CMORG
CROSS MARKET OPERATIONAL RESILIENCE GROUP

FS-ISAC

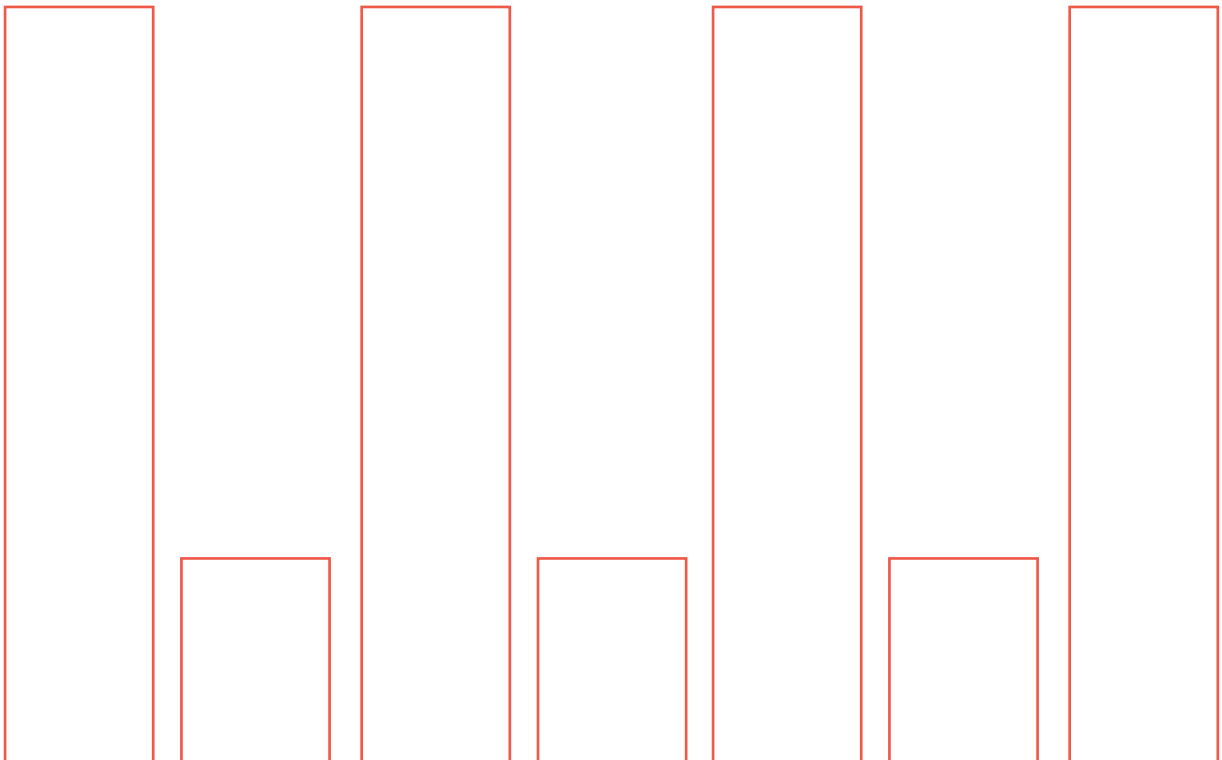
FSCCC
Financial Sector Cyber Collaboration Centre

CDA
CYBER DEFENCE ALLIANCE

FIRM GUIDANCE FOR FRONTIER AI

AI TASKFORCE

VERSION 1.0 | JUNE 2026| TLP CLEAR



CONTENTS

CMORG-endorsed capabilities (including good practice guidance, response frameworks and contingency tools) have been developed collectively by industry to support the operational resilience of the UK financial sector. The financial authorities support the development of these capabilities and collective efforts to improve sector resilience. However, their use is voluntary, and they do not constitute regulatory rules or supervisory expectations; as such, they may not necessarily represent formal endorsement by the authorities.

1 EXECUTIVE SUMMARY

Frontier AI models, defined by the National Cyber Security Centre (NCSC) as the most advanced artificial intelligence systems in development. These systems, which sit at the frontier of what technology can currently do, can perform complex tasks which can generate benefits and risks to cyber security. These models, such as Anthropic's Claude Mythos and other recent large-scale AI systems, are compressing the time between vulnerability discovery and exploitation, enabling attackers to operate at greater speed, scale and sophistication than previously observed. While these technologies will also deliver long-term defensive benefits, the immediate challenge for financial institutions is clear: they must adapt now to maintain resilience in an environment where the pace, volume and accessibility of cyber capabilities are increasing faster than existing defensive models were designed to accommodate.

This guidance, developed by the Cross Market Operational Resilience Group (CMORG), consolidates a broad and growing body of industry and public authority thinking into a single reference point, enabling financial institutions of different sizes and levels of maturity to draw on leading practices**. It reflects a strong and emerging consensus on the core capabilities required to maintain resilience, setting out a structured approach across governance, protection, response, automation and collective resilience, which must operate as a coherent system rather than in isolation.**

In the near term, financial institutions should expect a material increase in both the scale and speed of vulnerability discovery and exploitation. Remediation timelines will need to compress from weeks to days, and in some cases hours, requiring organisations to operate with much greater urgency, coordination and discipline. **This will require a shift in risk calculus, with financial institutions in some instances needing to consider a stronger emphasis on rapid patch deployment balanced against other potential impacts, such as service availability.** This reflects a necessary evolution in responsible decision-making, as firms work alongside regulators and policymakers to adapt to changing threat conditions.

However, patching alone will not be sufficient. The challenge presented by frontier AI extends beyond vulnerability discovery to how cyber risk is managed across the organisation as a whole. **Effective response will require coordinated action across governance and leadership, operating models, technology architecture, detection and response capabilities, and the management of supply chain and ecosystem risk.** Financial institutions must be able to operate at increased speed, contain compromise effectively, and continuously adapt to evolving threat conditions, recognising that the primary challenge is no longer understanding what good looks like, but executing it consistently at pace and at scale.

Many of the controls and capabilities set out in this document represent established best practice. What has changed is not the nature of these controls, but the speed, scale and intensity with which they must now be implemented. **As frontier AI continues to evolve, specific techniques and recommendations will also develop; however, the need for organisations to act quickly, manage risk dynamically, and operate with greater technical and organisational agility will remain constant.**

Given the pace of development in frontier AI, this guidance will necessarily evolve over time. However, it is intended to provide a practical and actionable baseline for firms to assess their current capabilities and accelerate their response. Financial institutions should work from the basis that the window to strengthen resilience ahead of further acceleration in threat capability is limited and time-sensitive, and act with urgency to ensure readiness for these changes in the landscape.

2 TAKE CONTROL

2.1 GOVERNANCE AND LEADERSHIP

The speed and scale of frontier AI-driven cyber threats require urgent action across firms and their wider ecosystem. Strong leadership, with clear executive-level ownership and oversight, is critical to an effective response. This includes close interaction with regulators, who play a central role in maintaining sector resilience. Measures taken to mitigate this evolving threat will inevitably introduce disruption; governance, risk management, and compliance processes must therefore be capable of operating effectively at increased speed.

Leadership must also ensure that risk positions and threat landscape assessments are updated to reflect AI-driven attack dynamics, while supporting the responsible and risk-assessed use of AI-enabled defence. Governance approaches should be evidence-led, drawing on cyber threat intelligence and peer exchange to mitigate the risks of both over- and underreaction. Boards and senior committees should assess whether their financial institution is equipped to respond at the required pace and ensure they develop sufficient understanding of associated cyber security risks.

Key Takeaways – Firms Should:

- Establish clear executive ownership, with aligned leadership expectations for frontier AI cyber preparedness across all relevant functions.
- Ensure governance and oversight are evidence-led, embedding AI-linked security risks in the risk register with a clear distinction between observed threats and more speculative scenarios.
- Update risk appetite and governance frameworks to reflect AI-driven disruption, including explicit support for rapid remediation decisions and associated trade-offs between resilience and service availability.
- Strengthen critical risk metrics, prioritising remediation speed, key exposure indicators, and improved measurement of risk reduction over time.
- Strengthen leadership capability through targeted executive education, ensuring accountability for cyber resilience is clearly embedded in performance objectives.

And Ask Themselves:

- Do our risk and threat models adequately reflect AI-driven attack dynamics, and does our organisational culture support the required urgency, speed and discipline?
- Are we deploying AI-assisted defence in a controlled, disciplined and risk-assessed manner?
- Is our approach evidence-led and proportionate, aligned to cyber threat intelligence and peer information exchange?
- Are we effectively governing AI-driven cyber risk, ensuring clear executive ownership, aligning decision-making and escalation processes to operate at speed, and maintaining effective engagement with regulators and third parties during periods of rapidly evolving threat?

2.2 OPERATING MODEL SHIFT

Speed, scale and volume are central to an operating model capable of countering frontier AI-enabled attacks. Cyber operations must function at machine speed across detection and initial response, rapidly identifying and containing threats, while prevention and recovery processes are accelerated and tightly coordinated to minimise impact. In the near term, vulnerability management will be critical, reducing the time from discovery to remediation from weeks to days, while operating at significantly greater scale.

Delivering this requires a broader operating model shift, with some existing control environments and delivery models requiring fundamental redesign to operate at the necessary speed and scale. This includes embedding DevSecOps and shift-left security practices (embedding testing and mitigation

into the earliest phases of software development), ensuring that security is designed into software and vulnerabilities are identified and remediated during development rather than retrofitted later. Firms should also revisit change management and governance frameworks, linking urgent remediation and large-scale patching more closely to incident response, recognising that rapid change can introduce operational disruption requiring clear escalation, coordination, and oversight. In parallel, assurance should evolve from static assessments to more continuous evaluation, supported by code and component inspection, to maintain real-time visibility of assets, dependencies, and exposures.

Key Takeaways – Firms Should:

- Operate cyber and technology functions at increased speed and scale, reducing timelines across prevention, detection, response, and recovery to match AI-driven attack dynamics.
- Embed DevSecOps and shift-left practices, ensuring vulnerabilities are identified and remediated during development and that security assurance is integrated into standard delivery processes.
- Align vulnerability remediation, incident response, and change management, enabling rapid and coordinated action during periods of elevated threat, supported by clear escalation and decision pathways.
- Evolve assurance from periodic assessments to continuous evaluation, supported by integrated data across development, infrastructure and operations, enabling faster and more confident decisions under pressure.

And Ask Themselves:

- Is our operating model designed to function at machine speed, with clear decision pathways that enable rapid action while maintaining appropriate oversight and accountability?
- Are we embedding security early enough in the technology lifecycle to prevent weaknesses from reaching production at scale?
- Are change, remediation, and response processes sufficiently aligned to support high-tempo patching and containment activity delivered safely and consistently?
- Is assurance continuous and sufficiently evidence-led to support rapid decision-making under persistent, high-velocity attack?

3 PROTECT YOUR ORGANISATION

3.1 ATTACK SURFACE REDUCTION

Financial institutions should promptly, carefully, and continuously manage their attack surface by eliminating or isolating unsupported technologies, unnecessary exposure, weak configurations, stale identities and excessive privilege. This must be supported by continuous exposure validation across assets, configurations, and identities, with clear ownership and accountability for remediation. External exposure should be independently validated and linked to accountable owners, service criticality and supplier dependencies, recognising attack surface reduction as a first-order control.

This approach must extend to AI systems, which form part of the attack surface where they access data, call tools, execute code, or operate through identities and APIs. These should be governed as privileged applications, with tightly controlled access, clearly defined purpose, robust logging, human oversight for high-impact actions, and effective containment mechanisms, particularly where they interface with critical environments, data or business services.

Key Takeaways – Firms Should:

- Maintain comprehensive, continuously updated visibility of internet-facing assets, identities, AI systems, third-party integrations and business service dependencies, with clear ownership and accountability.

- Aggressively reduce exposure by eliminating unnecessary services, restricting privilege, and remediating misconfigurations, stale identities and over-permissioned access paths.
- Govern AI on the basis that it operates with privileged access: tightly control high-risk capabilities (including tools, code execution and APIs), enforce scoped permissions, robust logging, human oversight, and effective kill-switch mechanisms.
- Independently validate and prioritise exposure risk by continuously scanning the external attack surface, linking exposure to critical services and suppliers, and applying urgent remediation to high-impact assets.

And Ask Themselves:

- What is externally visible and exploitable today, and which of these assets support important business services?
- Where are our highest-risk exposures, including unsupported, misconfigured, vulnerable, or unowned assets, and what would be the impact of exploitation?
- Which suppliers, APIs, Software as a Service (SaaS) platforms and AI systems introduce privileged access paths into our environment?
- How quickly can we remove, isolate or mitigate exposures, particularly those affecting sensitive data, critical services or high-impact business workflows?

3.2 ARCHITECTURE AND RESILIENCE

Firms should ensure systems are architected and tested to enable containment, recovery and sustained operation under persistent, high-velocity attack. Environments must be designed on an assumption of breach, limiting propagation and maintaining critical services even where controls fail. Strong containment, real-time detection and response, and the rapid and systematic reduction of legacy or unsupported technology are critical pillars.

Rather than duplicating controls described elsewhere, architecture should be treated as the mechanism through which governance (see Governance and Leadership), high-tempo cyber operations (see Operating Model Shift), and systematic exposure reduction (see Attack Surface Reduction) are integrated and operationalised. Architectures must continuously evolve to reflect accelerating threat conditions, increasing system interdependencies, and the need to operate at speed under AI-driven attack dynamics.

Key Takeaways – Firms Should:

- Design for containment and least privilege, applying Zero Trust principles, segmentation, and strong identity and access controls to restrict access paths and limit lateral movement.
- Strengthen real-time detection and response through continuous monitoring, centralised logging and advanced detection capabilities to rapidly identify and contain threats.
- Operate at high tempo with resilience by regularly stress-testing incident response, including concurrent exposures and AI-enabled attack scenarios, supported by clearly defined emergency pathways.
- Continuously evolve architecture and reduce fragility by iterating controls in line with accelerating threat conditions and eliminating or isolating legacy and unsupported technologies.

And Ask Themselves:

- Have we engineered our architecture to deliver containment and resilience, applying Zero Trust, segmentation, isolation and least privilege, alongside controlled failure modes, to limit lateral movement and sustain critical services?

- Can we operate effectively at machine speed, including the ability to rapidly detect, prioritise and remediate vulnerabilities, and to manage multiple critical exposures simultaneously through predefined emergency pathways?
- Do we have embedded, real-time detection and response capabilities at scale, supported by integrated telemetry, centralised visibility and secure automation, enabling containment at the required pace within clear governance guardrails?
- Are we continuously modernising and evolving our architecture, reducing legacy complexity and technical debt, and adapting designs in response to threat intelligence, incidents and regulatory expectations to keep pace with AI-driven threats?

4 PREPARE TO RESPOND AT PACE

4.1 DETECTION AND RESPONSE

Financial institutions should operate on the basis that threat actors will gain access to their networks. Intelligence-led-defence-in-depth is therefore critical to enabling rapid detection and limiting threat actor activity. Cyber threat intelligence should be translated into practical action, informing what to detect, what telemetry is required, how to respond, and how to validate that controls are effective. AI can support analysts in increasing speed and scale; however, high-impact decisions, including incident containment and operational changes, should remain controlled, explainable, and auditable.

Key Takeaways – Firms Should:

- Ensure core monitoring coverage across identity systems, cloud services, endpoints, email, web activity, DNS, proxies, data access, developer platforms, and payment and fraud processes.
- Use MITRE ATT&CK as a common behavioural framework to describe adversary activity and support detection engineering and testing, while recognising that local telemetry and context remain essential.
- Ensure front-line analysts are appropriately trained and regularly refreshed, so they can accurately assess the significance and severity of detections.
- Use AI to support analysts, summarising evidence, correlating activity, and generating investigation hypotheses, with appropriate access controls and human oversight.
- Explore ways to automate initial incident response and containment, using pre-defined and tested playbooks to isolate assets, revoke access, and execute early defensive actions at speed, reducing the risk of rapid attacker propagation.

And Ask Themselves:

- Do we have the right monitoring and telemetry to detect threats quickly across our most important systems and services?
- Are incident response decisions controlled, explainable, and auditable, particularly where AI is involved?
- Have we tested our response plans under realistic stress conditions, including scenarios involving supplier disruption and reduced operational capacity?

4.2 VULNERABILITY MANAGEMENT TRANSFORMATION

Vulnerability management should be continuous, intelligence-led, and focused on real-world risk. Firms should prioritise vulnerabilities that are actively exploited, externally exposed, affect key suppliers, or have clear business impact. The focus should move beyond severity scores towards what is exposed, exploitable, and most consequential. Accelerating remediation will increase the risk of operational disruption, both intentional and unintentional, but is necessary to reduce overall exposure and prevent greater harm.

Key Takeaways – Firms Should:

- Set risk-based remediation targets, measured in hours or days, for vulnerabilities that are actively exploited, internet-facing, or linked to important business services, while ensuring sufficient capacity to manage sustained surges in remediation activity.
- Connect vulnerability data to asset inventories, accountable owners, configuration records, cloud inventories, software component lists, supplier products, deployment plans, and change windows.
- Agree emergency remediation pathways in advance with technology, change, resilience, legal, communications, and business-service owners.
- Apply compensating controls where immediate patching is not possible, including isolation, service disablement, configuration changes, virtual patching, enhanced monitoring, and access restriction.
- Report exposure and remediation outcomes, including how long the financial institution remained exposed, what residual risk was accepted, and whether remediation was fully validated, rather than relying solely on counts of overdue vulnerabilities.
- Maintain visibility and control over third party and open-source dependencies, including up to date inventories of libraries and components, continuous monitoring of vulnerability disclosures, and the ability to rapidly remediate or revert to secure versions while preserving system integrity.

And Ask Themselves:

- Are vulnerabilities prioritised based on exploitability, exposure, and business impact, rather than inherent severity alone?
- Can we demonstrate how long systems were exposed and whether remediation was fully validated?

4.3 RESPONDING AT PACE

Automation and AI should be used to reduce organisational latency across security, engineering, operations, and recovery functions. Well-designed automation improves speed, consistency, and evidence capture, while well-governed use of AI can increase coverage and help defenders keep pace with faster-moving threats. As automation and AI operate with increasing privilege, business impact, or irreversibility, firms should apply correspondingly stronger controls, ensuring accountability remains clear and high-impact decisions are appropriately governed.

Key Takeaways – Firms Should:

- Use automation and AI to reduce organisational latency across security, engineering, operations, and recovery functions, improving speed, consistency, and evidence capture.
- Automate routine tasks such as alert and context enrichment, event correlation, owner identification, ticket routing, status reporting, and evidence capture to increase operational throughput at scale.
- Use automation and AI to prioritise work based on exploitability, exposure, asset criticality, dependency context, compensating controls, and business impact.
- Design automation and AI agents to fail safely, applying least privilege, scoped tool access, comprehensive logging, rate limits, approval steps, rollback options, and kill-switch mechanisms.
- Monitor automation and AI agents for unsafe or unexpected behaviour, including manipulated inputs, unusual tool use, data leakage, or unreliable outputs, and measure effectiveness through outcomes such as reduced exposure time, improved validation, and stronger operational resilience.

And Ask Themselves:

- Where can automation or AI meaningfully improve the speed or quality of our defensive activities today?
- Are we applying AI early enough in the lifecycle to identify and remediate weaknesses before deployment?
- Are AI-assisted outputs tested and validated to the same standard as human-generated work?
- Is clear human accountability maintained for high-impact decisions affecting customers, services, or privileged access?
- Are automation and AI agents actively monitored as potential malicious insiders for unsafe behaviour, unreliable outputs, or data leakage?

5 WORK COLLECTIVELY

5.1 SUPPLY CHAIN AND ECOSYSTEM RISK

AI-enabled vulnerability discovery will further increase the importance of understanding risk across suppliers and the broader ecosystem. Suppliers, software dependencies, shared infrastructure, cloud services, open-source components, and AI providers should be treated as part of the sector's overall attack surface. Firms remain accountable for the resilience impact of third party failures and should be prepared to coordinate urgent remediation across organisational boundaries.

Key Takeaways – Firms Should:

- Maintain visibility of critical suppliers and dependencies, including software components, cloud services, open-source usage, AI services, and software bills of materials where available.
- Strengthen contractual and operational expectations, requiring risk-based remediation, secure software updates, timely incident notification, and evidence that controls are effective.
- Engage early with vendors and sector partners, confirming impact, obtaining fixes, and prioritising externally exposed or business-critical components.
- Plan for large-scale, coordinated patching and contingency activity across interconnected systems, suppliers, and dependencies.
- Participate in trusted sector collaboration and intelligence sharing, strengthening collective defence against ecosystem-wide threats.

And Ask Themselves:

- Do we have sufficient visibility over critical suppliers, software components, cloud services, and AI dependencies?
- Can we identify which third-party dependencies pose the greatest risk to important business services?
- How quickly can we engage vendors and partners to confirm impact and obtain fixes when urgent vulnerabilities emerge?
- Is our cyber threat intelligence sufficiently timely, relevant and actionable to inform decisions on supplier risk and vulnerability prioritisation?
- Are we exposed to concentration risk or single points of failure across critical suppliers, services or technologies, and do we have viable alternatives or contingency plans?

6 APPENDIX – REFERENCE DOCUMENTS

1) **Australian Signals Directorate (ASD)** – Frontier models and their impact on cyber security (09 April 2026)
   View: Frontier models and their impact on cyber security [cyber.gov.au]
2) **Australian Signals Directorate (ASD) –** Frontier models and their impact on cyber security (30 April 2026; updated 8 May 2026)
   View: Frontier models and their impact on cyber security [cyber.gov.au]
3) **Bank of England, FCA and HM Treasury** - joint statement on Frontier AI models and cyber resilience (15 May 2026)
   View: BofE, FCA and HM Treasury - joint statement on Frontier AI models and cyber resilience
4) **CERT-EU** – AI and vulnerability discovery: Defenders should adapt now (21 April 2026)
   View: AI and vulnerability discovery – Defenders should adapt now [cert.europa.eu]
5) **Cloud Security Alliance (CSA) –** The AI vulnerability storm: Building a "Mythos-ready" security programme (April 2026)
   View: The AI Vulnerability Storm – Cloud Security Alliance
6) **Cyber Security Agency (CSA) Singapore** – Risks associated with frontier AI models (15 April 2026)
   View: Advisory on risks associated with frontier AI models [csa.gov.sg]
7) **FS-ISAC** – Preparing the enterprise for AI-enabled vulnerability discovery (20 April 2026)
   View: Preparing the enterprise for AI-enabled vulnerability discovery [fsisac.com]
8) **FS-ISAC** – AI-enabled vulnerability detection & remediation: third parties (20 April 2026)
   View: AI-enabled vulnerability detection & remediation – perspectives on third parties [fsisac.com]
9) **JPMorgan Chase** – Fortifying the enterprise: 10 actions to take now for AI-ready cyber resilience (17 April 2026)
   View: Fortifying the enterprise: 10 actions to take now for AI-ready cyber resilience
10) **National Cyber Security Centre (NCSC)** – Retaining defensive advantage against frontier AI (15 April 2026)
    View: Retaining defensive advantage in the age of frontier AI cyber capabilities [ncsc.gov.uk]
11) **National Cyber Security Centre (NCSC)** – Supporting AI adoption for UK cyber defence (23 April 2026)
    View: Supporting AI adoption for UK cyber defence [ncsc.gov.uk]
12) **National Cyber Security Centre (NCSC)** – Preparing for a 'vulnerability patch wave' (1 May 2026)
    View: Preparing for a vulnerability patch wave [ncsc.gov.uk]
13) **National Cyber Security Centre (NCSC)** – Impact of AI on cyber threat from now to 2027 (7 May 2025)
    View: Impact of AI on cyber threat from now to 2027