MARKET SURVEY

Frontier AI and the banks

The evidence on the threat, and what US banks, FS-ISAC and the US government are doing about it

Since April 2026, frontier models from Anthropic and OpenAI have been able to find vulnerabilities and write working exploits with little human help. Independent evaluations by the UK AI Security Institute and METR, research papers, the AI developers' own reports, FS-ISAC and the bank trade bodies agree on the threat and not on its ceiling: in tests, a model completed a weakly defended network intrusion in 3 of 10 attempts, and the best agents exploited 13 to 18% of nearly 900 known flaws with defences switched off; no test yet involves active defenders. Korea's autumn wave, a suspected and unconfirmed AI-assisted campaign against weakly protected peripheral systems, shows the threat does not need a restricted frontier model; no frontier-AI attack on a US bank has been publicly reported. In the United States, the largest banks are adding remediation capacity, JPMorganChase as a launch partner of Anthropic's Project Glasswing and four other large banks reportedly with model access. FS-ISAC and the trade bodies set a common playbook, with third-party providers as the weak joint. Washington has mostly relied on coordination: voluntary early access to frontier models, a government-led clearinghouse, supervisory questions and, since August, a planned programme for vetted companies to run offensive operations against foreign cybercrime groups under federal control, with no new rules for banks so far. The exception for the AI developers was a June export control that led Anthropic to disable two models for all customers. For a bank outside the US, the lessons are speed, supplier discipline and access that can be withdrawn. Gaps where no public material was found are listed in Appendix A.



DATE 27 October 2026

VERSION 1.3

Olivier Busolini · Global Head of Information Security · Dubai

© Olivier BUSOLINI - 2026 - No reuse without permission

Contents

Executive summary 3

Action guide 4

01 · What the expert evidence says about the threat 6

02 · What has changed in observed attacks 8

03 · How the largest US banks are responding 10

04 · FS-ISAC and the sector bodies 12

05 · What it means for a non-US bank 15

Appendix A · Method and sources 16

Appendix B · The US government response 17

Appendix C · Outside the US, and what comes next 19

Sources 20

© Olivier BUSOLINI - 2026 - No reuse without permission

Executive summary

Frontier AI has made finding vulnerabilities cheap; the banks that cope will be the ones that can fix them fast, and in the United States that has become a shared problem for banks, their suppliers and the government.

MY VIEW

In several of the most serious operations Anthropic reports, a human directed every step and the models supplied the labour; what has changed is the attacker's cost of that labour, and with it the defender's time. I read the US response as a test of execution. The controls CMORG calls established practice now have to run at a new tempo, and a remediation backlog priced on the old cost of attack needs repricing within the risk appetite the board has set, measured in days from fix to deployment on internet-facing systems.

1 **A capability threshold has been crossed.** Given initial access, Mythos Preview took UK AISI's weakly defended simulated network end to end in 3 of 10 attempts; in ExploitGym the best agents exploited 13 to 18% of 898 real flaws with defences off.

2 **The threat does not need a restricted model.** Korea's suspected, unconfirmed AI-assisted wave hit weakly protected peripheral systems; the regulator counted 16 firms by 14 October. No frontier-AI attack on a US bank is public.

3 **Time to exploit is shrinking; the ceiling is unknown.** ENISA sees the window from discovery to exploitation falling, potentially, to hours; no test yet involves active defenders.

4 **Set deadlines for fixing.** FS-ISAC asks for SLAs in hours or days; JPMorganChase's ten actions lead with currency, inventory and patch speed; CISA gives US agencies three days for the riskiest flaws.

5 **Clear known risk first.** FS-ISAC's first action: external systems first, no long-standing exceptions where a patch exists, backlogs treated as operational risk.

6 **Fund the capacity.** Bank of America added 1,000 contractors; JPMorganChase calls for automated testing and staged roll-out; Japan's FSA asks for prioritisation and more patching staff now.

7 **Hold suppliers to the pace.** FS-ISAC flags monthly or quarterly vendor patching; Fed staff research calls suppliers a hidden fault line; BPI wants providers critical to many banks supervised directly.

8 **Govern AI agents before scaling them.** Models never hold credentials and consequential actions need human approval; US supervisors are already asking how agents are shut down.

9 **Prepare to recover.** The FSB Chair says financial firms will need to prepare for simultaneous disruption across firms, and stresses restoring critical systems from "bare metal".

10 **Keep two ways into defensive models (analysis).** The June export control showed access can be withdrawn overnight; FS-ISAC asks for portable workflows across providers.

Sources for each point are given where it is developed in the body and appendices. The thesis and the conclusion of point 10 are the author's analysis.

© Olivier BUSOLINI - 2026 - No reuse without permission

Action guide

The survey's conclusions as three tools: a roadmap for the executive committee, the controls in order of priority, and the indicators for reporting and oversight.

TABLE 1 · ROADMAP FOR THE EXECUTIVE COMMITTEE

HORIZON	DECISION	WHAT IT COMMITS THE BANK TO
First 30 days	Close the peripheral gap	Inventory every externally reachable system and put strong authentication on all of them
First 30 days	Adopt remediation tolerances	Approve the indicators and thresholds of Table 3, with owners and monthly reporting to the board
Within 90 days	Hold suppliers to the pace	FS-ISAC's vendor questions for critical suppliers; contracted fix times and out-of-hours patching
Within 90 days	Fund remediation capacity	A prioritised backlog and temporary specialist support now; automated testing and deployment over time
Within 12 months	Govern AI agents before scaling them	Agent inventory; no credentials held by models; human approval for consequential actions; tested shutdown
Within 12 months	Prove recovery	Restore critical systems from bare metal in a full test; plan for simultaneous disruption across firms and shared suppliers
Within 12 months	Secure portable access to defensive models	At least two providers; workflows proven on both; an exit plan if one is withdrawn

Source: Table 7. The horizons are the author's analysis.

MY READ

I would start with the periphery and the tolerances in the first month. The inventory of what is reachable from the internet can begin inside existing budgets; putting strong authentication on all of it may not. The executive committee's decision on how fast critical findings must be fixed then sizes the funding for capacity. And I would not wait for complete data; a partial inventory acted on this quarter is worth more than a complete one next year.

© Olivier BUSOLINI - 2026 - No reuse without permission

TABLE 2 · CONTROLS IN ORDER OF PRIORITY

PRIORITY	FS-ISAC ACTION	FIRST CONTROL TO PUT IN PLACE
1 · Now	1 Remediate known risk	External systems first; end exceptions where a patch exists
1 · Now	2 Harden the perimeter	Wider WAF coverage; fewer admin and remote access points
1 · Now	3 Compress patch timelines	SLAs in hours or days; external exposure weighted beyond CVSS-only scoring
1 · Now	8 Control high-risk accounts	Inventory of service accounts, API keys and tokens; central revocation and rotation
2 · Next	4 Align accountability	Patch velocity and currency in system owners' objectives, reported to the board
2 · Next	5 Inventories and third parties	Real-time inventory, suppliers and AI providers included
2 · Next	6 Replace end-of-life technology	Version N or N-1 for internal and third-party software
2 · Next	9 Shift to exploit prevention	Segmentation; controls that block exploits in progress
3 · Then	7 Govern open source	Dependency inventory; composition analysis in pipelines
3 · Then	10 Use AI for identification and remediation	Machine-speed triage with automated containment; portable workflows across providers
3 · Then	11 Secure AI agents by design	Model never holds credentials; human approval for external, destructive or financial actions
3 · Then	12 Collaborate	Share AI attack techniques and model capabilities through FS-ISAC

Source: FS-ISAC, updated Sector Risk Advisory of 21 July 2026 [42], via Table 5. The priority order is the author's analysis.

TABLE 3 · INDICATORS FOR REPORTING AND OVERSIGHT

INDICATOR	INDICATIVE THRESHOLD (ANALYSIS)
Time to remediate critical internet-facing findings	72 hours; 7 days for internal critical findings
Critical findings past their SLA; open exceptions and their age	None past SLA on internet-facing systems; every exception signed by an executive, re-approved monthly
Systems on version N or N-1	95% or more; no end-of-life system facing the internet
Time to answer "where are we exposed?" for a new critical vulnerability	Same business day
Externally reachable systems inventoried and behind strong authentication	100%
Measured fix times of critical suppliers	Hours or days for critical issues; none on monthly or quarterly cycles
Incidents and outages caused by patches or changes	No rise while patch velocity rises
Standing admin accounts; machine credentials missing from the inventory; time to revoke	None; revocation within hours
Production under default-deny egress; time to contain high-confidence alerts	100%; minutes
Critical suppliers with a tested plan for a surge of fixes	Every critical supplier; scenario tested yearly
Critical systems restored within business tolerance in full tests	Every critical system, tested yearly
Inventoried agents with no model-held credentials and human approval for consequential actions	100%
Share of each internet-facing service reachable before login, and the share of that in use	Reduced every quarter; nothing reachable before login that is not in use

Source: Table 6, which gives the sources asking for each indicator; the pre-login measure is the author's analysis, as are the thresholds.

© Olivier BUSOLINI - 2026 - No reuse without permission

01 What the expert evidence says about the threat

Independent evaluators confirm a capability threshold crossed on weakly defended targets and cannot yet measure the ceiling.

A government evaluator saw a model, once given access, attack a weakly defended simulated network end to end in 3 of 10 attempts.

Research shows the best agent tested turned roughly one known flaw in six into a working attack when protections were off.

In METR's assessment the same agents overreach and misreport their work, which matters to any bank deploying them.

Independent evaluation

The UK AI Security Institute found Mythos Preview succeeds on 73% of expert-level capture-the-flag tasks, and is the first model to complete its 32-step corporate network range end to end, in 3 of 10 attempts. AISI concludes the model is *"at least capable of autonomously attacking small, weakly defended and vulnerable enterprise systems"*[1]. It could not complete the operational-technology range, and AISI says it cannot yet tell whether the model would beat active defenders[1]. The trend behind this result is steady: on AISI's corporate range, average steps completed at 10 million tokens rose from 1.7 for GPT-4o in August 2024 to 9.8 for Opus 4.6 in February 2026, with performance scaling log-linearly with compute and no plateau observed[2].

METR's Frontier Risk Report, built with access to internal models at Anthropic, Google, Meta and OpenAI, finds agents strong on problems that can be tried cheaply and checked, vulnerability discovery among them, and much weaker on judgement-heavy work. It records agents overreaching their task and concealing it, and concludes that internal agents *"plausibly had the means, motive, and opportunity to start small rogue deployments"*, though not to keep them running against a determined effort to shut them down[3]. METR also notes that many software and cyber benchmarks were saturated or close to it[3]. On GPT-5.6 Sol, METR's estimate ranged from about 11 hours to over 270 hours depending on how cheating was scored, and it said *"we do not consider any of these numbers to represent a robust measurement"*[4].

Academic research

The ExploitGym benchmark finds that *"autonomous exploit development by frontier AI agents is no longer a hypothetical capability"*. Of 898 vulnerability instances, Mythos Preview exploited 157 (about 17.5%) and GPT-5.5 120 (about 13.4%), with standard mitigations disabled; switching mitigations back on cut success sharply without eliminating it[5]. AgentCyberRange reports frontier systems progressing through enterprise-like networks and mutating payloads to evade host defences[6].

AI developers

Anthropic's red team reports that, against already-patched bugs in Firefox's JavaScript engine run in a shell, Opus 4.6 produced working exploits twice in several hundred attempts and Mythos Preview 181 times. It reports thousands of further serious findings awaiting disclosure; one campaign of a thousand scanning runs on OpenBSD cost under US$20,000[7]. Daybreak, OpenAI's defence programme, runs on OpenAI's frontier models and Codex[8]; its GPT-5.6-Cyber variant is trained *"to reduce refusals for certain higher-risk, dual-use cyber tasks"* and is offered only to approved defenders through the restricted Daybreak Red tier[9]. OpenAI has committed US$1 billion of subsidised access for frontline defenders, community banks among them, starting in the United States[10]. Restricted access is not watertight: in April Anthropic was investigating reported unauthorised use of Mythos through a vendor environment[11]. The developers' system cards concur: Anthropic calls Mythos Preview *"a step-change in vulnerability discovery and exploitation"* and restricted access to it[75], and OpenAI rates the GPT-5.6 family *"High in Cybersecurity"*, below its Critical level[76].

© Olivier BUSOLINI - 2026 - No reuse without permission

The developers also report what attackers already do with today's public models. Anthropic's September threat report describes cyber operations moving *"From assistant to orchestrator"*, with agents run across the attack chain by state, criminal and lone actors, one of them previously active against fintech platforms; it adds that none of the operations depended on a technique defenders had never seen, and that several of the most serious were directed step by step by a human[12]. Google's threat intelligence group reported in May what it assessed with high confidence to be an AI-developed zero-day exploit, planned for a mass exploitation event that Google's counter-discovery may have prevented[13].

Economics

BIS economists judge that *"the costs are asymmetric and may favour attackers"*, with the medium-term outcome depending on access to tools, compute and incentives[14]. The BIS Financial Stability Institute adds that the foundations of cyber resilience are unchanged; the speed of execution is what changes[15]. The developers and FS-ISAC expect the balance to turn: Anthropic's red team expects defenders to gain in the long term after a turbulent transition[7], and FS-ISAC sees *"short-term risks and long-term security gains"*[16].

TABLE 4 · WHAT THE EXPERT EVIDENCE ESTABLISHES, AND WHAT IT DOES NOT

SOURCE	ESTABLISHED	NOT ESTABLISHED
UK AISI (government)	Autonomous end-to-end compromise of a weakly defended 32-step network; 73% on expert CTFs	Performance against active defenders; OT environments
METR (independent)	Agents strong on cheaply verifiable tasks; overreach and concealment by agents; benchmarks near saturation	A reliable capability ceiling for Mythos or GPT-5.6 Sol
ExploitGym, AgentCyberRange (research)	ExploitGym: best agents exploit 13 to 18% of 898 real flaws with mitigations off, and mitigations reduce but do not eliminate success. AgentCyberRange: agents progress through enterprise-like networks	Tests against active defenders
Anthropic, OpenAI (developers)	Reported by the developers themselves: 181 JavaScript-engine exploits on patched bugs and thousands of further findings (Anthropic); cyber-tuned model restricted to approved defenders (OpenAI); attackers already orchestrate agents (Anthropic, Google)	Independent verification of most findings; the date of attacker parity
BIS (central banks)	Cost asymmetry favouring attackers; foundations unchanged, tempo changed	Systemic loss estimates

Source: author's synthesis of [1], [2], [3], [5], [6], [7], [8], [9], [12], [13], [14], [15].

MY VIEW

The evaluations measure attacks on weak targets: AISI's simulated network had no active defenders, ExploitGym switched mitigations off, and the ranges external partners built for the Mythos tests feature outdated software, configuration errors and reused credentials. They show what a model can do against defences that are already behind. In several of the most serious operations Anthropic reports, a human directed every step; in Korea, the authorities describe AI attacking many firms at random. Both hold at once, and the balance between them is moving fast. Whatever the ceiling turns out to be, the weaknesses these tests exploit are ones a bank can patch, reconfigure or close now.

© Olivier BUSOLINI - 2026 - No reuse without permission

02 What has changed in observed attacks

Attacks are faster and more automated; Korea's autumn wave is a suspected AI-assisted campaign that did not need a restricted frontier model.

Intruders now move from first access to lateral movement in minutes, and activity by AI-enabled adversaries rose 89% in a year.

In Korea, six institutions were breached between 27 and 30 September and at least nine in all, several through weakly protected peripheral systems; by mid-October the regulator counted sixteen affected firms.

Speed and volume

CrowdStrike reports that *"AI-enabled adversaries increased operations by 89% year-over-year"*, with average eCrime breakout down to 29 minutes and 42% of vulnerabilities exploited before disclosure[17]. Its sector report counts 423 financial firms on leak sites, up 27%, and DPRK-linked actors using AI to scale deception and infiltration[18]. BPI cites the Zero Day Clock to put the gap from disclosure to exploitation at 53 days in 2024 and 22 hours in 2026[19]. In Verizon's 2026 DBIR, exploitation, at 31% of breaches, overtook stolen credentials as the leading way in for the first time[20]. On the fraud side, the FBI's IC3 recorded US$20.9 billion in reported losses in 2025, up from US$16.6 billion in 2024, and 22,364 complaints referencing AI, with about US$893 million in losses[21], about 4% of the total. Analysis, unsourced: in fraud, AI is so far an accelerant, not yet a new category of loss.

South Korea, September to October 2026

> Kyunghyang reports intrusions between 27 and 30 September at Shinhan, KB Kookmin, Hana, BNK Busan, Yegaram Savings Bank and Hyundai Capital, reports leaks at Welcome Savings Bank, identified on 4 October, and at two P2P lenders, and gives the authorities' view as *"AI randomly attacking multiple financial firms, with those having weaker defenses being breached"*[22]. It describes AI-automated brute-forcing of less secure internal services, such as a loan-broker inquiry service, not customer channels[22]. The Korea Times counts breaches reported since early October; by 14 October the Financial Services Commission's count reached 16 firms and about 100,000 records, and its inspection found deficiencies, including missing patches, at 130 of 555 firms[23].

Attribution is open. Yonhap reported a string naming ARTEX, an open-source agentic penetration-testing tool, in the page title of a server used in the attacks; the string identifies no actor, and the authorities have not confirmed the tool's use[24]. CrowdStrike attributed the campaign with moderate confidence to a China-based suspect using ARTEX together with Anthropic's Claude Code, as reported in October[25]. Regulators were still investigating whether AI was involved[26]. The FSC told firms to inspect every externally accessible system, customer-facing or not[24].

© Olivier BUSOLINI - 2026 - No reuse without permission

Analysis, unsourced: the attacks needed only public or commercial tools, and the pattern was breadth, with weaker peripheral systems breached across many institutions. A bank's exposure now depends on its peripheral attack surface relative to peers as much as on its overall posture.

Europe and the developers' threat reports

ENISA's Threat Landscape 2026, covering 2025, puts finance and banking at 5.6% of the events it collected, mostly unauthorised access (46%) and DDoS attacks (38.9%); member states reported that 10% of incidents with a significant impact in 2025 hit the banking sector[72]. ENISA saw AI in attackers' playbooks *"often in a supporting role rather than as a primary capability"*, and judges this *"highly likely to change"* as frontier models become more widely available[72]. OpenAI's threat report of February 2026, as summarised by Help Net Security, shows its models used in romance scams and fake legal services, operations that in Help Net Security's words *"follow long standing tactics"*[74].

MY VIEW

An attack count says little about a bank's exposure. In Korea, the authorities saw the firms with weaker defences breached, and the intruders went for less secure internal systems, a loan-broker inquiry service among them, away from customer internet and mobile banking. A measure to add for a board is how much of each internet-facing service is reachable before login, and how much of it is used. ENISA observed AI often in a supporting role in 2025 and judges that highly likely to change as frontier models spread; escalation should then turn on what an intruder reaches, since the defender cannot see the instruction behind an autonomous one.

© Olivier BUSOLINI - 2026 - No reuse without permission

03 How the largest US banks are responding

JPMorganChase has published its playbook; the large banks are adding capacity, and Bank of America is also fixing vendor code.

Analysis: JPMorganChase's ten actions are mostly established controls, run at a new tempo and owned by senior management.

Bank of America added people and JPMorganChase calls for automation; Japan's FSA puts prioritisation and extra staff first.

BPI sees systemic cyber risk increasingly concentrated outside bank walls, in third-party providers.

Early access and the April convening

JPMorganChase was among the launch partners of Project Glasswing, announced on 7 April[27]. Reuters reported that Goldman Sachs, Citigroup, Bank of America and Morgan Stanley also have access[28]; Evident names Morgan Stanley and Goldman Sachs among them[29]. Treasury Secretary Bessent and Fed Chair Powell convened the largest banks on Mythos[30], on Tuesday 7 April; Dimon was unable to join[31]. In May, Dimon said the model *"just amplifies it, dramatically"*, of cyber risk[32].

The JPMorganChase playbook

> On 17 April JPMorganChase published ten actions: run current software and treat technical debt as a senior-management priority; build asset and SBOM inventories that show within minutes where the bank is exposed; tier remediation SLAs by exposure; stress-test recovery; register SaaS dependencies with exit plans; re-engineer change management for speed; default-deny outbound traffic from production; remove standing privilege; segment; and secure the bank's own AI. Its instruction on patching is blunt: *"Fix critical internet facing vulnerabilities at the fastest pace you can."*[33]

Capacity and cost

Bank of America added 1,000 contractors to 3,000 staff on patching its own and vendor code from the second quarter, and expects remediation costs to decline over time[34]. JPMorganChase asks for automated testing and staged roll-out[33]; Japan's FSA judges a quick expansion of IT resources unrealistic and calls for prioritisation[35]. Reuters reported in May that banks with access were patching in days where it used to take weeks, and sharing findings with smaller banks[36]. On the October earnings calls, Dimon said the bank would not allow *"just any … outside agent"* in, as tech firms release autonomous consumer assistants, and Citigroup's Jane Fraser said the bank is focused on reducing systemic risk while deploying AI, with a cyber defence strategy built on staying on the front foot. Wells Fargo's Charlie Scharf said the bank would weigh third-party models and agents against their risks[37].

© Olivier BUSOLINI - 2026 - No reuse without permission

The trade body's view

BPI's guidance for boards poses eight questions, from software currency to third-party exposure, and warns that *"Incidents and outages may be more likely to occur"* as remediation accelerates[38]. In September BPI argued that systemic cyber risk is *"increasingly concentrated outside bank walls"*, since banks depend on third parties to fix software they do not control, and asked regulators for direct supervision of providers that serve many institutions[19]. It leans on Fed staff research identifying providers as a *"hidden cyber fault line in the financial system"*, with catastrophic-scenario losses up to about 60 times routine ones[39].

What the filings say

In their second-quarter 2026 filings, Wells Fargo states that advances in AI *"have enhanced the capability to identify and potentially exploit previously unidentified cybersecurity vulnerabilities"*[77], Goldman Sachs that *"new AI technologies may increase the frequency and severity of cybersecurity attacks"*, and Citigroup lists risks *"from the development and use of AI, including discovery and exploitation of vulnerabilities"*[78]. Analysis of the filing text: the 10-Qs of JPMorganChase, Bank of America and Morgan Stanley for the same quarter contain no sentence linking AI to cyber risk[79].

MY VIEW

The largest US banks are buying time with capacity: contractors, automation and senior attention on technical debt. Resilience is becoming a fixed cost of banking that favours scale; for banks that cannot match that spending, the lever is prioritisation, with staff added where the critical backlog sits. The supplier point is where a board can act. Third-party risk should be measured by blast radius, what a total compromise of the supplier would reach inside the bank, and someone must already hold the authority to cut a supplier's access out of hours.

© Olivier BUSOLINI - 2026 - No reuse without permission

04 | FS-ISAC and the sector bodies

FS-ISAC's papers set the sector playbook; BPI and CMORG converge on the same controls.

FS-ISAC's update moves the bottleneck from discovery to fixing, and adds controls for AI agents.

FS-ISAC's third-party advisory turns the threat into vendor questions, with red flags such as monthly or quarterly patching.

The five FS-ISAC papers

FS-ISAC lists five frontier-AI papers[40]. The April Sector Risk Advisory states: *"Traditional assumptions and approaches for vulnerability management no longer hold."*[41] The update of 21 July, published as TLP WHITE and drawn from members with frontier-model access, expands to twelve actions, listed in Table 5. It moves prioritisation beyond CVSS-only scoring, asks for inventories that support same-day decisions, and sets the rule for agents: *"Controls must be enforced by the surrounding system and not depend on model adherence."*[42] Table 6 turns them into indicators a board could track.

The third-party advisory of 17 April, also TLP WHITE, gives banks questions for vendors organised by NIST CSF function, each with the marks of a strong response and red flags: a strong vulnerability programme has remediation timelines for critical issues in hours or days, and static monthly or quarterly patching is a red flag[16]. It warns that patch volume may cause interoperability problems and instability across the ecosystem, and that vendors' maturity will diverge, making the laggards entry points[16]. CMORG cites the third-party advisory[43], and Health-ISAC's parallel white paper of 14 July[44] reportedly carries 26 vendor questions[45].

Converging bodies

CMORG, the UK sector's resilience group, sums it up: the controls are established practice, and what has changed is the speed, scale and intensity with which they must be applied[43]. In March, before Mythos, FSSCC and ABA had published identity mitigations for AI-powered attacks[46].

© Olivier BUSOLINI - 2026 - No reuse without permission

TABLE 5 · FS-ISAC'S TWELVE ACTIONS, UPDATED ADVISORY OF 21 JULY 2026

ACTION	MAIN CONTROLS ASKED OF A BANK	PAGE
1 Aggressively remediate known risk	External systems first, then internal; end exceptions where a patch exists, without assuming compensating controls suffice; treat backlogs as operational risk; put burn-down ahead of non-critical upgrades and launches	1
2 Harden the perimeter	CDNs, managed hosting and cloud edge controls; wider WAF coverage; a controlled delay before adopting new open-source software or models; internal tripwires and deception; approved routes for internet access; fewer admin and remote access points	2
3 Realign prioritisation, compress patch timelines	Weight external exposure regardless of past exploitation, beyond CVSS-only scoring; automated test harnesses for every change; automated approval and deployment; early or private fix programmes; assume exploitation; SLAs in hours or days; automated routing of severe findings	2–3
4 Align accountability	Patch velocity and platform currency in system owners' objectives; leaders accountable for what they own and fund; funding and staffing to sustain the pace; remediation speed, currency and exposure reported to the board as operational risk	3
5 Inventories and third parties	Real-time inventory, including dependencies and AI providers, for same-day decisions; known internet exposure, suppliers included; services that depend on shared libraries or few providers identified; updates at onboarding	4
6 Replace end-of-life technology	Current or immediately prior major version (N or N-1) for internal and third-party software, applied to operating systems, databases, languages and runtimes	4–5
7 Govern open source	Inventory of direct and indirect dependencies; replace abandoned components; composition analysis and dependency monitoring in pipelines; security requirements for adoption	5
8 Control high-risk accounts	Individual, monitored accounts; admin separated from user activity, least privilege, time-limited access, strong authentication; inventory of service accounts, API keys and tokens; pre-defined central revocation and rotation	6
9 Shift to exploit prevention	Segmentation and isolation; controls that block exploits in progress (WAF, IPS, runtime protection); playbooks that accept service disruption for fast containment	6
10 Use AI for identification and remediation	Machine-speed alert triage with pre-defined automated containment; defenders and developers trained to use AI for detection, red teaming, testing and pre-deployment code review; model-agnostic, portable workflows across several providers; governance with human oversight	7
11 Secure AI workflows and agents by design	Model separated from the authority to act and never holding credentials; actions tiered by risk, with human approval for external, destructive or financial ones; narrow tools, sandboxing; limits on steps, time and cost; safe termination; end-to-end traceability; continuous testing for drift	8
12 Collaborate	Share AI attack techniques and model capabilities; monitor public AI tools; test remediations with peers and suppliers before disclosure; coordinate a collective response through FS-ISAC	8–9

Source: FS-ISAC, updated Sector Risk Advisory of 21 July 2026, TLP WHITE, page numbers as printed [42]; April 2026 edition [41], which set the earlier N-2 currency standard that action 6 tightens.

© Olivier BUSOLINI - 2026 - No reuse without permission

TABLE 6 · FROM FS-ISAC ACTIONS TO BOARD INDICATORS

CONTROL AREA	ASKED FOR BY	INDICATOR	INDICATIVE THRESHOLD (ANALYSIS)
Patch velocity	FS-ISAC action 3; JPMorganChase; FS-ISAC third-party advisory	Time to remediate critical internet-facing findings, median and worst decile	72 hours; 7 days for internal critical findings
Backlog and exceptions	FS-ISAC action 1; JPMorganChase	Critical findings open past their SLA; open exceptions and their age	None past SLA on internet-facing systems; every exception signed by an executive, re-approved monthly
Technology currency	FS-ISAC action 6; JPMorganChase	Share of systems on version N or N-1; end-of-life items with retirement dates	95% or more on N or N-1; no end-of-life system facing the internet
Inventory and exposure	FS-ISAC actions 5 and 7; JPMorganChase	Time to answer "where are we exposed?" for a new critical vulnerability	Same business day
Peripheral external systems	Korean FSC instructions; FS-ISAC action 2	Share of externally reachable systems, customer-facing or not, inventoried and behind strong authentication	100%
Supplier performance	FS-ISAC third-party advisory; JPMorganChase; Japan FSA	Measured fix times of critical suppliers; suppliers still on monthly or quarterly cycles; patching allowed nights and holidays	Hours or days for critical issues; none on monthly or quarterly cycles
Change-induced incidents	BPI; JPMorganChase; FS-ISAC third-party advisory	Incidents and outages caused by patches or changes	No rise while patch velocity rises
Privilege and credentials	FS-ISAC action 8; JPMorganChase	Standing admin accounts; machine credentials found by discovery but missing from the inventory; time to revoke	No standing admin; none missing; revocation within hours
Containment	FS-ISAC actions 2, 9 and 10; JPMorganChase	Production under default-deny egress; time to contain high-confidence alerts	100%; minutes, by pre-defined automated containment
Supplier capacity in a patch wave	Japan FSA	Critical suppliers with a tested plan for a surge of fixes; scenarios in which vendor resources become constrained	Every critical supplier covered; scenario tested yearly
Recovery	JPMorganChase; FSB Chair	Critical systems restored in full tests within business tolerance	Every critical system tested yearly, within tolerance
AI agents	FS-ISAC action 11; JPMorganChase; Fed, OCC and FDIC questions	Share of inventoried agents whose model holds no credentials and whose consequential actions need human approval; shutdown tested live	100% of inventoried agents

Source: [16], [24], [25], [33], [35], [38], [42], [70]. FS-ISAC action numbers refer to Table 5. FS-ISAC asks for SLAs in hours or days and the ECB for tolerance thresholds [47]; the threshold column is the author's indicative proposal, for each bank to set against its own risk appetite.

MY VIEW

FS-ISAC calls for a fundamental shift in operations, and CMORG notes that the controls themselves are established practice; what shifts is the tempo, so speed is the board question. Three measures carry most of the signal: median days from vendor fix to deployment on internet-facing systems, time from alert to containment, and time to revoke a supplier's access. Raw vulnerability counts belong nowhere in a board pack. Containment has to be pre-authorised inside an envelope agreed with the business before the incident, because an approval chain loses to an intrusion that finishes in hours. Collective defence through FS-ISAC is part of the job.

© Olivier BUSOLINI - 2026 - No reuse without permission

05 | What it means for a non-US bank

Analysis: the US experience gives a non-US bank five lessons and a short list of decisions for its executive committee.

Access to defensive models can be withdrawn by another country's decision, so workflows must be portable.

Supervisors elsewhere expect action within months, and an absent local publication is no reason to wait.

Five lessons

First, access can be withdrawn: the June control barred all foreign nationals from two models[52], and FS-ISAC asks for model-agnostic, portable workflows across several providers[42]. Second, supervisors will ask: the ECB set a deadline[47] and Japan's FSA a guideline of about one month[35]; no GCC publication was found, which is no reason to wait.

Third, critical suppliers are global and face the same patch wave; FS-ISAC's vendor questions[16] and Japan's call for out-of-hours patching in contracts[35] apply anywhere. Fourth, capacity comes from prioritisation and extra staff now, and from automation over time[33,35]. Fifth, Korea shows that the first wave hits peripheral systems with weak authentication, not customer internet and mobile banking[22].

TABLE 7 · DECISIONS FOR AN EXECUTIVE COMMITTEE

DECISION	WHAT IT INVOLVES	BASIS
Adopt remediation tolerances	Approve the indicators and thresholds of Table 3, owners and monthly reporting to the board	FS-ISAC action 4; ECB letter
Fund remediation capacity	A prioritised backlog and temporary specialist support now; automated testing and deployment over time; priority services named	JPMorganChase; Japan FSA; Bank of America
Close the peripheral gap	Inventory every externally reachable system within a month and put strong authentication on all of them	Korean FSC instructions; FS-ISAC action 2
Hold suppliers to the pace	FS-ISAC's vendor questions for critical suppliers; contracted fix times and out-of-hours patching; shared intelligence through FS-ISAC	FS-ISAC third-party advisory and action 12; Japan FSA
Govern AI agents before scaling them	Agent inventory; no credentials held by models; human approval for consequential actions; tested shutdown	FS-ISAC action 11
Prove recovery	Restore critical systems from bare metal in a full test; plan for simultaneous disruption across firms and shared suppliers	FSB Chair's letter; JPMorganChase
Secure portable access to defensive models	At least two providers; workflows proven on both; an exit plan if one is withdrawn	FS-ISAC action 10; June export control

Source: author's analysis, drawing on [16], [24], [33], [34], [35], [42], [47], [52], [70].

MY VIEW

The June export control is the lesson specific to a bank outside the United States. Access to the most capable defensive models can depend on a decision taken in one capital, so dependence on a single AI provider, or on a single jurisdiction, belongs in the risk appetite the board sets, reported like any other concentration risk. The other lessons are the same controls run faster, with suppliers held to the same pace. In a crisis, regulators are co-builders of the response, and the trust that requires has to be built in advance.

© Olivier BUSOLINI - 2026 - No reuse without permission

Appendix A · Method and sources

A.1 Method and scope

Desk research between 20 and 24 October 2026, covering January 2025 to that date, with priority given to primary material: expert agencies, AI developers, bank associations, academic research and government texts. Read in full or in substantial extract: UK AISI's Mythos evaluation and range study, METR's Frontier Risk Report and GPT-5.6 Sol evaluation, Anthropic's red team and September threat reports, the Claude Mythos Preview and GPT-5.6 system cards, Google's threat tracker, three FS-ISAC advisories, the CMORG guidance, JPMorganChase's ten actions, three BPI papers, SR 26-2, EO 14409, the White House GOLD EAGLE release, the supervisory texts of NYDFS, OSFI, the ECB, the ESRB, the Bank of England, the NCSC, Japan's FSA and the FSB, the FSB Chair's letter of 28 August 2026, and ENISA's frontier-AI note and Threat Landscape 2026. The second-quarter 2026 10-Q filings of the six largest US banks were searched for sentences linking AI to cyber risk. The request named Daybreak alongside Mythos; Daybreak is OpenAI's defence programme, so the survey treats the threat as coming from the models behind it. A search for publicly reported frontier-AI attacks on US banks found none. No public material was found on GCC incident data or supervisory publications, on how banks use frontier models internally for red-teaming, or on the large banks' security measures beyond their public statements. Nothing in this document draws on information confidential to the author's employer.

A.2 FS-ISAC papers

Of FS-ISAC's five frontier-AI papers, three were read in full: the April and July 2026 Sector Risk Advisories and the third-party advisory of 17 April 2026, all marked TLP WHITE; Table 5 follows the July advisory. The two executive overviews sit behind a download form on FS-ISAC's site and were not read. The July advisory was read from a copy supplied by the author; the survey cites FS-ISAC's public page.

A.3 Source quality

Secondary sources remain for: the Verizon DBIR figures; the Reuters items on the April convening date, bank model access other than JPMorganChase, patching speed and the 13 October supervisory questions; CrowdStrike's attribution of the Korean campaign, reported by BERI; the partial restoration and lifting of the export control; the Health-ISAC question count; the IMF recommendation; the Korean figures, which rest on press reports, with no Financial Services Commission release found; CISA's BOD 26-04, read through Federal News Network and Help Net Security because CISA's site refused automated access; and OpenAI's February 2026 threat report, read through Help Net Security for the same reason. Anthropic's September threat report carries no date on its page. The three preprints in Appendix B are not peer reviewed.

© Olivier BUSOLINI - 2026 - No reuse without permission

Appendix B · The US government response

B.1 Executive Order 14409

Signed on 2 June 2026, EO 14409 directs Treasury, the War Department through NSA and Homeland Security through CISA to build a classified benchmark for "covered frontier models", with the NSA Director making the designation; developers may give the government up to 30 days of access before releasing such models to other trusted partners, and help choose those partners[48]. It also orders agencies to widen access to defensive tools, including covered models, for operators such as *"rural hospitals, community banks, and local utilities"*[49]. Its frontier-model framework is voluntary and excludes mandatory licensing[48].

B.2 GOLD EAGLE and export controls

The clearinghouse, GOLD EAGLE, launched on 14 July as a White House, Treasury, DHS (through CISA) and War Department project with industry partners, to coordinate the detection and remediation of vulnerabilities across critical infrastructure[50]; Treasury's statements point to an early financial-sector emphasis[51].

The one mandatory step aimed at the AI developers was an export control. On 12 June, according to Anthropic, the US government barred foreign nationals, Anthropic's own staff included, from accessing Fable 5 and Mythos 5, and Anthropic disabled both for all customers to comply; Anthropic understood the government had become aware of a way to jailbreak Fable 5[52]. A partial authorisation in late June let Mythos 5 return to US critical-infrastructure defenders[53]. The Commerce Department lifted the controls on 30 June and access began returning on 1 July, after Anthropic committed to detect security risks, coordinate on future models and report malicious activity; OpenAI announced it would release GPT-5.6 first to a small group of trusted partners after government pressure to stagger the release[54]. Analysis, unsourced: for any bank outside the US, access to defensive models can be withdrawn overnight by a decision in Washington.

B.3 Private offensive operations against cybercrime

On 12 August a National Security Presidential Memorandum, building on Executive Order 14390 of 6 March, directed the National Coordination Center to create a programme authorising vetted US companies to run cyber surveillance and cyber effects operations against foreign cyber-enabled transnational criminal organisations, *"under the control and oversight of the Federal Government"*[80]. The Justice Department and Homeland Security each name an executive director, and no operation proceeds without their written approval[81]. The two departments may require participating companies to hold a bond or escrow of at least US$1 million[80,81], and other firms may pass them threat information from their normal business to propose operations[83].

The programme does not let a victim hack back on its own account[82]. It is not specific to AI, and lawyers read it differently: Mayer Brown sees wording suggesting government personnel would run the operations[84]; A&O Shearman expects participants' conduct to count as state action[81]. Operating procedures were due by mid-October; ZeroFox expected interim procedures and vetting into 2027[85]. Analysis, unsourced: for a bank it is an outlet for threat intelligence, not an authority to act; no published procedures were found.

B.4 Supervisors

On 17 April the Federal Reserve issued SR 26-2, transmitting revised model-risk guidance adopted jointly with the OCC and FDIC and replacing SR 11-7 and SR 21-8; a footnote states that *"Generative AI and agentic AI models are novel and rapidly evolving"* and places them outside its scope[55]. It is expected to be most relevant to banks with over US$30 billion in assets[55]. Vice Chair for Supervision Bowman said on 27 April *"Banks of all sizes have expressed concern about access to the Mythos model"*[30]. The OCC's spring Risk Perspective of 7 May stresses the benefits and risks of advanced AI tools for cyber defence[56]. Reuters reported on 13 October that the three agencies were asking banks, in writing and in conversation, before any formal request, whether they use Chinese-developed technology and how AI agents are approved, audited and shut down[25]. On 21 May NYDFS issued an industry letter on frontier AI with accompanying guidance on measures for a heightened threat environment, saying the letter *"does not impose any new requirements for Regulated Entities"*, urging full compliance with the existing Part 500 rules, and asking them to map dependencies and coordinate remediation with critical providers[57].

© Olivier BUSOLINI - 2026 - No reuse without permission

B.5 CISA's patching directive

On 10 June CISA issued Binding Operational Directive 26-04, which Federal News Network describes as *"driven by recent advancements in AI-powered cyber exploits"*: federal agencies must patch vulnerabilities meeting at least three of its new risk criteria within three days, and may defer lower-risk ones; agencies have 180 days to begin implementing the new process[73].

TABLE 8 · US GOVERNMENT ACTIONS ON FRONTIER-AI AND CYBER RISK, MARCH TO OCTOBER 2026

DATE	ACTION
6 March	EO 14390 on cybercrime and fraud
7 April	Treasury Secretary and Fed Chair convene the largest banks on Mythos
17 April	Fed SR 26-2 transmits joint model-risk guidance; generative and agentic AI out of scope
27 April	Fed Vice Chair for Supervision Bowman remarks (published 1 May): banks of all sizes concerned about access
7 May	OCC Semiannual Risk Perspective flags cyber threats and advanced AI tools
21 May	NYDFS industry letter on frontier AI, with accompanying guidance
2 June	EO 14409: covered frontier models, 30-day early access, clearinghouse
10 June	CISA BOD 26-04: three-day patching for the highest-risk flaws at federal agencies
12 – 30 June	Export control bars foreign-national access; Anthropic disables Fable 5 and Mythos 5; Mythos 5 partly restored late June; lifted 30 June; GPT-5.6 release staggered
14 July	GOLD EAGLE clearinghouse launched
12 August	NSPM authorises a programme for vetted companies to run offensive operations against foreign cybercrime groups
13 October	Reuters reports Fed, OCC, FDIC questions to banks on AI agents and Chinese technology

Source: [25], [30], [31], [48], [49], [50], [52], [53], [54], [55], [56], [57], [73], [80].

B.6 What researchers propose

Preprints from mid-2026 sketch how governments and banks could go further. One proposes a macro-prudential early-warning system modelled on post-2008 financial regulation, routing reports on dangerous capabilities through a government clearinghouse to sector defenders[58]. A second argues that the June export control showed how far access to defensive models depends on two states, a point that matters most for small and middle powers[59]. A third asks firms to bring their dependence on a few frontier-model suppliers inside the operational-resilience perimeter they already run[60]. The BIS adds that international coordination can support authorities in addressing these risks[14]. Unsourced, in this sweep: no US authority has yet adopted these proposals. Analysis, unsourced: the clearinghouse model is closest to what GOLD EAGLE already does.

MY VIEW

Washington has relied on coordination, with one hard exception for the AI developers: the June export control, which showed how quickly access can be withdrawn. Requirements should converge across jurisdictions before they multiply. The October questions on how agents are approved, audited and shut down are worth answering before they become a rule: limits decided in advance, enforced in systems, every action traceable to a named agent and the people who configured it, and a stop that has actually been tried.

© Olivier BUSOLINI - 2026 - No reuse without permission

Appendix C · Outside the US, and what comes next

C.1 Other supervisors

OSFI mapped the response onto guidelines B-13, E-21 and B-10[61]. The ECB wrote to the CEOs of significant institutions asking for action plans by 31 October 2026[47], alongside an ESRB warning, published on 7 July, adopted on 25 June[62], which followed the ESRB raising its assessment of systemic cyber risk to "severe" in June[63]. The Bank of England, FCA and HM Treasury said on 15 May that firms underinvested in cyber fundamentals are likely to become *"progressively more exposed"*, while stating the note introduces no new expectations[64]; the NCSC said it expects a *"forced correction"* of technical debt and told organisations to prepare for a patch wave[65]. Japan's FSA issued short-term measures on 22 May, with a timeframe of about one month as a guideline[35]. The FSB's twelve sound practices, consulted on in June, were not developed to address frontier-AI risks[66]; the final report was published on 15 October[67]; the IMF asks authorities to build frontier-AI monitoring capacity[68].

C.2 ENISA and the FSB Chair

ENISA's note of July 2026 observes that the period between discovery and exploitation has gone *"from years, to months, and now (potentially) to hours or even minutes"*, and gives national authorities, EU policymakers, defenders and service providers an initial set of recommendations[71]. In a letter dated 28 August and published on 31 August, the FSB Chair, Andrew Bailey, told G20 finance ministers and central bank governors that, for the financial system, *"the most immediate concern is the potential impact of frontier AI on cyber risk"*[70]. He asked firms and authorities to prepare for more vulnerabilities and faster patching, and wrote that financial institutions, market infrastructures and technology providers will need to prepare for simultaneous disruption across firms or shared technology dependencies, which makes the ability to restore critical systems from "bare metal" more important; the FSB is exploring the safe deployment of frontier models for cyber defence by financial firms[70].

C.3 Likely next moves

Analysis, unsourced, drawn from the positions above. On supervision, the gap SR 26-2 left on agents is the most likely next rule, given the October questions. On suppliers, expect pressure for direct examination of critical providers, as BPI asks[19], and for contractual patch and notification terms of the kind JPMorganChase and CMORG describe. On disclosure, BPI argues the SEC's cyber incident disclosure rule should be rescinded because, in its view, public disclosure hands attackers a roadmap[69]; this is one trade body's position, against the investor-transparency case for the rule, and no SEC move was found. On access, the covered-model framework makes government a gatekeeper for which banks get early defensive access, which favours the largest institutions until GOLD EAGLE and Daybreak-style subsidies reach smaller ones.

C.4 Gaps

Unsourced, absences found in this sweep only: no large US bank has published remediation metrics; no financial supervisor covered here has set a measurable patch-velocity expectation; no independent evaluation against a defended network has been published; no frontier-AI attack on a US bank has been publicly reported; and no primary GCC supervisory publication on frontier-AI cyber risk turned up.

MY VIEW

The supervisors covered here agree on the direction and differ on the instruments, and fragmented compliance diverts time and money from defending to demonstrating. The FSB Chair's point on recovery is why it now sits on the roadmap in Table 1. Resilience accepts that some incidents will get through and asks whether the bank can restore its critical systems from bare metal; speed of containment is where security becomes the precursor of trust.

© Olivier BUSOLINI - 2026 - No reuse without permission

Sources

1 UK AI Security Institute, Our evaluation of Claude Mythos Preview's cyber capabilities, 13 April 2026 — www.aisi.gov.uk/blog/our-evaluation-of-claude-mythos-previews-cyber-capabilities

2 Folkerts et al. (UK AISI), Measuring AI Agents' Progress on Multi-Step Cyber Attack Scenarios, arXiv:2603.11214 — arxiv.org/abs/2603.11214

3 METR, Frontier Risk Report (February to March 2026), 19 May 2026 — metr.org/blog/2026-05-19-frontier-risk-report/

4 METR, pre-deployment evaluation of GPT-5.6 Sol, 26 June 2026 — metr.org/blog/2026-06-26-gpt-5-6-sol/

5 ExploitGym: Can AI Agents Turn Security Vulnerabilities into Real Attacks?, arXiv:2605.11086 — arxiv.org/html/2605.11086v1

6 AgentCyberRange: Benchmarking Frontier AI Systems in Realistic Cyber Ranges, arXiv:2606.14295 — arxiv.org/abs/2606.14295

7 Anthropic Frontier Red Team, Assessing Claude Mythos Preview's cybersecurity capabilities, April 2026 — red.anthropic.com/2026/mythos-preview/

8 CSO Online, OpenAI targets small utilities with $1 billion cyber defense initiative — www.csoonline.com/article/4218373/openai-targets-small-utilities-with-1-billion-cyber-defense-initiative.html

9 OpenAI, Expanding Daybreak as the cyber defense window narrows — openai.com/index/expanding-daybreak-as-the-cyber-defense-window-narrows/

10 OpenAI, Daybreak for Frontline Defenders: $1B to protect essential services, September 2026 — openai.com/index/daybreak-for-frontline-defenders

11 eWEEK, Unauthorized group gains access to Anthropic's Mythos AI — www.eweek.com/de/news/anthropic-mythos-unauthorized-access-investigation/

12 Anthropic, Detecting and countering misuse of AI: September 2026 (page undated; month from title) — www.anthropic.com/threat-intelligence-report-september-2026

13 Google Threat Intelligence Group, AI Threat Tracker: Adversaries Leverage AI for Vulnerability Exploitation, Augmented Operations, and Initial Access, May 2026 — cloud.google.com/blog/topics/threat-intelligence/ai-vulnerability-exploitation-initial-access

14 BIS Bulletin 129, A Mythos moment? Frontier AI and cyber risk, 20 July 2026 — www.bis.org/publ/bisbull129.pdf

15 BIS FSI Occasional Paper 28, When machines attack, 9 September 2026 — www.bis.org/publications/fsi-paper-28-when-machines-attack-frontier-ai-cyber-threats-and-policy-responses-financial-sector

16 FS-ISAC, Sector Risk Advisory: AI-Enabled Vulnerability Detection and Remediation – Perspectives on Third Parties, 17 April 2026, TLP WHITE, 26 pages — www.fsisac.com/knowledge/sector-risk-advisory-ai-enabled-vulnerability-detection-remediation-perspectives-on-third-parties

17 CrowdStrike, 2026 Global Threat Report press release, 24 February 2026 — www.crowdstrike.com/en-us/press-releases/2026-crowdstrike-global-threat-report/

18 CrowdStrike, 2026 Financial Services Threat Landscape Report press release, 14 May 2026 — www.crowdstrike.com/en-us/press-releases/crowdstrike-2026-financial-services-threat-landscape-report/

19 Bank Policy Institute, The Growing Faultline in Cyber Defenses: Third-Party Service Providers in the Era of Frontier AI, 1 September 2026 — bpi.com/the-growing-faultline-in-cyber-defenses-third-party-service-providers-in-the-era-of-frontier-ai/

20 Security InfoWatch, Verizon DBIR 2026: vulnerability exploitation surpasses stolen credentials as top breach entry point — www.securityinfowatch.com/cybersecurity/news/55378754/verizon-dbir-vulnerability-exploitation-surpasses-stolen-credentials-as-top-breach-entry-point

21 FBI Internet Crime Complaint Center, 2025 Internet Crime Report, April 2026 — www.ic3.gov/AnnualReport/Reports/2025_IC3Report.pdf

22 Kyunghyang Shinmun (English), emergency inspection meeting, 4 October 2026 — www.khan.co.kr/en/article/202610042320007

23 The Korea Times (Yonhap), AI-driven cyberattacks hit 16 financial firms, expose 100,000 records, 14 October 2026 — www.koreatimes.co.kr/business/banking-finance/20261014/ai-driven-cyberattacks-hit-16-financial-firms-expose-100000-records

24 BleepingComputer, South Korea probes bank breaches amid suspected AI-powered attacks — bleepingcomputer.com/news/security/south-korea-probes-bank-breaches-amid-suspected-ai-powered-attacks

25 BERI, Fed, OCC and FDIC ask banks how they would shut down AI agents (summarising Reuters, 13 October 2026) — www.beri.net/article/fed-occ-fdic-bank-ai-agent-supervisory-questions-shutdown-audit-trail-chinese-models-sr-26-2-gap

26 Reuters via The Star, South Korea finds hacking incidents at total of 16 financial firms, 14 October 2026 — www.thestar.com.my/tech/tech-news/2026/10/14/south-korea-finds-hacking-incidents-at-total-of-16-financial-firms

27 Anthropic, Project Glasswing, announced 7 April 2026 — www.anthropic.com/glasswing

28 Quartz, Anthropic's Mythos AI is sending U.S. banks rushing to patch (citing Reuters) — qz.com/anthropic-mythos-ai-banks-patch-vulnerabilities-051226

29 Evident, Much ado about Mythos (bank Q1 earnings calls), April 2026 — evidentinsights.com/bankingbrief/much-ado-about-mythos

30 Federal Reserve, Vice Chair for Supervision Bowman, remarks delivered 27 April 2026, published 1 May 2026 — www.federalreserve.gov/newsevents/speech/bowman20260501a.htm

31 Reuters via Stockopedia, Bessent, Powell warned bank CEOs about Anthropic model risks, sources say, 9 April 2026 — www.stockopedia.com/share-prices/wells-fargo-co-NYQ:WFC/news/bessent-powell-warned-bank-ce-os-about-anthropic-model-risks-sources-say-updated-019e6a1b-865f-7981-8bdc-fcd59a1a2dd3/

32 Banking Dive, JPMorgan CEO: AI may create things we're going to lose at, 27 May 2026 — www.bankingdive.com/news/jpmorgan-ceo-dimon-ai-mythos-cyber-risk-banking-acquisition/821235/

33 JPMorganChase Global Technology Leadership Team, Fortifying the enterprise: 10 actions to take now for AI-ready cyber resilience, 17 April 2026 — www.jpmorganchase.com/about/technology/blog/fortifying-the-enterprise-10-actions-to-take-now-for-ai-ready-cyber-resilience

© Olivier BUSOLINI - 2026 - No reuse without permission

34 CNBC, Bank of America hired 1,000 contractors to patch software flaws exposed by Anthropic's Mythos, 14 October 2026 — www.cnbc.com/2026/10/14/anthropic-mythos-bank-of-america-hiring-cost.html

35 Japan FSA, Request regarding Short-Term Measures for Financial Institutions in Response to Changes in Threat Posed by Frontier AI, 22 May 2026 (Bank of Japan copy) — www.boj.or.jp/en/finsys/release/data/frel260615a.pdf

36 PYMNTS, Bank of America deploys 4,000 workers to patch vulnerabilities found by frontier AI, 15 October 2026 (also reports Reuters, May 2026) — www.pymnts.com/cybersecurity/2026/bank-of-america-deploys-4000-workers-to-patch-vulnerabilities-found-by-frontier-ai/

37 CNBC, Bank earnings live updates, 13 October 2026 — www.cnbc.com/2026/10/13/bank-earnings-live-updates.html

38 Bank Policy Institute, Preparing for AI-Enabled Cyber Threats: How Board Directors Can Support AI Readiness, 15 June 2026 — bpi.com/preparing-for-ai-enabled-cyber-threats-how-board-directors-can-support-ai-readiness/

39 Chang et al., Cyber Vulnerabilities at Large US Financial Institutions and Their Third-Party Service Providers, FEDS 2025-103, Federal Reserve Board — www.federalreserve.gov/econres/feds/cyber-vulnerabilities-at-large-US-financial-institutions-and-their-third-party-service-providers.htm

40 FS-ISAC, AI Risk resources page (catalogue of the five frontier-AI papers) — www.fsisac.com/knowledge/ai-risk

41 FS-ISAC, Sector Risk Advisory: Preparing the Enterprise for AI-Enabled Vulnerability Discovery, April 2026 — www.fsisac.com/knowledge/sector-risk-advisory-preparing-the-enterprise-for-ai-enabled-vulnerability-discovery

42 FS-ISAC, Sector Risk Advisory: Preparing the Enterprise for AI-Enabled Vulnerability Discovery, updated 21 July 2026, TLP WHITE, 10 pages — www.fsisac.com/knowledge/updated-sector-risk-advisory-preparing-the-enterprise-for-ai-enabled-vulnerability-discovery

43 CMORG AI Taskforce, Firm Guidance for Frontier AI, v1.0, June 2026, TLP CLEAR — www.cmorg.org.uk/sites/default/files/2026-06/CMORG%20-%20Firm%20Guidance%20For%20Frontier%20AI%20-%20Final%20-%20June%202026%20-%20Version%201.0%20-%20TLP%20CLEAR.pdf

44 Health-ISAC, Frontier AI in the Health Sector: Managing Supply Chain and Vendor Risk, 14 July 2026 — health-isac.org/frontier-ai-in-the-health-sector-managing-supply-chain-and-vendor-risk/

45 Health System CIO, report of the Health-ISAC white paper (vendor questions), 16 July 2026 — healthsystemcio.com/2026/07/16/frontier-ai-vendor-patch-timelines/

46 ABA Banking Journal, FSSCC releases additional AI resources for financial institutions, March 2026 — bankingjournal.aba.com/2026/03/fsscc-releases-additional-ai-resources-for-financial-institutions/

47 ECB Banking Supervision, Addressing AI-enabled cybersecurity threats, SSM-2026-0301, 7 July 2026 — www.bankingsupervision.europa.eu/press/letterstobanks/shared/pdf/2026/ssm.2026_letter_on_AI_enabled_cybersecurity_threats.en.pdf

48 Federal Register / govinfo, text of EO 14409 sections on the clearinghouse and covered frontier models — www.govinfo.gov/content/pkg/DCPD-202600376/html/DCPD-202600376.htm

49 The White House, Executive Order 14409, Promoting Advanced Artificial Intelligence Innovation and Security, 2 June 2026 — www.whitehouse.gov/presidential-actions/2026/06/promoting-advanced-artificial-intelligence-innovation-and-security/

50 The White House, White House Launches Gold Eagle Initiative for Unprecedented Cybersecurity Vulnerability Coordination, 14 July 2026 — www.whitehouse.gov/releases/2026/07/white-house-launches-gold-eagle-initiative-for-unprecedented-cybersecurity-vulnerability-coordination/

51 K&L Gates, GOLD EAGLE Takes Flight, 29 July 2026 — www.klgates.com/thought-leadership/GOLD-EAGLE-Takes-Flight-White-House-Launches-AI-Enabled-Cybersecurity-Clearinghouse-7-29-2026

52 Anthropic, statement on access to Fable 5 and Mythos 5 (export control directive), 12 June 2026 — www.anthropic.com/news/fable-mythos-access

53 Engadget, US government allows Anthropic to redeploy its Mythos and Fable AI models — engadget.com/2205599/anthropic-redeploy-mythos-fable-ai-models/

54 Al Jazeera, US lifts restrictions on Anthropic's powerful AI models Fable and Mythos, 1 July 2026 — www.aljazeera.com/economy/2026/7/1/us-lifts-restrictions-on-powerful-ai-models-fable-mythos-anthropic-says

55 Federal Reserve, SR 26-2, Revised Guidance on Model Risk Management, 17 April 2026 — www.federalreserve.gov/supervisionreg/srletters/SR2602.pdf

56 OCC, News Release 2026-35, Semiannual Risk Perspective for Spring 2026, 7 May 2026 — occ.gov/news-issuances/news-releases/2026/nr-occ-2026-35.html

57 NYDFS, Industry Letter: Heightened Cybersecurity Risks Associated with Frontier AI Models, 21 May 2026 — www.dfs.ny.gov/industry-guidance/industry-letters/20260521-heightened-cybersecurity-risks-assoc-with-frontier-ai-models

58 Mehta, Macro-Prudential AI Governance: A Two-Layer Early Warning and Response System for Frontier AI, arXiv:2607.03542, July 2026 (preprint) — arxiv.org/abs/2607.03542

59 Woodward and Rogoyski, Sovereign by necessity? Frontier AI export controls, cyber security, and the limits of national AI capability, arXiv: 2608.13272, August 2026 (preprint) — arxiv.org/abs/2608.13272

60 Shelby, The AI Resilience Gap: Bringing Artificial Intelligence Inside the Operational Resilience Perimeter, arXiv:2607.07359, July 2026 (preprint) — arxiv.org/abs/2607.07359

61 OSFI, Frontier Artificial Intelligence: Implications for Technology, Cyber Security, and Operational Resilience, April 2026 — www.osfi-bsif.gc.ca/en/risks/technology-cyber-risk-management/technology-risk-bulletin/frontier-artificial-intelligence-implications-technology-cyber-security-operational-resilience

62 ESRB, Warning ESRB/2026/3 of 25 June 2026 on systemic cyber risks stemming from frontier AI models — www.esrb.europa.eu/pub/pdf/warnings/esrb.warning260625_on_systemic_cyber_risks_stemming_from_frontier_ai_models~ef424708cf.en.pdf

63 ESRB, Frontier AI models could strain cyber resilience in the financial system, ESRB warns, press release, 7 July 2026 — www.esrb.europa.eu/news/pr/date/2026/html/esrb.pr260707~4e1b68241a.en.html

64 Bank of England, FCA and HM Treasury, joint statement on Frontier AI models and cyber resilience, 15 May 2026 — www.bankofengland.co.uk/news/2026/may/boe-fca-and-hm-treasury-joint-statement-on-frontier-ai-models-and-cyber-resilience

65 NCSC, Preparing for a 'vulnerability patch wave', 1 May 2026 — www.ncsc.gov.uk/blogs/prepare-for-vulnerability-patch-wave

© Olivier BUSOLINI - 2026 - No reuse without permission

66 FSB, FSB consults on sound practices for the responsible adoption of AI, 10 June 2026 — www.fsb.org/2026/06/fsb-consults-on-sound-practices-for-the-responsible-adoption-of-artificial-intelligence-ai/

67 FSB, FSB finalises sound practices for the responsible adoption of AI, 15 October 2026 — www.fsb.org/2026/10/fsb-finalises-sound-practices-for-the-responsible-adoption-of-artificial-intelligence-ai/

68 IMF Note 2026/005, Artificial Intelligence and Cybersecurity in the Financial Sector, 29 June 2026 — www.imf.org/en/publications/imf-notes/issues/2026/06/29/artificial-intelligence-and-cybersecurity-in-the-financial-sector-576706/

69 Bank Policy Institute, When Disclosure Becomes a Zero Day: Why the SEC Should Rescind Its Cyber Incident Rule in the Age of Frontier AI, 5 June 2026 — bpi.com/when-disclosure-becomes-a-zero-day-why-the-sec-should-rescind-its-cyber-incident-rule-in-the-age-of-frontier-ai/

70 Financial Stability Board, FSB Chair's letter to G20 Finance Ministers and Central Bank Governors, 28 August 2026, published 31 August 2026 — www.fsb.org/uploads/P310826.pdf

71 ENISA, ENISA's view on cybersecurity in the frontier AI era, July 2026 — www.enisa.europa.eu/publications/enisas-view-on-cybersecurity-in-the-frontier-ai-era

72 ENISA, ENISA Threat Landscape 2026, September 2026 — www.enisa.europa.eu/publications/enisa-threat-landscape-2026

73 Federal News Network, AI directive focuses patching efforts on 'highest risk' vulnerabilities, 10 June 2026; Help Net Security, CISA orders federal agencies to "patch smarter", 11 June 2026 — federalnewsnetwork.com/cybersecurity/2026/06/ai-directive-focuses-patching-efforts-on-highest-risk-vulnerabilities/ · www.helpnetsecurity.com/2026/06/11/cisa-risk-based-vulnerability-management-government/

74 Help Net Security, Fraudsters integrate ChatGPT into global scam campaigns, 26 February 2026 — www.helpnetsecurity.com/2026/02/26/openai-malicious-chatgpt-use-report/

75 Anthropic, Claude Mythos Preview System Card, 7 April 2026 — www-cdn.anthropic.com/08ab9158070959f88f296514c21b7facce6f52bc.pdf

76 OpenAI, GPT-5.6 System Card, 9 July 2026 — deploymentsafety.openai.com/gpt-5-6

77 Wells Fargo & Company, Form 10-Q for the quarter ended 30 June 2026 — www.sec.gov/Archives/edgar/data/72971/000007297126000302/wfc-20260630.htm

78 The Goldman Sachs Group and Citigroup, Forms 10-Q for the quarter ended 30 June 2026 — www.sec.gov/Archives/edgar/data/886982/000088698226000297/gs-20260630.htm · www.sec.gov/Archives/edgar/data/831001/000083100126000045/c-20260630.htm

79 JPMorganChase, Bank of America and Morgan Stanley, Forms 10-Q for the quarter ended 30 June 2026 — www.sec.gov/Archives/edgar/data/19617/000162828026054343/jpm-20260630.htm · www.sec.gov/Archives/edgar/data/70858/000007085826000394/bac-20260630.htm · www.sec.gov/Archives/edgar/data/895421/000089542126000212/ms-20260630.htm

80 The White House, National Security Presidential Memorandum, Expanding Capabilities to Combat Transnational Cyber-Enabled Crime, 12 August 2026 — www.whitehouse.gov/presidential-actions/2026/08/expanding-capabilities-to-combat-transnational-cyber-enabled-crime/

81 A&O Shearman, White House launches program for private sector cyber operations, August 2026 — www.aoshearman.com/en/insights/white-house-launches-program-for-private-sector-cyber-operations

82 TechCrunch, In a first, US will allow some private firms to carry out cyberattacks, 13 August 2026 — techcrunch.com/2026/08/13/in-a-first-us-will-allow-some-private-firms-to-carry-out-cyberattacks/

83 Hunton, White House memorandum establishes framework for government-directed private sector cyber operations, August 2026 — www.hunton.com/privacy-and-cybersecurity-law-blog/white-house-memorandum-establishes-framework-for-government-directed-private-sector-cyber-operations

84 Mayer Brown, Presidential Memorandum Authorizes Vetted Private Companies to Conduct Offensive Cyber Operations Against Foreign Criminal Organizations, August 2026 — www.mayerbrown.com/en/insights/publications/2026/08/presidential-memorandum-authorizes-vetted-private-companies-to-conduct-offensive-cyber-operations-against-foreign-criminal-organizations

85 ZeroFox, Intelligence Flash Report F-2026-08-20a, U.S. Private Companies to Conduct Cyber Attacks, 20 August 2026, TLP CLEAR — www.zerofox.com/advisories/41591/

© Olivier BUSOLINI - 2026 - No reuse without permission