WHITE PAPER

CROWDSTRIKE

Five Steps for Frontier AI Security Readiness

Table of Contents

Executive Summary

Frontier AI is reshaping how organizations must think about cyber risk. Frontier models are a new class of highly capable AI systems that can reason across complex tasks, analyze software, identify vulnerabilities, accelerate exploit development, and support increasingly sophisticated security workflows. Anthropic's Claude Mythos and OpenAI's GPT-5.4-Cyber are early examples of this shift, showing how quickly AI is expanding both offensive and defensive capability.

As vulnerabilities are discovered and exploited on shorter timelines, traditional security approaches built on periodic assessments, severity scores, and human-paced response are becoming less effective. Defenders need a new model centered on exploitability, continuous validation of exposure, stronger prevention, cross-domain visibility, decisive response, and governed use of AI.

This shift changes what defenders need to do. The challenge is no longer just finding vulnerabilities faster than adversaries. It is figuring out which weaknesses are truly exploitable, reducing the conditions that turn them into real risk, and responding as quickly as attackers can move. As AI speeds up both discovery and exploitation, organizations need to move from periodic assessment to continuous, intelligence-driven exposure management so they can determine what really matters, prioritize real risk, and quickly coordinate remediation. They must also prepare for a surge in vulnerability discovery and patch activity that many organizations are not operationally prepared to absorb. The goal is no longer just better hygiene. It is resilience in an environment where offensive capability is improving faster than traditional security programs were built to handle.

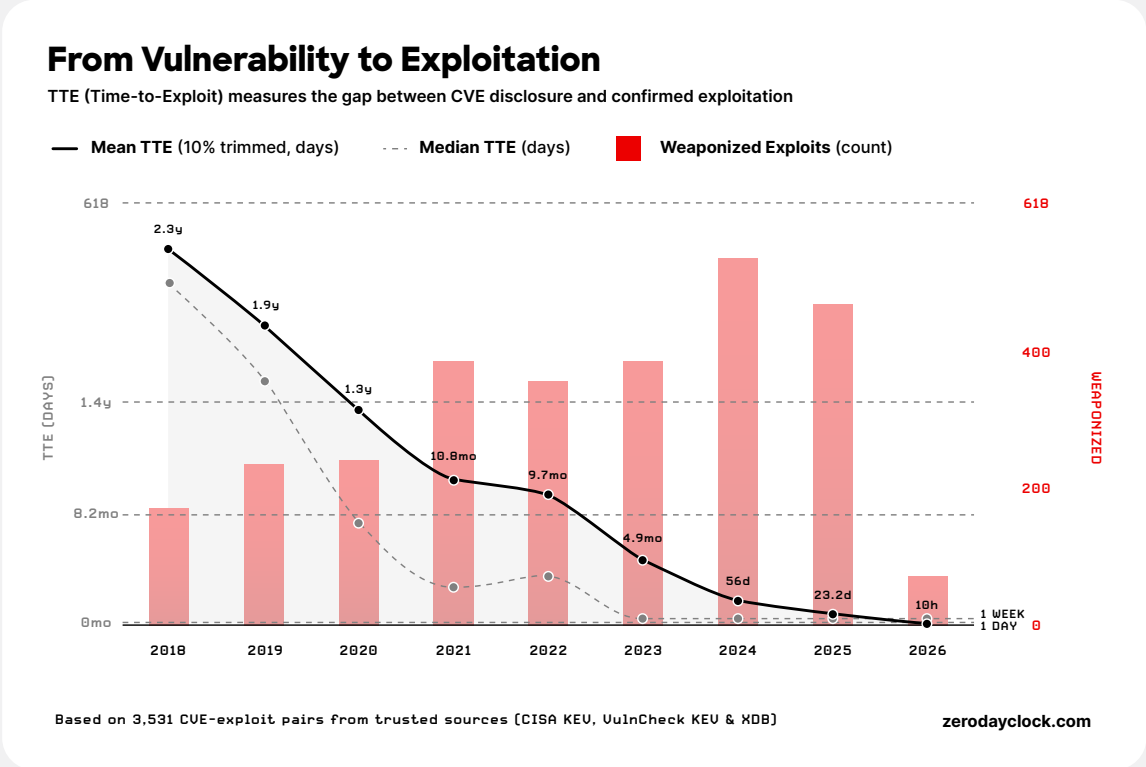
Organizations do not need to wait to improve readiness. They can act now by tightening remediation and recovery workflows, running regular validation exercises, reducing attack surface and telemetry blind spots, and improving how risk is prioritized. That means focusing less on severity scores alone and more on exploitability, business impact, adversary behavior, and attack path relevance. Most importantly, leaders should treat this as a business resilience issue, not just a security issue, and align security, IT, engineering, and executive teams around timely decisions, clearer ownership, and tighter coordination.

The New Reality: Offensive Capability Is Accelerating

For years, defenders operated with at least some assumption of friction on the attacker side. Discovering a vulnerability was one thing. Turning it into a usable exploit, chaining it into a broader attack, and operationalizing it against a live environment took skill and time. That delay created a window, however imperfect, for patching, mitigation, and detection.

Frontier AI models are collapsing that window.

The CrowdStrike 2026 Global Threat Report reinforces the scale of the shift. In 2025 alone, CrowdStrike observed an 89% year-over-year increase in attacks by AI-enabled adversaries and a 42% increase in zero-day vulnerabilities exploited before public disclosure. Frontier AI is lowering the cost and skill barrier for attackers faster than defenders can patch, compressing the time between exposure and exploitation (See Figure 1). It is also accelerating vulnerability discovery itself, creating a coming surge of disclosures, fixes, and patch decisions that will strain already overloaded teams.



Source: Zero Day Clock

Figure 1. The window from vulnerability to exploitation is collapsing from years to hours.

AI systems are beginning to assist with tasks that materially improve offensive velocity: analyzing code, identifying vulnerable logic, generating proof-of-concept exploits, mapping misconfigurations into attack paths, and accelerating the translation of raw exposure into usable intrusion techniques. The same capabilities can also improve defense. Frontier AI can help organizations analyze code more effectively, identify vulnerabilities earlier, test fixes faster, map attack paths more accurately, and prioritize the exposures most likely to matter. The issue is not whether these capabilities will be used. It is whether defenders operationalize them fast enough, and in the right way, to keep pace with adversaries.

Even where current systems remain imperfect, the direction is clear. Offensive workflows are becoming more automated, more scalable, and less dependent on elite expertise. The result is not just faster exploitation. It is broader access to offensive capability.

This matters because the defensive model in many organizations is still built around delay. Teams collect findings, sort them by severity, review them in cycles, and triage them with finite human capacity. They then need time to assess business impact, identify ownership, work through change control, and decide on patching, mitigation, or compensating controls. In this environment, every patch that cannot be applied quickly becomes a clear signal to attackers about what to target next. In a world where attack preparation and execution can move at machine-speed, those processes become increasingly misaligned with risk.

Recent public reporting and early industry analysis reinforce this shift. Frontier AI models are beginning to demonstrate autonomous vulnerability discovery and exploitation across major operating systems and browsers, compressing the patch-to-exploit window from weeks to minutes. In that environment, reactive security models break down. Discovery alone is not enough. A single attack may begin on the endpoint, transition to identity compromise, and persist in the cloud, meaning defense must be unified and cross-domain. The risk is not limited to internet-first intrusion paths. As open source components, internal dependencies, and AI infrastructure expand, organizations must also prepare for attacks that begin inside trusted environments and reduce available attack opportunities. Defenders need systems that can perceive, prioritize, intervene, remediate, and validate continuously.

Anthropic’s Mythos frontier model and OpenAI's GPT-5.4-Cyber should be viewed in that context. They are not the end state, but early indicators of the power of frontier AI and the acceleration of offensive capabilities it enables. The broader implication is clear: The gap between discovery and exploitation will continue to shrink unless defenders redesign how they prioritize, validate, and respond.

The Core Shift: From Managing Vulnerabilities to Managing Exposure and Risk

The old model optimized for volume. Find more vulnerabilities. Scan more assets. Prioritize by CVSS. Reduce backlog. That model was already straining under the weight of modern environments. Frontier AI makes its limitations even more obvious.

Organizations do not need more findings if those findings cannot be converted into better decisions. AI will help surface more vulnerabilities than most teams can ever process. The question is not how many exposures exist. The question is which ones can actually be used against the organization in a way that creates meaningful business impact.

This requires a move from vulnerability management to exposure management.

Threat intelligence is essential to making that shift. It helps organizations move beyond theoretical severity by showing which vulnerabilities align to real attacker behavior, which tradecraft is active, and which exposures are most likely to be weaponized in the current environment. In a world where AI can accelerate discovery and exploitation, threat intelligence becomes the link between raw findings and meaningful prioritization.

Exposure management asks different questions. Is the vulnerable asset reachable? Is there a viable attack path? Are there identity weaknesses that make exploitation more consequential? Can the issue be chained with another weakness to produce privilege escalation, lateral movement, or persistence? Is there evidence that adversaries are targeting this technique in the wild?

This is the lens security programs need now. Risk must be measured in terms of real-world attack paths, adversary behavior, and environment-specific conditions, not theoretical severity alone. Otherwise, defenders will drown in findings while attackers continue to focus only on the few exposures that actually matter.

The market is already moving in this direction. Exposure management platforms, validation tools, and aggregators are all trying to solve the same problem. Organizations do not need more findings. They need to know which risks are actually exploitable in their environment and what to do next. Discovery is becoming easier. The real differentiators are validation and action.

Five Requirements for Closing the Gap

1 Measure what matters: exploitability

As AI increases the rate of discovery, organizations will face an even greater flood of vulnerabilities, weak configurations, and potential attack paths. The real challenge is determining which exposures matter most and what action to take first. Teams need to know not just what exists, but what can actually be weaponized in their environment.

This is where exposure management becomes essential. It helps organizations understand where they are exposed, which weaknesses are actually reachable and exploitable, and the most effective way to reduce risk through patching, mitigation, or compensating controls. Just as importantly, it helps validate that those actions actually worked. As the exploit window compresses, organizations need faster patch cycles, more automation, and pre-approved mitigation steps to reduce risk while permanent fixes are underway. In practice, this means preparing not only for more vulnerabilities but for more fixes, more urgent patch decisions, and more operational pressure on remediation teams.

In this environment, severity alone is no longer enough. Security teams need a way to assess whether an exposure is reachable, chainable, and likely to be targeted. They need to understand the role of identity, privilege, segmentation, and workload context in determining whether a vulnerability is merely present or truly dangerous.

A mature program reduces noise by ranking exposures according to operational risk. That means combining asset criticality, reachability, identity pathways, attacker tradecraft, and signs of active exploitation. In practice, the most important vulnerability is rarely the one with the highest score. It is the one most likely to become a breach.

Traditional metrics such as CVSS and patch SLAs are becoming less meaningful on their own. As AI-assisted exploit generation reduces the time between disclosure and weaponization, even issues that appear moderate on paper may become urgent in practice. Teams need to prioritize based on exploitability, asset value, and attack path relevance, not severity alone. When patching cannot happen immediately, they also need to know which mitigations and compensating controls will reduce risk fastest.

2 Continuously validate exposure from the “inside out” and “outside in”

Periodic scanning and external advisories are no longer enough. Organizations need more than a point-in-time view of exposure. They need a current, accurate understanding of how exposures change, connect, and create viable attack paths across the environment.

In a faster offensive environment, organizations need continuous validation of their own exposure. That means moving from static assessments to a current, inside out understanding of what assets exist, what weaknesses are present, how those weaknesses connect, and which conditions create viable attack paths.

This matters because attackers do not assess risk the way defenders often do. They do not care about reporting cadence or inventory completeness. They care about what is reachable right now, what can be chained, and what yields access. Defenders need to evaluate exposure the same way, and they need to do it continuously.

That requires integrating internal telemetry, configuration state, identity relationships, network reachability, and workload behavior into a unified model of exposure. It also requires testing assumptions continuously. Controls that look strong on paper may fail in practice. Segmentation may not be enforced consistently. Privileged access may be broader than expected. Exposure management must become dynamic, evidence-based, and specific to the environment.

Continuous validation should confirm not only that an exposure is real, but whether existing controls, detections, and response processes are effective against the attack paths most likely to matter. For many organizations, the most useful form of validation is understanding attack paths, identifying likely routes of lateral movement, and pinpointing the choke points where an adversary can be stopped. This is often more actionable than full simulated exploitation, especially for teams trying to prioritize risk across large, complex environments.

Continuous validation also depends on aggregation. Most organizations are not abandoning existing scanners, configuration management databases (CMDBs), posture tools, or remediation workflows. Instead, they need a way to unify fragmented exposure data across on-premises, SaaS, cloud, applications, identity, and the external attack surface into a single view of risk. This view must also account for the AI supply chain, internal dependencies, and the telemetry coverage gaps that create blind spots for prevention, detection, and response.

3 Design for prevention, identity control, and containment with zero standing privileges

More vulnerabilities are inevitable. The real question is whether they lead to impact.

Organizations should assume that some exposures will remain unresolved for some period of time. The defensive objective is to make exploitation harder, reduce the chance of meaningful access, and contain adversaries before they can move laterally or escalate privileges.

Identity sits at the center of this problem. Many successful attacks do not end with the initial exploit. They become dangerous when they allow an adversary to assume a trusted identity, obtain credentials, or abuse excessive privileges. That means prevention is no longer just about patching. It requires a commitment to continuous identity, transforming the security posture of all identities (human, non-human, and AI) from a point-in-time decision into a real-time control system. It includes enforcing zero standing privileges, continuously verifying access, limiting credential exposure, and connecting identity posture to endpoint and workload context in real-time.

Containment must also be deliberate by design. If an attacker reaches a vulnerable system, what stops them from moving further? What prevents them from accessing sensitive workloads, assuming privileged roles, or persisting across environments? Continuously evaluating identity risk is a critical Zero Trust control, limiting lateral movement by enforcing least privilege and dynamically restricting access to applications and resources. This approach is highly effective at mitigating credential-based breaches. Organizations that tie identity signals, endpoint controls, and workload telemetry together are better positioned to break the attack chain before impact. This becomes even more important as attacks increasingly move through trusted software, internal systems, AI infrastructure, and other inside out paths that reduce traditional prevention opportunities.

In many environments, the most realistic path is not immediate remediation but risk reduction through compensating controls. The critical question is not just whether a vulnerability exists, but whether existing controls can prevent exploitation from becoming meaningful access. Continuous identity principles — dynamic access, zero standing privileges, and a unified control plane — become especially important when patching cannot happen within the shrinking exploitation window.

Cyber resiliency is also becoming foundational. As exploitation windows shrink, rapid recovery, low-disruption patching, and containment without business interruption become core defensive requirements. In a frontier AI threat model, resiliency is no longer a differentiator layered on top of prevention. It is part of prevention.

4 Operate at machine-speed across detection and response

As discovery accelerates, response must keep pace.

The traditional model separates detection, investigation, and containment into distinct stages, often with handoffs and delays between them. That is increasingly untenable. Defenders need a continuous operating pipeline that moves from signal to context to action with minimal delay.

Machine-speed response does not mean eliminating humans. It means ensuring the system can gather context, correlate signals, and initiate appropriate actions fast enough to matter. Detection must account for activity across endpoints, identities, and cloud environments. Investigation must reconstruct attack chains quickly enough to support decisive action. Containment must happen in minutes, not hours, when the risk justifies it.

This matters because modern intrusions are rarely isolated to one domain. A single attack may begin with an exposed asset, transition into credential abuse, pivot into identity compromise, and establish persistence in cloud infrastructure. Operating at machine-speed means connecting these cross-domain signals fast enough to disrupt the adversary before they achieve their objective.

Speed matters not only in detection and containment, but in decision-making. Security teams need to know who owns the issue, what action is possible, whether remediation has succeeded, and what compensating controls are available if it has not. Effective operations depend on compressing both analysis and mobilization, not just alert handling.

This also requires breaking down the traditional separation between the SOC and vulnerability or exposure management teams. In many organizations, the SOC may have the data, but plays little role in prioritization, remediation, or validation. That model is too slow for a machine-speed threat environment. Security operations and exposure management must become more unified, and agentic SOC capabilities will be increasingly important in helping organizations connect detection, prioritization, remediation, and validation at speed.

AI-assisted discovery does not just increase the number of findings. It exposes the weakness of partial workflows. Security teams cannot afford a model where one tool identifies risk, another files a ticket, and a third eventually validates the fix. The operating model has to become continuous: detect, prioritize, remediate, and validate in one loop. Products that stop at alerting will increasingly fail to keep pace with adversaries operating at machine-speed.

Faster detection alone is not enough. Defenders need automated response: the ability to disrupt exploitation, enforce containment, and validate remediation before an attacker can reuse the same path.

5 Apply AI with control and intent

AI is part of the answer, but it also expands the attack surface and introduces new forms of risk.

Organizations need AI to scale analysis, prioritization, and response. It can help compress investigation timelines, identify likely attack paths, and reduce manual effort in triage. AI delivers the most value when it is applied deliberately inside governed workflows, not left to operate without oversight.

This means embedding AI into workflows where it augments human decision-making, while maintaining governance, policy controls, and oversight. Organizations also need visibility into shadow AI tools, models, and agents, since unmanaged AI adoption can expand the attack surface and introduce new security risks. It also means securing the AI stack itself: monitoring how models are used, governing the tools and systems AI agents can access, restricting unauthorized access, validating outputs, and designing controls around prompt injection, model misuse, and sensitive data leakage.

The organizations that benefit most from AI will not be the ones that deploy it everywhere first. They will be the ones that apply it deliberately, align it to real operational needs, and keep it governed from the start.

The most useful applications of AI today are often operational rather than fully autonomous. AI is proving valuable in explaining risk, summarizing attack paths, identifying likely asset ownership, recommending mitigation options, validating whether patches were executed, and confirming whether remediation actually reduced exposure. Its value comes from accelerating existing workflows and improving decision quality, not replacing governance.

Organizations should also avoid overcommitting to a single AI provider or model architecture. The most durable approach is model-agnostic by design, with the flexibility to route workloads across multiple models based on trust, governance, and task suitability. This reduces lock-in risk and allows security teams to adapt as frontier models evolve.

Mapping the Five Requirements to CrowdStrike

1 Measure what matters: exploitability

Customer Challenge

Teams are overwhelmed by findings and cannot tell which exposures are reachable, exploitable, or most urgent.

Desired Outcome

Understand where the organization is exposed, identify which exposures are exploitable, take the most effective action through patching or compensating controls, and prove that those actions reduce risk.

How CrowdStrike Delivers This Requirement

CrowdStrike closes the gap between exposure and action with real-time insight and exploitability-driven prioritization. This helps teams focus on the risks most likely to be weaponized and respond faster through patching or mitigation.

2 Continuously validate exposure from the “inside out” and “outside in”

Customer Challenge

Periodic scanning is not enough. Customers need attack path context, choke-point analysis, and continuous validation across fragmented tools and environments.

Desired Outcome

Continuously validate what is exposed, what is reachable, how an attacker could move, and whether controls can block the attack paths that matter most.

How CrowdStrike Delivers This Requirement

CrowdStrike helps customers aggregate and contextualize exposure across assets, identities, cloud, and external attack surface, supporting passive validation and continuous risk assessment.

3 Design for prevention, identity control, and containment with zero standing privileges

Customer Challenge

Even when vulnerabilities are known, organizations often lack the controls to stop exploitation from turning into privileged access, lateral movement, or persistence.

Desired Outcome

Stop exploitation earlier, continuously verify trust, and contain attackers before impact spreads.

How CrowdStrike Delivers This Requirement

CrowdStrike links endpoint, identity, and workload context to strengthen prevention, enforce zero standing privileges, and enable real-time containment, including when compensating controls are required.

4 Operate at machine-speed across detection and response

Customer Challenge

Detection, investigation, ownership, and containment are too slow and disconnected, giving attackers time to move across endpoint, identity, and cloud.

Desired Outcome

Move from signal to context to action faster, reducing time to contain from hours to minutes.

How CrowdStrike Delivers This Requirement

CrowdStrike unifies detections, investigation, automation, and response across domains, helping teams detect, investigate, assign, and contain threats faster.

5 Apply AI with control and intent

Customer Challenge

Customers want AI to improve explainability, workflow orchestration, and remediation validation without creating new governance gaps or unmanaged risk.

Desired Outcome

Embed AI into workflows to improve speed and scale while maintaining control, oversight, and policy enforcement.

How CrowdStrike Delivers This Requirement

CrowdStrike applies AI to scale prioritization, detection, investigation, and response while keeping humans in control and aligning automation to governed workflows.

What a Modern Defensive Model Looks Like

Closing the exposure window requires a more proactive operating model.

1. Organizations need a unified understanding of risk that combines asset criticality, exposure, identity, high-fidelity behavioral signals, and adversary context. Without it, prioritization will remain fragmented and reactive.

2. Validation must be continuous. Teams need to know not just what they are exposed to, but how those exposures could actually be used against them. Static reporting will not keep pace with machine-speed adversaries.

3. Security teams need controls that reduce blast radius even when prevention is imperfect. This requires adopting a continuous identity strategy where limiting credential exposure, enforcing zero standing privileges, and implementing real-time authorization and policy enforcement are no longer secondary disciplines, they are core to blocking exploits before lateral movement or privilege escalation can occur.

4. Detection and response need to function as one pipeline. High-fidelity context and rapid action matter more than sheer alert volume. The goal is not to process more alerts. It is to reach the right conclusion and contain the threat faster.

5. AI must be incorporated into security operations as an amplifier, not a replacement for judgment. Human expertise remains essential, especially when distinguishing benign anomalies from true malicious intent. AI helps scale the process. It does not remove the need for disciplined operational control. Together, these capabilities shift security from reactive workflows to a more proactive model built to anticipate, reduce, and contain risk before it becomes impact.

This shift is not only technical. It is organizational. Many of the constraints that slow patching, mitigation, and containment today are rooted in business process, ownership ambiguity, change approval friction, and the difficulty of aligning security, IT, engineering, and operations teams around urgent risk reduction. In the frontier AI era, organizations will need a new operating mindset that treats rapid exposure reduction as a shared business priority, not just a security function.

This is also where the market is moving. Customers are no longer asking only for better discovery. They are asking for validation, context, and action. They want to understand how an attacker would move through their environment, which controls can block that path, who owns the risk, and what to do when patching is not immediately possible. In this model, the value of a platform is not the number of findings it produces. It is its ability to turn fragmented exposure data into prioritized, actionable decisions at machine-speed.

WHAT FRONTIER AI MEANS FOR BUSINESS RISK

Frontier AI is changing more than the speed of cyberattacks. It is changing the amount of time organizations have to identify, assess, and reduce risk before an exposure becomes an incident. As the gap between discovery and exploitation shrinks, the risk is no longer limited to security teams falling behind on patching. It becomes a broader business problem that affects operational resilience, change management, incident readiness, and the ability to protect critical systems without disrupting the business.

For business leaders, the implication is clear: This is not just a technical shift. It is a change in the operating environment. Organizations that cannot prioritize real risk, move quickly across security and IT workflows, and contain threats before they spread will face greater exposure to disruption, data loss, recovery costs, and reputational damage. The organizations best positioned for this shift will be those that treat exposure reduction, identity control, and machine-speed response as business priorities, not just security priorities.

Questions Security Leaders Should Be Asking

Security leaders should assess whether their current programs are designed for this new environment. A few questions can help frame the issue:

- Are we **prioritizing exposures based on exploitability in our environment**, or still relying mainly on severity and backlog reduction?
- Are **we maintaining a complete inventory of assets**, including shadow technology, developer-provisioned resources, SaaS dependencies, and AI systems?
- Are we **keeping critical systems current** and patching quickly enough to reduce real exposure?
- Are we **scanning continuously across environments and development pipelines**, or still relying on periodic assessments?
- Are we **connecting endpoint, identity, and cloud context** quickly enough to understand the full attack path?
- Are our **prevention and identity controls, including zero standing privileges**, strong enough to stop an exposure from turning into lateral movement, privilege escalation, or breach?
- Are we **stress-testing incident response under realistic conditions**, and are our **detection and response workflows** moving fast enough to matter?
- Are we **clear on who owns the risk**, what mitigation options are available, and whether remediation actually worked?
- Are we **governing AI appropriately** by monitoring AI assets, validating AI-generated code, testing for abuse, and enforcing guardrails around deployment and use?
- Are we **using AI to improve security operations** while maintaining clear governance and control?

If the answer to these questions is inconsistent or unclear, that is usually the real signal. The problem is not a lack of data. It is a mismatch between the speed of the threat and the design of the program.

Conclusion

Frontier AI is not simply making attackers more efficient. It is reshaping the defensive timeline. The interval between discovery and exploitation is compressing, and the cost of operating with slow, fragmented, or volume-driven security processes is rising.

Anthropic's Mythos and OpenAI's GPT-5.4-Cyber serve as early indicators of what this future looks like. They will not be the last. Offensive capability will keep improving in sophistication, accessibility, and speed. Organizations cannot respond by trying to out-count the problem. They need to out-prioritize, out-validate, and out-operate it. They also need to be prepared for a world in which vulnerability discovery, patch pressure, and attack execution all accelerate at the same time.

This starts with five disciplines: focus on exploitability, continuously validate exposure, design for prevention and identity-centric control, operate at machine-speed, and apply AI deliberately. Organizations that get these right will be better positioned to absorb the next wave of offensive capability without surrendering the initiative.

Meeting this challenge requires more than better discovery. It requires a platform that can connect exploitability, identity, telemetry, and response in one operating model. It also requires organizations to treat frontier AI not only as a source of new risk, but as a tool to improve how they identify, prioritize, and reduce exposure CrowdStrike is uniquely positioned to help by combining frontline adversary intelligence, cross-domain visibility, machine-speed automation, and integrated protection across endpoint, identity, and cloud.

Frontier AI is shrinking the time between discovery and attack. Defenders need a faster way to respond.

Learn more: https://www.crowdstrike.com/en-us/solutions/frontier-ai-readiness-requirements/

Disclaimer: This white paper includes discussion of unreleased services and features. Any references to unreleased features reflect our current plans only and do not constitute a promise or commitment to deliver such features. These items may change or may not be made available in all regions. Customers should make purchase decisions based on features currently available.

About CrowdStrike

CrowdStrike (NASDAQ: CRWD), a global cybersecurity leader, has redefined modern security with the world's most advanced cloud-native platform for protecting critical areas of enterprise risk — endpoints and cloud workloads, identity and data.

Powered by the CrowdStrike Security Cloud and world-class AI, the CrowdStrike Falcon platform leverages real-time indicators of attack, threat intelligence, evolving adversary tradecraft, and enriched telemetry from across the enterprise to deliver hyper-accurate detections, automated protection and remediation, elite threat hunting, and prioritized observability of vulnerabilities.

Purpose-built in the cloud with a single lightweight-agent architecture, the Falcon platform delivers rapid and scalable deployment, superior protection and performance, reduced complexity and immediate time-to-value.

CrowdStrike: We stop breaches.

Learn more: https://www.crowdstrike.com/en-us/

Start a free trial: https://www.crowdstrike.com/free-trial-guide/

Follow us: Blog | X | LinkedIn | Facebook | Instagram