REQUEST FOR COMMENT RESPONSE

Singapore - Securing Agentic AI – An Addendum to the Guidelines and Companion Guide on Securing AI Systems

31 December 2025

I. INTRODUCTION

CrowdStrike welcomes the Cyber Security Agency of Singapore's (CSA) draft *Securing Agentic AI – An Addendum to the Guidelines and Companion Guide on Securing* AI *Systems* ('Addendum'). We strongly support CSA's decision to address the specific risks of agentic AI systems through a risk-based, lifecycle-oriented addendum to the 2024 Guidelines and Companion Guide on Securing AI Systems.

The Addendum is a timely and well-founded response to the emergence of self-managing, goal-driven AI systems that can plan, act and iterate across multiple steps. In particular, we welcome its capability-centric framing (cognitive, interaction and operational capabilities), its adoption of autonomy levels and workflow-based analysis, and its alignment with community efforts such as the agentic AI threat taxonomy from the Open Worldwide Application Security Project (OWASP) and the inclusion of concrete case studies. Together, these elements provide a practical mental model that reflects how agentic AI systems are actually being designed and deployed.

We also welcome the community-driven nature of the document and the public consultation process inviting feedback from both local and international stakeholders.

The draft Addendum provides a strong foundation for 'security for AI'. Our submission provides ways to deepen its operational impact, especially in the important context of 'AI for security'. It also suggests ways to ensure that the Addendum enables, rather than unintentionally constrains, the responsible use of agentic AI for cyber defence. Given the breadth of controls in the Addendum, CSA may also wish to highlight a baseline subset and an optional 'enhanced' tier as autonomy and impact increase.

II. COMMENTS

Connecting 'security for AI' with 'AI for security'

The Addendum provides a strong foundation for organisations to ensure that agentic AI systems are deployed safely and securely. This is 'security for AI'. We encourage CSA to explicitly situate 'security for AI' alongside complementary work focused on 'AI for security'. 'AI for security' refers to using AI and agents to improve cyber defence operations, and to detect and respond to threats that themselves leverage AI and agentic automation.

This framing explicitly reinforces that securing agentic AI systems is one part of a broader, coherent strategy for AI and cybersecurity, and encourages organisations to treat AI both as an asset to be protected and as an essential tool for defence.

We also note the emergence of AI Detection and Response (AIDR) solutions that monitor AI behaviour and interactions, secure AI agent identities and access, and guard against sensitive data leakage from AI systems. We recommend CSA consider referencing AIDR in the Addendum as an emerging and important component of AI for security for organisations.

Strengthen guidance on runtime detection, telemetry and incident response

The Addendum rightly identifies 'Continuous monitoring and logging' (control 4.3) as a critical control and recommends routing agent-initiated calls through a centralised middleware enforcement plane such as an API gateway, MCP gateway or service mesh, with logs streamed into a SIEM for SOC monitoring, and the ability to revoke access on anomaly. CSA should also recommend that organisations ultimately treat agentic activity logs as security telemetry.

To further bridge the current guidance in the Addendum with concrete expectations for detection, investigation and response, CSA could also include examples of useful attributes for an agentic AI log schema.

Orchestration/Enforcement planes, Telemetry and MCP/Agent-to-Agent communication

Orchestration and enforcement planes in agentic AI systems, whether implemented via central gateways, service meshes, MCP-based coordination layers or bespoke Agent-to-Agent (A2A) mechanisms, mediate human-to-agent, agent-to-agent and agent-to-tool interactions across the lifecycle. They are a key architectural component for Agentic AI deployments and are security-critical components in their own right.

We recommend that CSA explicitly encourage system owners to identify these orchestration planes in threat models and to apply strong identity, access control and change-management to orchestration policies. They should also ensure appropriate logging including which agents invoked which tools, via which MCP/A2A endpoints, under what autonomy level and with what approvals, and maintain a trusted registry of MCP servers and other A2A endpoints with appropriate isolation.

To reinforce the idea that agentic AI security involves architectural decisions in addition to model-selection decisions and prompt-engineering risks, we also recommend CSA work with stakeholders to develop a reference architecture for deploying agentic AI systems with external or high-risk capabilities (e.g. internet access, business transactions, system management); or for higher-risk, multi-tool, or multiagent systems.

Extending this even further, several of the Addendum's recommendations including taint tracing, centralised enforcement planes, continuous monitoring of agent decisions, and threat-informed detection engineering depend on rich, timely and well-governed telemetry. In our experience, such capabilities are significantly easier to implement when organisations operate a unified, AI-ready security data layer that consolidates telemetry from endpoints, identities, cloud workloads, applications and third-party tools into a single, governed model.

We suggest that CSA explicitly recognise unified data layers as an architectural enabler for securing agentic AI systems, and encourage system owners to consider platform approaches that provide both the data foundation and the governance needed for safe, large-scale agentic orchestration.

Human-led agentic SOC as an exemplar deployment model

One of the most immediate areas agentic AI can improve cybersecurity practices is by deploying agents into the Security Operations Center (SOC). By deploying specialised agents to tackle time-intensive tasks, security teams can reclaim a speed advantage, close persistent labor and response gaps, and shift from reactive to proactive defense.

This increasingly occurs through curated 'agentic security workforces' as part of a 'human-led agentic SOC'. These are sets of mission-specific agents designed, trained and governed to handle particular workflows (for example, malware analysis, vulnerability triage or fraud case review) with embedded domain expertise and pre-defined guardrails.

In a human-led agentic SOC, security analysts use specialised AI-powered agents with specific domain knowledge that automate investigations, triage, and response, while remaining under human command. Such well-governed, purpose-built security agents' behaviour and impact can be more easily tested, monitored and constrained. Applying this kind of defined bounded autonomy allows organisations to set when and how automated actions occur, keeping AI-driven automation trusted, accountable, and under human control. This model is preferable to granting generic security agents broad access to systems. The human-led agentic SOC model aligns strongly with the Addendum's emphasis on human-in-the-loop oversight for high-impact actions, careful management of autonomy levels, and clear scoping of agent capabilities.

We encourage CSA to recognise such deployments as leading examples of using AI for security, and to highlight that the Addendum's principles are directly applicable to designers of such human-led agentic SOCs.

We also recommend that CSA consider recognising the human-led agentic SOC models as an important emerging practice and the exemplar for AI for security in appropriate high-risk domains.

Threat-informed defence and mapping to adversary behaviours
As CSA is aware, adversaries are starting to use agentic orchestration to assist them with common campaign tasks. This makes it important to explicitly flag adversarial use of agentic AI as an important component of threat modelling. Organisations should include this in their threat intelligence collection requirements.

They should also include detection of agent-generated attack traffic (e.g. fast, tool-driven reconnaissance patterns) in their security monitoring strategies. OWASP's Agentic AI threat taxonomy is an excellent way to help with this. To make it even easier for security operations teams to operationalise the Addendum and incorporate agentic AI into existing risk, threat modelling and detection practices, CSA could consider encouraging organisations to leverage current threat intelligence (including AI-specific TTPs such as prompt injection campaigns, automated vulnerability scanning and agent-driven lateral movement) when updating their agentic threat risk models and detection content.

Better define shared responsibility and expectations for SaaS agentic AI
The Addendum recognises that activities such as deep taint tracing and internal threat modelling may be impractical for some SaaS deployments, and suggests using the framework to ask better questions of vendors, adopting red-teaming, and escalating residual risks for formal acceptance. To make the Addendum even more actionable for organisations heavily reliant on cloud and SaaS-delivered agentic platforms we recommend that CSA expands the guidance to include a short 'SaaS agentic AI provider expectations' checklist.

Organisations would also benefit from the inclusion of concrete examples for which controls fall primarily on the vendor versus the customer in typical SaaS scenarios in Section 4.3 (e.g. 2.1 Supply Chain Security, 2.7 Limit Agency, 3.3 Secure MCP, 4.3 Continuous Monitoring).

Global interoperability, data protection and product-neutrality
The Addendum and original Guidelines already reference resources such as MITRE ATLAS, OWASP Top 10 for LLMs, SLSA and national guidance from other authorities. CrowdStrike recommends that CSA continues emphasising interoperability with international frameworks by continuing to align the Addendum with frameworks such as the NIST AI RMF and relevant ISO/IEC standards, and to keep the core guidance product-neutral while referencing examples in non-normative materials.

The Addendum also helpfully references the Singapore PDPC Advisory Guidelines on the use of personal data in AI recommendation and decision systems. This linkage could be highlighted even more prominently where the Addendum discusses PII protection, logging, and validation of outputs.

Organisational readiness and the evolving role of the analyst

Finally, we observe that the move to agentic AI is not purely technical. The Addendum's emphasis on human-in-the-loop control, autonomy management, and clear assignment of responsibilities implies an evolution in skills and roles, especially within security operations staff. As organisations adopt agentic AI, analysts and engineers increasingly shift from executing manual tasks to designing, supervising and orchestrating fleets of agents. We recommend that CSA consider briefly acknowledging this and encouraging organisations to:

- define accountability for agent behaviour and outcomes for operators;
- invest in developing and running training for 'agent orchestrator' roles; and
- ensure that governance structures such as risk committees and change advisory boards are equipped to evaluate changes to agents, orchestration policies and autonomy levels.

III. CONCLUSION

CrowdStrike strongly supports CSA's leadership in developing practical, security-focused guidance for agentic AI systems. The Addendum's capability-centric view, use of autonomy levels, emphasis on workflows and taint tracing, and lifecycle-based control catalogue form a robust foundation that reflects the realities we see in complex enterprise environments.

We particularly welcome the Addendum as a core pillar of 'security for AI' and see clear opportunities to connect it with complementary efforts on 'AI for security' (using AI and agents to improve cyber defence operations, and to detect and respond to AI-enabled threats). In our experience, one of the most high-impact deployments of agentic AI is within security operations centres themselves, where human-led, agentic SOC models will increasingly allow analysts to orchestrate fleets of agents while retaining oversight over high-impact actions.

Unified, AI-ready security data layers and well-governed orchestration planes will be key architectural enablers for safely realising the potential of agentic AI at scale. Our recommendations are intended to deepen the operational aspects of the Addendum, especially around runtime detection and response, shared responsibility for SaaS, security-critical use cases such as agentic SOCs, and the design of orchestration and enforcement planes, while reinforcing alignment with international standards and maintaining product neutrality.

CrowdStrike appreciates the opportunity to contribute to this consultation and would be pleased to continue engaging with CSA and the wider community as the Addendum evolves and as organisations in Singapore and beyond operationalise secure, human-led agentic AI systems.

Public policy inquiries should be made to:

Drew Bagley VP & Counsel, Privacy and Cyber Policy	**Brian Fletcher** Director, Public Policy APJ

Email: policy@crowdstrike.com

IV. ABOUT CROWDSTRIKE

CrowdStrike (Nasdaq: CRWD), a global cybersecurity leader, has redefined modern security with one of the world's most advanced cloud-native platforms 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 prioritised 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.

Learn more: https://www.crowdstrike.com/.