AI Integration: CrowdStrike Incident Response Strategy

Camilo Castrillon, [redacted]

Team Six

Information Security Strategies & Policy - CS 6725

Professor [redacted]

19 March 2026

AI Integration: CrowdStrike Incident Response Strategy

Executive Summary

Artificial Intelligence threatens to change the cybersecurity landscape by exponentially increasing the speed of cyberattacks, resulting in the subsequent need for faster incident response. CrowdStrike now faces a threat environment in which adversaries can move laterally through entire networks in as little as 27 seconds. This report encourages CrowdStrike to continue expanding AI within incident response, particularly with enhancements to its current tools: Charlotte AI and the Falcon platform. AI has proven to significantly reduce manual analyst workload and breach lifecycles; these advantages elevate both operational efficiency and customer protection. At the same time, increased use of AI introduces serious risks, including hallucinations, prompt injection, and adversarial manipulation, among other risks. If not governed properly, these assets may lead to adverse effects, such as legal, security, and compliance concerns.

In light of this new era of threats, CrowdStrike should adopt a uniform approach that treats AI as a tool rather than a replacement for human judgment. AI should be supported by rigorous model validation. Much of AI's effectiveness depends on the quality of the underlying model, thereby reinforcing the need for continuous evaluation rather than unchecked expansion. Presently, AI should be applied to high-value areas such as automated threat detection, alert triage, and manual analysis, as human oversight remains essential to ensuring that automated decisions do not create unintentional security consequences. Moving forward, CrowdStrike's long-term success will ultimately hinge on how well it can incorporate and scale AI to deliver clear operational gains and maintain a strong posture in regulatory alignment and accountability.

Current Threat Landscape & AI Integration

The speed of cybersecurity attacks has steadily increased, with average breakout time (the time for an adversary to move laterally through a network) decreasing from roughly two hours in 2017 to 29 minutes in 2025, with the fastest recorded breakout at 27 seconds [1]. While organizations are incentivized to rapidly adopt AI to match this acceleration, doing so introduces risks if systems are deployed without sufficient validation. AI removes traditional bottlenecks in attack stages such as reconnaissance, phishing, and data exfiltration, enabling both defenders and adversaries to operate at unprecedented speed. This dual-use nature of AI creates a cultural tension: the same capabilities that enable faster defense also empower more sophisticated attacks.

AI also introduces new attack vectors, particularly through social engineering. One prominent example occurred in 2024 at the UK engineering firm Arup, where attackers used AI-generated video impersonations of executives to convince an employee to initiate 15 fraudulent transfers totaling $25 million [2] [3]. Such cases highlight that while rapid AI deployment may improve efficiency, insufficient safeguards can amplify organizational vulnerability. As attackers leverage AI to scale deception, CrowdStrike must balance accelerating defensive capabilities with ensuring system reliability and resilience.

AI can assist organizations in responding to threats, particularly by improving breach cost and breach lifecycle outcomes. According to IBM's Cost of a Data Breach Report, breaches with shorter lifecycles cost significantly less ($3.87 million versus $5.01 million) [4]. This creates a strong pressure to deploy AI quickly to reduce detection time. However, increased incident volume [5] means that overreliance on automated systems without proper oversight may introduce false positives or missed threats. While AI has proven effective in identifying breaches with high accuracy in both academic and industry settings [6], its probabilistic nature necessitates careful integration. Thus, while AI offers clear efficiency gains, organizations must ensure that speed does not come at the expense of accuracy and accountability in incident response.

AI for Incident Response: Internal AI Tools

CrowdStrike's Charlotte AI, built on the Falcon platform, is embedded into the incident response ecosystem, triaging alerts, automating investigations, and assisting analysts through natural language. This integration reportedly saves over 40 hours per week in manual triage work [7]. While this demonstrates the efficiency gains of rapid AI adoption, it also raises concerns about over-automation if outputs are not sufficiently validated before action.

Charlotte AI operates as a multi-agent system, where specialized agents handle tasks such as data retrieval, script generation, and threat intelligence analysis [8]. This architecture helps mitigate hallucination risk by introducing a validation agent that checks outputs against Falcon data before presenting them to analysts [9]. Though this design supports responsible deployment, it also reflects a deliberate tradeoff: adding validation layers may slightly slow system outputs but improves reliability and trustworthiness, reinforcing the need to balance speed with accuracy.

A primary cybersecurity risk for internal AI tools is prompt injection, ranked by OWASP as the top vulnerability for LLM applications [10]. Following its acquisition of Pangea, CrowdStrike has analyzed 300,000 adversarial prompts and 150 injection techniques [11], demonstrating the scale of defensive investment required to responsibly deploy

AI. This highlights that rapid AI-scaling must be matched with equally robust security testing to avoid introducing exploitable weaknesses.

Additional safeguards are necessary to maintain a defense-in-depth approach. Traditional cybersecurity layers, such as network segmentation, endpoint detection, and manual analysis, must continue to supplement AI-driven workflows so that failures in systems like Charlotte AI do not become single points of failure [12]. Human-in-the-loop oversight is particularly critical for high-impact actions, as AI outputs are probabilistic and may misclassify threats [Mohsin et al., 2025]. Studies show that while 93% of security professionals believe generative AI improves capabilities, 71% still prefer human review before action [13]. As such, while autonomous AI can significantly accelerate response times, responsible deployment requires retaining human oversight to prevent cascading errors.

AI for Incident Response: Customer-Facing AI

Customer-facing AI has transformed how companies detect and respond to incidents by enabling faster, real-time threat mitigation. AI allows organizations to "streamline their response strategies, allowing for quicker detection and remediation of incidents" [14]. This creates a strong incentive for rapid deployment, particularly in environments where response and speed directly impact damage mitigation. However, deploying customer-facing AI systems too quickly without sufficient testing may expose users to inaccurate outputs or incomplete protections.

AI-driven cybersecurity has also shifted risk management toward a more proactive model. Khadka (2024) describes this as a transition to predictive and preventative security, supported by tools such as CrowdStrike's Falcon Adversary Intelligence Premium, which uses AI to anticipate and counter cyberattacks [15]. Although these tools enhance efficiency and scalability, their effectiveness depends on the quality and reliability of underlying models, thereby reinforcing the need for continuous evaluation rather than unchecked expansion.

CrowdStrike reported that "AI-driven solutions contributed to a 25% increase in protected endpoints over the past fiscal year, highlighting operational reliance on automated analytics" [16]. While this demonstrates the benefits of scaling AI deployment, it also increases systemic risk if models fail or are compromised. As a result, companies must balance expanding AI capabilities with ensuring that systems remain accurate, secure, and resilient under evolving threat conditions.

AI in Cybersecurity: Legal & Ethical Considerations

The integration of AI into enterprise cybersecurity platforms has significantly enhanced the ability of security operations centers to detect and respond to threats; "Over 60% of enterprises integrating AI into cybersecurity report faster detection times and fewer missed incidents" [17]. While this creates pressure to accelerate AI adoption, it also expands legal and ethical responsibilities, particularly as AI systems process sensitive data and may initiate automated actions. The same capabilities that improve efficiency also increase the consequences of system errors, underscoring the need to consider governance in the deployment of these tools.

AI-enabled systems can process massive volumes of data generated by endpoints, cloud infrastructure, and network activity, allowing security teams to identify anomalous behavior and respond to threats more rapidly than traditional rule-based systems. Specifically, "AI supports threat intelligence by continuously processing system logs, network events, and external threat feeds to identify indicators of compromise" [18] [19]. Within platforms like CrowdStrike Falcon, tools such as Charlotte AI assist analysts with triage and investigation [20]. However, reliance on automated analysis introduces risks, as false positives or false negatives can undermine incident response and "jeopardize the integrity" of findings [21] [22]. Consequently, the growing reliance on AI in cybersecurity raises important concerns regarding the tension between efficiency and accuracy: faster systems are valuable, but only if their outputs are dependable.

In the United States, frameworks such as the NIST Artificial Intelligence Risk Management Framework and the NIST Cybersecurity Framework 2.0 emphasize structured risk management, transparency, and system security [22]. The NIST AI Risk Management Framework explains that "AI risk management offers a path to minimize potential negative impacts of AI systems while maximizing their benefits," supporting the development of trustworthy AI systems that are reliable, secure, and accountable [22]. Executive Order 14110 on Safe, Secure, and Trustworthy Artificial Intelligence outlines the need for testing, risk assessments, and safeguards for high-impact AI systems [23]. For CrowdStrike, these governance expectations translate into operational requirements: ensuring explainability, resilience to adversarial attacks, and human oversight in automated decision-making. These measures may slow deployment timelines but are necessary to ensure compliance and trust [12].

Ultimately, trustworthy AI systems must be "valid and reliable, safe, secure and resilient, accountable and transparent, explainable and interpretable, and privacy-enhanced" [22]. In AI-driven incident response, failures can disrupt critical infrastructure or expose sensitive data. Therefore, organizations must align innovation

with governance through continuous monitoring and human oversight, ensuring that increased speed does not compromise legal compliance or operational integrity.

Future Considerations: Policy Updates

As AI becomes more integrated into CrowdStrike's incident response operations, the company must anticipate new attack vectors, model vulnerabilities, and operational risks that are not fully addressed by existing cybersecurity policies [24]. **To strengthen the security and governance of AI-driven systems, several forward-looking policies should be considered:**

1. Third-party AI models should be rigorously vetted for backdoors, malicious behavior, and training data quality [25].
2. In-house AI development should include bias assessments, anonymization of personally identifiable information in training datasets, and verification of model supply chains.
3. Robust testing protocols, including adversarial testing, input fuzzing, and scenario simulations, can identify vulnerabilities in AI models before deployment [22]. Opposition to these measures due to impacts on speed, these steps will reduce long-term risk and improve system trustworthiness.
4. AI should be supplemented with traditional cybersecurity layers to maintain defense-in-depth [26].
5. Human-in-the-loop oversight should remain standard for critical actions, acknowledging that AI outputs are probabilistic [12].
6. Fostering interdisciplinary communication, or a DevSecOps approach across SOC analysts, AI engineers, and legal/compliance teams, ensures that vulnerabilities and risk mitigation strategies are identified and addressed throughout [27]. DevSecOps integrates security throughout the software and infrastructure lifecycle, so that security measures are "built in" rather than added later in the process [27].

Establishing an AI Governance Task Force composed of representatives from each relevant team would further support continuous oversight and facilitate proactive risk management. Together, these policies reflect a measured approach, enabling CrowdStrike to scale AI capabilities while maintaining the safeguards necessary for responsible and secure deployment.

Conclusion

While it may be tempting and lucrative for a cybersecurity company to focus primarily on rapid adoption of AI for efficiency and automation, it is critical to maintain a focus on secure and responsible implementation. For CrowdStrike, AI should be leveraged

in key areas such as automated threat detection, alert triage, and incident response, while maintaining a human-in-the-loop approach to address potential biases, inaccuracies, or unexpected model behavior.

Best practices include rigorous model validation, secure data handling, and mitigation of bias and adversarial risks. Human-in-the-loop oversight remains critical to ensure that automated decisions do not result in unintended consequences. Additionally, across applications, compliance with U.S. frameworks such as NIST AI RMF, CSF 2.0 is imperative to maintaining trust and regulatory alignment. In all, the integration of AI into incident response is not solely a technical challenge but a governance challenge. By balancing rapid innovation with comprehensive policies, structured oversight, and cross-agency collaboration, CrowdStrike can continue to leverage AI while ensuring that security, accuracy, and responsibility remain central to its operations.

References

[1] CrowdStrike, "Global Threat Report," CrowdStrike, [Online]. Available: CrowdStrike Global Threat Report.

[2] World Economic Forum, "Cybercrime: Lessons learned from a $25m deepfake attack," Feb. 4, 2025. [Online]. Available: https://www.weforum.org/stories/2025/02/deepfake-ai-cybercrime-arup/.

[3] Financial Times, "Arup lost $25mn in Hong Kong deepfake video conference scam," May 16, 2024. [Online]. Available: https://www.ft.com/content/b977e8d4-664c-4ae4-8a8e-eb93bdf785ea.

[4] IBM, "Cost of a Data Breach Report," IBM, [Online]. Available: https://www.ibm.com/forms/mkt-53830.

[5] SentinelOne, "Key Cyber Security Statistics for 2026," Jan. 15, 2026. [Online]. Available: https://www.sentinelone.com/cybersecurity-101/cybersecurity/cyber-security-statistics/.

[6] M. Rahmati, "Towards Explainable and Lightweight AI for Real-Time Cyber Threat Hunting in Edge Networks," *arXiv preprint* arXiv:2504.16118, Apr. 2025. [Online]. Available: https://arxiv.org/abs/2504.16118.

[7] E. Zaitsev, "Agentic AI innovation in cybersecurity: Charlotte AI detection triage," CrowdStrike, Feb. 13, 2025. [Online]. Available: https://www.crowdstrike.com/en-us/blog/agentic-ai-innovation-in-cybersecurity-charlotte-ai-detection-triage/.

[8] Exabeam, "CrowdStrike Charlotte AI: Solution Overview, Pros and Cons," *Exabeam Explainers*, [Online]. Available: https://www.exabeam.com/explainers/crowdstrike/crowdstrike-charlotte-ai-solution-overview-pros-and-cons/.

[9] OWASP, "LLM01 – Prompt Injection," OWASP GenAI Risk, [Online]. Available: https://genai.owasp.org/llmrisk/llm01-prompt-injection.

[10] J. Gamble, "Indirect prompt injection attacks: Hidden AI risks," CrowdStrike, Dec. 4, 2025. [Online]. Available: https://www.crowdstrike.com/en-us/blog/indirect-prompt-injection-attacks-hidden-ai-risks/.

[11] MIT Cybersecurity and Internet Defense, "Cybersecurity defense-in-depth: layered strategies for comprehensive protection," [Online]. Available: https://cyberir.mit.edu/site/cybersecurity-defense-depth-layered-strategies-comprehensive-protection/.

[12] A. Mohsin, H. Janicke, A. Ibrahim, I. H. Sarker, and S. Camtepe, "A Unified Framework for Human–AI Collaboration in Security Operations Centers with Trusted Autonomy," *arXiv preprint* arXiv:2505.23397 [cs.SI], May 29, 2025. [Online]. Available:

https://arxiv.org/pdf/2505.23397.

[13] A. Tan, "How CrowdStrike is leveraging AI to empower security teams," *Computer Weekly*, Aug. 2, 2024. [Online]. Available: https://www.computerweekly.com/news/366599716/How-CrowdStrike-is-leveraging-AI-to-empower-security-teams.

[14] H. Umar and R. Marchant, "Effective response to data breaches: AI-assisted solutions for modern enterprises," Oct. 2023. [Online]. Available: https://www.researchgate.net/profile/Robert-Marchant-3/publication/384227098_Effective_Response_to_Data_Breaches_AI-Assisted_Solutions_for_Modern_Enterprises/links/66eed7b76b101f6fa4fac1fe/Effective-Response-to-Data-Breaches-AI-Assisted-Solutions-for-Modern-Enterprises.pdf.

[15] K. Khadka, *Forecasting risks, challenges, and innovations: A cybersecurity perspective*, Dec. 21, 2024. [Online]. Available: http://dx.doi.org/10.2139/ssrn.5187659.

[16] CrowdStrike Holdings, Inc., "CrowdStrike reports fourth quarter and fiscal year 2026 financial results," Investor Relations, 2026. [Online]. Available: https://ir.crowdstrike.com/static-files/a315d192-32e0-4bb2-8be2-189336f80f3d.

[17] A. Devasia, "AI cyber threat statistics: The 2025 landscape of AI-powered cyberattacks," *The Network Installers*, Dec. 2, 2025. [Online]. Available: https://thenetworkinstallers.com/blog/ai-cyber-threat-statistics/.

[18] Cyble, "What is Cyber Threat Intelligence?," *Cyble Knowledge Hub*, [Online]. Available: https://cyble.com/knowledge-hub/what-is-cyber-threat-intelligence/.

[19] T. D. Le, T. Le-Dinh, and S. Uwizeyemungu, "Cybersecurity analytics for the enterprise environment: A systematic literature review," *Electronics*, vol. 14, no. 11, p. 2252, 2025, doi: 10.3390/electronics14112252.

[20] CrowdStrike, *Charlotte AI: Generative AI for cybersecurity operations*, 2025. [Online]. Available: https://www.crowdstrike.com/platform/charlotte-ai/

[21] S. Sudhakaran and N. Kshetri, "AI agents vs. human investigators: Balancing automation, security, and expertise in cyber forensic analysis," *arXiv preprint* arXiv:2601.14544, Jan. 2026. [Online]. Available: https://arxiv.org/abs/2601.14544.

[22] National Institute of Standards and Technology, *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, Gaithersburg, MD, USA: NIST, 2023. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

[23] The White House, "Executive Order 14110: Safe, secure, and trustworthy development and use of artificial intelligence," Oct. 30, 2023. [Online]. Available:

https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/.

[24] CrowdStrike, "2026 CrowdStrike Global Threat Report: AI accelerates adversaries and reshapes the attack surface," CrowdStrike Press Releases, Feb. 24, 2026. [Online]. Available: https://www.crowdstrike.com/en-us/press-releases/2026-crowdstrike-global-threat-report/.

[25] J. Klearman, "CrowdStrike and the future of AI in cybersecurity," *GraniteShares Research*, Nov. 2024. [Online]. Available: https://graniteshares.com/institutional/us/en-us/research/crowdstrike-and-the-future-of-ai-in-cybersecurity/.

[26] D. Kidd, "Cybersecurity defense-in-depth: layered strategies for comprehensive protection," *MIT Cybersecurity and Internet Defense*, [Online]. Available: https://cyberir.mit.edu/site/cybersecurity-defense-depth-layered-strategies-comprehensive-protection/.

[27] Red Hat, "What is DevSecOps?," Red Hat, [Online]. Available: https://www.redhat.com/en/topics/devops/what-is-devsecops.

Appendix

1. **AI Statement:** We have complied with the generative AI policy of this course.
2. **Link to Deliverable 2:** Here (Take-away document on Page 12)
3. **Deliverable 3** on Page 11

CrowdStrike: Audience Profile & Technical Expertise

This report is tailored for CrowdStrike's CISO, Justin Acquaro, with a secondary audience of senior leaders across security, technology, and risk operations. Given their deep familiarity with the domain, the report adopts a technical rather than purely business-oriented tone. These leaders are responsible for evaluating both the opportunities and implications of integrating AI into incident responses across operations, security, and governance dimensions. Accordingly, the report goes beyond high-level strategy to focus on issues most relevant to CrowdStrike, including the acceleration of modern attacks, triage burdens, and the latent risks of AI-driven systems. It also examines company-specific capabilities, Charlotte AI, and Falcon, to assess how leadership can balance operational value with risk when deploying AI within existing infrastructure.

Key Facts

Rise in AI-Enabled Attacks	*Average Breakout Time*	*Fastest Ever Breakout Time*
89%	**29 Min.**	**27 Sec.**

AI IS DUAL-USE: IT STRENGTHENS DEFENDERS, BUT ALSO IMPROVES AND SCALES ATTACK CAPABILITIES.

Overview

CrowdStrike faces cyberattacks that move faster than manual incident response can handle, making AI necessary for detection, triage, and investigation. Charlotte AI and Falcon improve speed and reduce analyst workload, but they also create risks such as hallucinations, prompt injection, and false decisions. **CrowdStrike should keep expanding AI in incident response under a governed human-in-the-loop framework.**

According to IBM's Cost of a Data Breach Report, **breaches with shorter lifecycles cost significantly less ($3.87 million versus $5.01 million)**. This creates a strong pressure to deploy AI quickly to reduce detection time.

Key Takeaways

Speed is the Constraint in the Era of AI-Enabled Attacks

It is getting increasingly difficult for defenses to match the pace of attackers. AI provides defenders with a reliable, solution to match the speed of attackers when responding to incidents

AI Initiatives Should Aim to Relieve Humans of Routine Tasks

The volume of attacks that defenders must respond to is increasing. AI can help efficiently defend against known attacks, which frees up humans to address novel, more complex threats.

Governance is at the Heart of Good Cybersecurity Posture

AI is not a one-time solution to building good cybersecurity posture. Human oversight and analysis of its work is essential to producing a safe and reliable tool for Crowdstrike and its clients