



WHITE PAPER



Re-Imagine Enterprise-Scale Cyber Defence Through Agentic AI Platforms

Spencer Izard, PAC
July 2026

Infosys®

Contents

Executive Summary on the New Velocity of Cyber Warfare	3
The Adversarial Landscape Shaped by Weaponised Agents and Grey State Threats	5
The Paradigm Shift Redefining Augmentation Through Agentic Swarms.....	6
The Three Pillars of Agentic Defence to Enable Observability, Protection, and Trust.....	7
Exploring Cyber Next Powered by Topaz Fabric – an Agentic AI First Approach	8
About Infosys.....	10
About PAC.....	10

The creation and distribution of this white paper was supported by Infosys. For further information please visit www.pacanalyst.com.

Disclaimer

The contents of this white paper were compiled with the greatest possible care. However, no liability for their accuracy can be assumed. Analyses and evaluations reflect the state of our knowledge in July 2026 and may change at any time. This applies in particular, but not exclusively, to statements made about the future. Names and designations that appear in this paper may be registered trademarks.

Usage rights

This white paper is protected by copyright. Any reproduction or dissemination to third parties, including in part, requires the prior explicit authorization of Infosys. The publication or dissemination of tables, graphics, etc. in other publications also requires prior authorization.

Executive Summary on the New Velocity of Cyber Warfare

In recent years, PAC has observed the understandable escalation of AI-led autonomous threats, marking a fundamental shift in the operational reality for modern enterprises across all vertical sectors. Traditional defences are, at best, striving for parity and, at worst, increasingly unable to keep pace with attacks orchestrated by highly sophisticated agentic AI. This is driving a transition from human-led incursions to fully automated, self-learning, and adaptive campaigns, with the time to compromise dropping from days to a mere handful of hours. This compression of the incident response window necessitates a radical rethink of how organisations structure their security operations centres and allocate their teams. In response, the concept of **digital immunity** is emerging as both a strategic imperative and a viable end state for companies seeking to survive in this new era of algorithmic warfare. Achieving such resilience requires moving away from a reactive posture towards a proactive ecosystem of autonomous agents capable of defending the perimeter without constant human oversight.

While AI has remained a significant topic of C-suite discussion for several years, the recent shift towards weaponised autonomy has fuelled a drive for **AI parity** between attackers and defenders. Malicious actors are no longer constrained by the physical limitations of human operators who must manually navigate network topologies and identify vulnerabilities. Instead, AI-driven agents are proliferating, deployed to conduct reconnaissance and exploit weaknesses at machine speed, potentially across thousands of nodes simultaneously. This development renders traditional time-based security metrics, such as mean time to detect or mean time to respond, largely obsolete. The defensive strategy must therefore pivot towards an architecture that prioritises the **autonomous neutralisation** of threats before they can propagate within an internal environment. Senior leaders must recognise that relying on human intervention to validate every alert is destined to fail in an environment increasingly defined by sub-second attack vectors.



The arrival of advanced frontier models, including **Claude Mythos, Claude Fable, and ChatGPT 5.5-Cyber**, has fundamentally redefined traditional vulnerability operations within the enterprise network. This manifest shift in risk emerged naturally as autonomous models began analysing source code to identify isolated software bugs and hidden executables. What has fundamentally changed with the latest frontier architectures is an autonomous agent's ability to extract multiple low-severity bugs that would traditionally remain invisible in a corporate backlog. Rather than leaving these minor discrepancies unexploited, bad actors can, with access to such models, systematically chain them together into a single, highly destructive enterprise exploit. Managing this specific exposure requires deploying sophisticated platforms that provide **AI-harness, security context, and agentic AI capabilities** to support rigorous and continuous red-team activities.

The move towards an agentic AI platform is not merely a technical upgrade but a fundamental shift in the relationship between human expertise and machine capability. Many organisations have spent the past decade accumulating a vast array of security tools that demand significant manual effort to manage and correlate. These legacy capabilities often create **analytical paralysis**, in which the volume of telemetry

data overwhelms a security team's capacity to identify genuine threats. PAC advocates integrating autonomous agentic solutions or services into the core of a security architecture to identify patterns that would be invisible to even the most experienced human analysts. This transition allows an enterprise to focus its most valuable human resources on high-level strategic planning rather than routine ticket management, with the objective of creating a defensive layer that is as agile and adaptive as the AI adversaries it is designed to encounter daily.

Success in this new type of cybersecurity operating environment depends on an organisation's ability to move beyond the initial exploration phase towards a fully composable **agentic AI cyber defence stack**. This requires significant investment in both the underlying data infrastructure and the governance models that oversee autonomous decision-making, as operational friction associated with legacy solutions and services can be a major barrier to adopting advanced AI technologies. Leaders must balance the need for rapid technical innovation with organisational realities and the inherent resistance to change within large security departments. The path to achieving total digital resilience is fraught with technical and cultural challenges that require sustained commitment from the highest levels of the corporate hierarchy.

PAC recommends the following key takeaways:



Redefine response metrics: Security leaders must move beyond traditional time-based metrics to reflect the instantaneous nature of autonomous, algorithmic attacks.



Implement swarm architectures: IT and security teams should begin designing collaborative agent networks to share threat intelligence across disparate functional silos in real time.



Evaluate adversarial parity: A formal gap analysis is required to assess how existing defensive controls compare with the latest generation of weaponised autonomous tools.



Strategic resource allocation: Investment strategies should prioritise developing internal agentic capabilities over maintaining traditional manual monitoring contracts.



Establish agent ethics: Governance bodies must define the acceptable boundaries for autonomous decision-making to ensure automated remediation actions remain within corporate risk appetites.

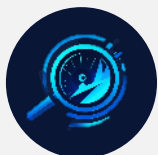
The Adversarial Landscape Shaped by Weaponised Agents and Grey State Threats

Modern adversaries are already deploying, and increasingly deploying, **weaponised agents** that utilise deep learning (DL) to bypass conventional security controls with alarming frequency. Such adversarial tools can generate highly targeted phishing campaigns and custom malware variants designed to evade detection by existing endpoint protection capabilities. Unlike traditional malware, DL AI agents adapt their behaviour to the specific defensive measures they encounter within a target network. This dynamic, self-modifying behaviour significantly reduces the effectiveness of signature-based detection on its own, because rapidly changing variants can evade static indicators and require continual updates. The scale of this problem is further exacerbated by the increasing availability of **adversarial AI capabilities packaged as SaaS products for attackers**, allowing even low-level actors to launch sophisticated campaigns.

A particularly sophisticated development in adversarial tactics is the emergence of **grey state intelligence** within malicious software payloads. These AI-driven attacks can be designed to detect when they are being executed in a simulated or sandbox environment used by security researchers. If the malware detects that it is in a test box, it will remain entirely dormant and perform no malicious actions to avoid triggering any alarms. Once the binary or code is deployed in a live production network, it activates its malicious functionality and begins navigating the environment with high precision. This level of environmental awareness demonstrates the rapidly evolving sophistication of modern **DL-based AI autonomous malware** and the limitations of current isolation techniques. Understandably, PAC considers that organisations must develop more advanced methods to identify these dormant threats before they cause widespread disruption to critical business operations.

While high-profile threats like deepfakes often dominate public conversation, the most significant risks to an enterprise are often subtle and technically focused. Issues such as non-human identity theft and the manipulation of automated supply chain processes pose a much greater threat to the long-term stability of large organisations. Attackers increasingly target the **autonomous interconnections** between enterprise applications to move laterally through networks undetected. These campaigns typically leverage compromised credentials from automated service accounts, a trend that is already increasing as agentic AI adoption grows. These accounts often lack the same level of oversight as human user accounts. Senior security leaders must ensure that their defensive strategy addresses these enterprise-specific risks rather than focusing exclusively on the most visible or publicised types of AI threats.

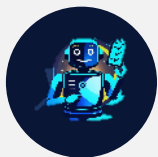
The reality of modern cyber warfare suggests that the distinction between state-sponsored actors and commercial criminal groups is increasingly blurred. This convergence of capabilities means that even mid-sized companies can now be targeted by attacks, once the exclusive domain of national intelligence agencies. The deployment of **Ransomware-as-a-Service (Raas)** platforms by dark web communities is evolving to incorporate autonomous negotiation and exfiltration agents, significantly lowering the barrier to entry for complex cybercrime. This is forcing enterprises to operate in a permanent state of conflict, where the only constant is the relentless pursuit of vulnerabilities by agent-led adversaries. Developing a robust understanding of these **adversarial agentic AI** techniques is essential for any organisation seeking to build a future-proof security posture. To address this, PAC recommends the following key takeaways:



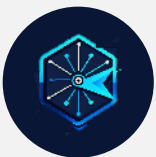
Expand sandbox capabilities: Security teams must develop more sophisticated detection environments that effectively simulate production traffic to trigger dormant malicious agents.



Monitor adversarial trends: Security teams need to actively track the evolution of deep-learning-based automated attack platforms to anticipate the next generation of weaponised AI.



Audit service accounts: Organisations should conduct regular, comprehensive reviews of all non-human identities to ensure they are not exploited as entry points for autonomous lateral movement.



Integrate behavioural analytics: Defensive strategies should shift away from static signatures and towards behavioural models that can detect the subtle patterns of autonomous adversarial reconnaissance.



Secure supply chains: Procurement and IT leaders should establish rigorous security standards for all automated third-party integrations to prevent exploitation of trusted interconnections.

The Paradigm Shift Redefining Augmentation Through Agentic Swarms

The primary obstacle to effective cyber defence in most large organisations remains the persistent reliance on **fragmented functional silos** that, at best, limit a unified response to threats and, at worst, prevent one. All too often, individual security tools operate as isolated islands of data, with their own telemetry formats and limited ability to share context with other solutions or services. This fragmentation is a major vulnerability because sophisticated attacks almost always involve a sequence of actions spanning multiple domains and control points. When capabilities for identity management or data protection cannot communicate with each other, they cannot identify the broader context of a multi-stage intrusion. The move towards a **collaborative agentic cybersecurity swarm** aims to address this fundamental weakness by creating a connected ecosystem in which agents share intelligence in real time.

In a collaborative swarm model, each agent can offer a distinct perspective on the network's state while coordinating its actions with other specialised agents. For example, an identity agent could alert a data protection agent upon detecting an anomalous login attempt from a high-privilege service account. The data protection agent could then proactively increase encryption levels or further restrict access to sensitive databases until the situation is fully investigated. This level of **inter-agent communication** enables a defensive system to simulate and counteract complex threat scenarios that would overwhelm a disconnected set of tools. The goal is to achieve a state of **swarm intelligence** in which the collective effectiveness of the agents is significantly greater than the sum of their individual parts.

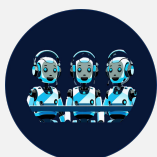
The role of a cybersecurity professional must also evolve to meet the demands of this new collaborative

paradigm, shifting from a manual operator to an **orchestrator of autonomous swarms**. Human cybersecurity analysts should no longer be expected to perform basic triage or initial investigations of low-level security alerts. Instead, they should focus on defining high-level strategic goals and the ethical guardrails that govern the behaviour of **autonomous agents**. This transition enables a security team to operate at a scale that would be physically impossible through manual effort alone. It also reduces the risk of human error and burnout, both of which are significant contributors to security failures in high-pressure environments. However, the orchestrator role requires a new set of skills that combine deep technical knowledge of AI with a strategic understanding of **organisational risk management**.

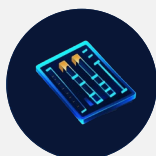
While the promise of collaborative security agent swarms is significant, the technical challenges of integrating disparate systems into a unified framework should not be underestimated. Legacy infrastructure often lacks the APIs or data standards needed to support the high-velocity communication required by **autonomous, agentic ecosystems**. There is also a significant risk of introducing new points of failure if the communication layer between agents becomes a target for adversarial manipulation. Ensuring the integrity and security of the **inter-agent protocol**, and adopting a **machine-zero-trust model**, must be top priorities for any organisation building a swarm-based defensive model. These complexities suggest that the transition will be incremental, requiring careful planning and a phased approach to implementation. For this, PAC identifies the following key success factors for scaling collaborative agentic AI cybersecurity capabilities across an organisation's operational footprint:



Unified Telemetry Fabric: Enterprise-scale deployment requires a standardised data abstraction layer that enables diverse security tools to inject normalised contextual information into the swarm in real time.



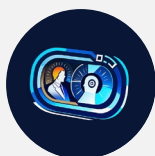
Inter-Agent Protocol Standardisation: Success depends on the engineering team adopting universal open communication standards that enable multi-vendor agents to safely coordinate operations across complex multi-cloud infrastructure.



Dynamic Guardrail Frameworks: Building this infrastructure across a large organisation requires central policy engines that dynamically translate corporate compliance requirements into local runtime constraints for autonomous agents.



Agent Skill Libraries: Enterprises need skill libraries built for Cyber Operations that enable the development of reusable skill sets. Such libraries provide workflows, enterprise context, and best practices that turn a general-purpose agent into a specialist cyber defender.



Machine Zero Trust Verification: Achieving enterprise-level security requires continuous verification of inter-agent communication channels to prevent malicious injection or manipulation of the swarm.

The Three Pillars of Agentic Defence to Enable Observability, Protection, and Trust

PAC considers that achieving a high degree of defensive effectiveness in an agentic AI operating model depends entirely on the quality and depth of **contextual observability** available to the solution or service. Autonomous agents cannot make accurate decisions in a vacuum and require access to a wide range of data sources to understand an enterprise's normal operating state. This includes not only real-time telemetry from endpoints and networks but also historical data that provides a baseline for long-term behavioural analysis. Many organisations possess **decades of historical data** from their security operations centres that could inform the decision-making of modern deep-learning (DL) AI agents. Leveraging this institutional memory, if architected correctly, provides a learning model that allows agents to understand an organisation's specific evolution and the unique risks inherent to its industry.

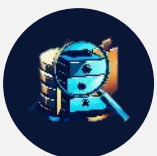
The **first pillar** of this structural framework is to create a unified security context data layer that can ingest and normalise data from across the enterprise. Without this AI harness, AI agents will struggle to correlate events across network segments or across multiple cloud environments. Observability must also extend to agents' internal states to ensure their actions are transparent and justifiable to human overseers. This requirement for agentic transparency is essential to building the trust needed to enable autonomous solutions and services to take high-stakes actions without manual approval for every decision.

The **second pillar** of the framework focuses on developing automation through reusable cyber-skilling libraries that can scale across the entire security operations lifecycle. Traditional automation has often relied on rigid playbooks that cannot adapt to the subtle variations that occur during real-world attacks. Agentic

automation uses deep learning techniques to identify the best security course of action based on the current context and the historical outcomes of similar events. This enables a much more flexible and effective response to **dynamic threat environments**, where the rules of engagement are constantly changing. However, the nascent state of agentic AI adoption carries the risk that automated errors could cause unintended business disruption, which remains a significant concern for many business, IT, and cybersecurity leaders. Implementing an autonomy framework in which agents gain greater authority as they demonstrate consistent accuracy can help mitigate these risks over time for critical workflows such as cybersecurity.

Establishing **digital trust and content engineering** is the **third and final pillar**, ensuring cybersecurity agents remain aligned with a business's strategic objectives and ethical standards. This involves creating sophisticated guardrails and playbooks tailored to an enterprise's specific needs. For instance, a cybersecurity AI agent operating in a highly regulated financial services environment will require a different set of **compliance guardrails** from one working in a retail setting. Content engineering enables the precise definition of what constitutes an acceptable action and ensures that agents always operate within the appropriate industry and cybersecurity bounds of corporate policy. This structured approach to trust is the optimal way to ensure that the deployment of **fully autonomous** cybersecurity agents remains a net benefit to the organisation rather than a source of new, unmanaged risk.

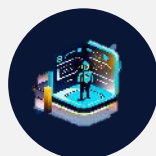
PAC advocates that the following key takeaways are core to establishing the three pillars:



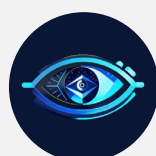
Centralise historical data: Data engineering teams should consolidate all relevant historical security telemetry into a single repository to support agent training and behavioural analysis.



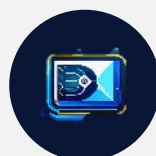
Deploy graduated autonomy: Security architects should implement a phased approach to autonomous action, granting agents increasing levels of authority as their performance proves.



Design contextual guardrails: Compliance and legal teams must collaborate with cybersecurity experts to define the specific operational boundaries for autonomous defensive agents.



Enhance agent transparency: Developers should prioritise creating detailed audit logs for every autonomous action to facilitate post-incident review and regulatory compliance.



Invest in content engineering: Organisations must allocate resources to develop enterprise-specific playbooks that guide decision-making for a swarm of cybersecurity AI agents to operate consistently.

Exploring Cyber Next Powered by Topaz Fabric – an Agentic AI First Approach

PAC considers that engaging a company like Infosys can significantly accelerate an enterprise's adoption of agentic AI cybersecurity capabilities at scale. Large enterprises face a combination of technical, skills, organisational, and governance challenges when introducing autonomous security agents, particularly when those agents must operate across complex environments and integrate with existing security operations. **Infosys brings established platforms, mature delivery models, and experience** in aligning security initiatives with broader enterprise AI programmes, helping organisations avoid the all-too-common challenge of fragmented implementations. As Infosys has built and continues to evolve agentic capabilities within its established security (Infosys Cyber Next) and AI (Infosys Topaz) Fabric, enterprises can leverage these to implement coordinated agent behaviour, shared threat context, auditable actions, and consistent governance from the outset. This approach reduces the risks associated with developing bespoke agentic stacks and supports faster progression towards large-scale production deployments. This can ensure that security automation evolves in step with responsible AI standards, regulatory expectations, and long-term operational sustainability.

The deployment of the three-pillar strategy within this specialised architectural framework provides organisations with a comprehensive roadmap for navigating architectural complexity. By establishing a unified security context data layer as the primary foundation, the platform effectively eliminates the persistent challenge of fragmented data silos across hybrid cloud environments. This profound level of observability enables autonomous solutions and services to correlate disparate alerts at machine speed while simultaneously providing the transparency required to build long-term institutional trust. Furthermore, the transition to dynamic agentic AI automation enables defensive protocols to adapt fluidly to rapidly changing threat landscapes without relying on rigid manual playbooks. Ultimately, the integration of robust digital trust and rigorous content engineering ensures that all autonomous actions remain strictly aligned with the enterprise's specific corporate policies and compliance mandates.

Strategic implementation often hinges on leveraging existing platforms, and this is where **Infosys' Cyber Next** can provide a consistent foundation for deploying and managing autonomous security agents. By integrating agentic capabilities directly into an existing security operating fabric, organisations can achieve a

much higher level of coordination and efficiency than with a piecemeal approach. Such an integrated model enables seamless sharing of context across agents and ensures that all autonomous actions are recorded and auditable. As AI agent adoption is likely to increase exponentially over the coming years, as the technology matures, reaching a benchmark of **100+ agents** is a significant milestone that demonstrates an enterprise's ability to operate at a truly autonomous scale. Using a proven platform, such as Infosys Cyber Next, also reduces the technical risks associated with building an entirely new agentic stack from scratch in an environment where speed is essential.

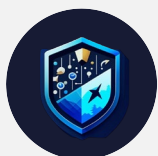
Additionally, alignment with a broader corporate AI initiative is essential to ensure that security goals are not sidelined by competing technical priorities within an organisation. For example, using a consistent framework such as the **Infosys Topaz fabric** ensures that security agents are built to the same standards and governed by the same models as other enterprise AI applications. This consistency can reduce the overhead of managing multiple AI technologies and enable an organisation to leverage its existing **responsible AI (RAI)** investments. It can also ensure that security agents remain compliant with an enterprise's broader ethical and regulatory obligations across all functional domains. PAC advocates that senior leadership ensure their security department is a key participant in developing an overall **corporate AI strategy** to prevent new security blind spots.

One of the most significant challenges in operationalising agentic AI is the requirement for **continuous model refinement** to stay ahead of evolving adversarial tactics. Agents that are not regularly updated with the latest threat intelligence will quickly lose effectiveness as attackers develop new ways to bypass their defences. This requires a dedicated team of data scientists and security researchers to monitor agent performance and identify areas for improvement. The use of **automated retraining loops**, combining Infosys Topaz and Infosys Cyber Next, can streamline this process by enabling agents to learn from their successes and failures in real time. The final consideration for any organisation implementing an agentic cybersecurity strategy is maintaining a strong focus on the **responsible and ethical deployment of AI**. As agents gain greater autonomy over critical business functions, the potential for unintended bias or incorrect decision-making becomes a significant security and business risk. Establishing clear lines of accountability and implementing **transparent reporting mechanisms** is

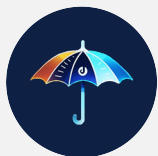
essential for maintaining the trust of both internal stakeholders and external regulators. This ethical foundation is not just a compliance requirement but a fundamental component of the overall security of an agentic stack.

In conclusion, PAC considers Infosys to have platforms, in the form of Infosys Cyber Next powered by Infosys Topaz, which provide a foundation needed to effectively deploy cybersecurity-focused AI agents to successfully navigate the next phase of autonomous cyber warfare.

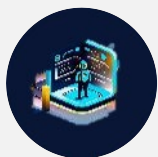
However, enterprise leaders must shift from isolated technological experimentation to a fully integrated and structured operational roadmap to:



Optimise Enterprise Risk Posture by leveraging a unified security context and deep system observability to eliminate fragmented silos and establish a clear baseline of operational risks.



Deploy Comprehensive Protection Protocols by rapidly scaling operational capabilities to provide means of maintaining hundreds of AI agents as an organisation scales usage, across all critical production environments to counter autonomous threats.



Establish Automated Prevention Guardrails by integrating reusable assets and robust engineering guardrails for AI, preventing adversarial exploits from disrupting core workflows.



Implement Continuous Retraining Loops within engineering departments to ensure automated feedback systems continuously update models with live threat intelligence feeds, maintaining defensive efficacy.



Address the Internal Skills Gap by aggressively reskilling and training current cyber defenders in advanced agentic threats and automated defensive capabilities to complement autonomous infrastructure.

About Infosys



Infosys is a global leader in next-generation digital services and consulting. We enable clients in more than 50 countries to navigate their digital transformation. With over four decades of experience in managing the systems and workings of global enterprises, we expertly steer our clients through their digital journey. We do it by enabling the enterprise with an AI-powered core that helps prioritize the execution of change. We also empower the business with agile digital at scale to deliver unprecedented levels of performance and customer delight. Our always-on learning agenda drives their continuous improvement through building and transferring digital skills, expertise, and ideas from our innovation ecosystem.

Visit www.infosys.com to see how Infosys (NSE, BSE, NYSE: INFY) can help your enterprise navigate your next.

About PAC



We are a content-based company with a consulting DNA. PAC is the leading European consulting and analyst firm supporting software & IT service vendors worldwide. Since 1976, we have helped our clients to understand market dynamics, grow their revenue and raise their profile. Our unrivalled understanding of European markets, and deep research coverage help key market players to define their strategy, optimize their go-to-market and increase market share. PAC is an analyst-led consultancy with a team of over 100 experts across Europe. We provide market research and analysis on more than 30 countries worldwide, delivered through our portfolio pillars, Guidance, Insights, and Visibility, and our renowned SITSI® research platform. More on www.pacanalyst.com.

Contact:

Spencer Izard
Research Director
s.izard@pacanalyst.com
+44 (0)7398 433071

PAC UK
2 Minton Place,
Victoria Road,
Bicester, Oxfordshire,
OX26 6QB, United Kingdom

www.pacanalyst.com
www.sitsi.com