity systems, networks, and applications; without it, they are either blind or noisy.
- **Behavior-based detection.** Static signatures and simple rules cannot keep pace with AI-mutated malware and never-before-seen attack chains, which is pushing organizations toward anomaly- and behavior-based models [11][12].
- **Automated containment.** Security orchestration, automation, and response (SOAR) tooling paired with AI can quarantine devices, disable accounts, revoke

tokens, and open incidents without waiting on a human queue, directly supporting the faster containment times documented in current breach research [10].

- **Secure-by-design AI pipelines.** Joint guidance from CISA, the U.S. National Security Agency, the UK National Cyber Security Centre, and international partner agencies recommends controlling data provenance, restricting model access, and monitoring AI systems throughout their lifecycle to mitigate poisoning, evasion, and misuse [13], complementing MITRE's ATLAS knowledge base for structuring defenses against AI-specific attack techniques [9].

Together, these form a defensive swarm: software that constantly patrols, correlates, and reacts, with human judgment focused where it adds the most value.

9. Fixing Complexity Before Scaling AI

There is a catch: dropping AI into a chaotic environment tends to amplify the chaos rather than resolve it. AI tools assume a baseline of operational hygiene — reasonably accurate asset inventories, well-managed identities, consistent logging, and predictable network behavior. Without that baseline, models mislearn, automations misfire, and analysts stop trusting the system.

This is not a hypothetical risk. IBM's 2025 research found that 97% of organizations with an AI-related security incident lacked proper access controls over their AI systems, and that 63% had no AI governance policy in place at all [10]. CISA, NSA, and NCSC-UK guidance treats secure-by-design practices, governance, and data vetting as prerequisites for safe AI adoption, not optional add-ons layered in afterward [13].

So before — or at minimum, in parallel with — “AI-enabling” everything, organizations should:

- Retire redundant tools and policies that obscure real risk.
- Simplify architectures where possible, and segment rigorously where they cannot be simplified.
- Standardize identity and access management around a single source of truth.
- Treat configuration management and infrastructure-as-code as core security controls, not DevOps conveniences.

Skip this step, and defensive AI amounts to strapping drones onto an already overloaded tank, with no guarantee they are aimed at the right target.

10. Does AI Change the Fundamentals?

Returning to the original question — does AI fundamentally change cybersecurity principles? — the most defensible answer, consistent with current government and

standards-body research, is: no, but it dramatically tightens execution tolerances and adds a distinct set of AI-specific safeguards [2][3][1][7].

The timeless basics still matter most: clear asset understanding, minimum necessary trust, disciplined operations, and resilience planning. What is new is the requirement that these basics operate at machine speed, and that organizations now must secure the AI systems themselves — models, data pipelines, prompts, and outputs — with the same rigor applied to any other critical system.

AI is less a new religion for security than a new set of physics: the same principles, under materially different constraints.

11. Recommendations for Security Leaders

In a tank-versus-drone environment, security leaders can tilt the asymmetry back in their favor by focusing on a small number of practical shifts:

- **Make technical debt visible and owned.** Tie architectural simplification, decommissioning, and configuration hygiene to explicit risk-reduction metrics, rather than treating them as background “cleanup projects.”
- **Invest in good data before clever models.** Ensure solid logging, tracing, identity signals, and asset inventories exist before adding another AI-branded product; peer-reviewed research is consistent that data quality, more than model sophistication, determines real-world AI effectiveness [11][12].
- **Deploy AI where latency hurts most.** Start with phishing detection, anomaly detection, alert triage, and automated containment — domains where current adoption data already shows the clearest gains in speed and accuracy [1][10].
- **Secure AI end-to-end.** Apply NIST's adversarial machine learning taxonomy, the OWASP LLM Top 10, MITRE ATLAS, and CISA/NSA/NCSC secure-AI-development guidance to assess AI risk, vet data sources, and scrutinize third-party AI providers [7][8][9][13].
- **Close the AI governance gap.** Formal AI governance and access-control policies remain the exception rather than the rule; closing that gap is one of the highest-leverage, lowest-cost steps available, given how directly it correlates with breach cost and severity in current research [10].
- **Re-skill teams for AI-augmented work.** Prepare analysts to collaborate with AI systems: understanding their strengths and failure modes, designing robust playbooks, and maintaining human oversight where it matters most.

12. Conclusion

The future of cybersecurity in the age of AI will not be won by whoever fields the biggest tank. Current government, standards-body, and peer-reviewed research points in a consistent direction: organizations that simplify their terrain, field trustworthy defensive AI capabilities, and continuously adapt at machine speed — without losing human judgment in the process — will be the ones best positioned to close the asymmetry that AI has opened up.

References

- [1] World Economic Forum. Global Cybersecurity Outlook 2026. January 2026.
<https://www.weforum.org/publications/global-cybersecurity-outlook-2026/>
- [2] G7 Cyber Expert Group. Statement on AI and Cybersecurity. U.S. Department of the Treasury, October 6, 2025. <https://home.treasury.gov/system/files/136/G7-Cyber-Expert-Group-Statement-AI-and-Cybersecurity-2025.pdf>
- [3] European Union Agency for Cybersecurity (ENISA). ENISA Threat Landscape 2025. October 2025.
<https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025>
- [4] MITRE. MITRE ATT&CK®. Accessed August 2026. <https://attack.mitre.org/>
- [5] Anthropic. “Disrupting the First Reported AI-Orchestrated Cyber Espionage Campaign.” 2025.
<https://www.anthropic.com/news/detecting-counteracting-misuse-aug-2025>
- [6] Verizon. 2026 Data Breach Investigations Report. May 2026.
<https://www.verizon.com/business/resources/reports/dbir/>
- [7] National Institute of Standards and Technology (NIST). Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations (NIST AI 100-2e2025). 2025.
<https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2025.pdf>
- [8] OWASP Foundation. OWASP Top 10 for Large Language Model Applications, 2025.
<https://owasp.org/www-project-top-10-for-large-language-model-applications/>
- [9] MITRE. MITRE ATLAS™ (Adversarial Threat Landscape for Artificial-Intelligence Systems). Accessed August 2026. <https://atlas.mitre.org/>
- [10] IBM Security, in partnership with the Ponemon Institute. Cost of a Data Breach Report 2025. 2025.
<https://www.ibm.com/think/x-force/2025-cost-of-a-data-breach-navigating-ai>
- [11] ACM Computing Surveys. “Exploring the Role of Artificial Intelligence in Enhancing Security Operations: A Systematic Review.” 2025. <https://dl.acm.org/doi/10.1145/3747587>
- [12] IEEE. “Empowering Security Operation Center with Artificial Intelligence and Machine Learning — A Systematic Literature Review.” IEEE Access, 2025. <https://ieeexplore.ieee.org/document/10850912/>
- [13] Cybersecurity and Infrastructure Security Agency (CISA), National Security Agency (NSA), UK National Cyber Security Centre (NCSC), and international partners. Guidelines for Secure AI System Development (November 2023) and Deploying AI Systems Securely (April 2024).
<https://www.cisa.gov/news-events/alerts/2023/11/26/cisa-and-uk-ncsc-unveil-joint-guidelines-secure-ai-system-development> ; <https://media.defense.gov/2024/apr/15/2003439257/-1/-1/0/csi-deploying-ai-systems-securely.pdf>