Deloitte.

Banking Trend Radar Webcast

Episode 20: Trust under Pressure: “Cyber Security in the Age of AI”

MAKING AN IMPACT THAT MATTERS since 1845

Unsere heutigen Referenten

Wir freuen uns auf Webcast & Diskussion!

Malko Steinorth

Partner

Cyber Defense & Resilience Lead Germany

msteinorth@Deloitte.de

Max Kaiser

Director

Cyber Banking & Capital Markets Lead Germany

mkaiser@Deloitte.de

Thomas Peek

Partner

Strategy, Risk & Transformation | Risk, Regulatory & Forensic

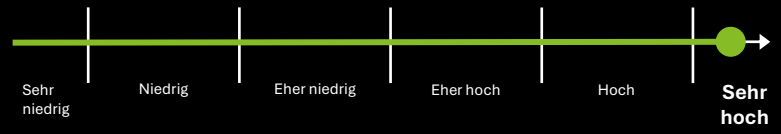
tpeek@Deloitte.de

Trend Assessment

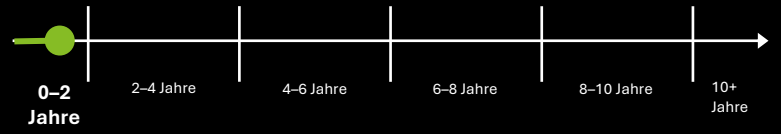
Vertrauen unter Druck: Cybersicherheit im Zeitalter der KI

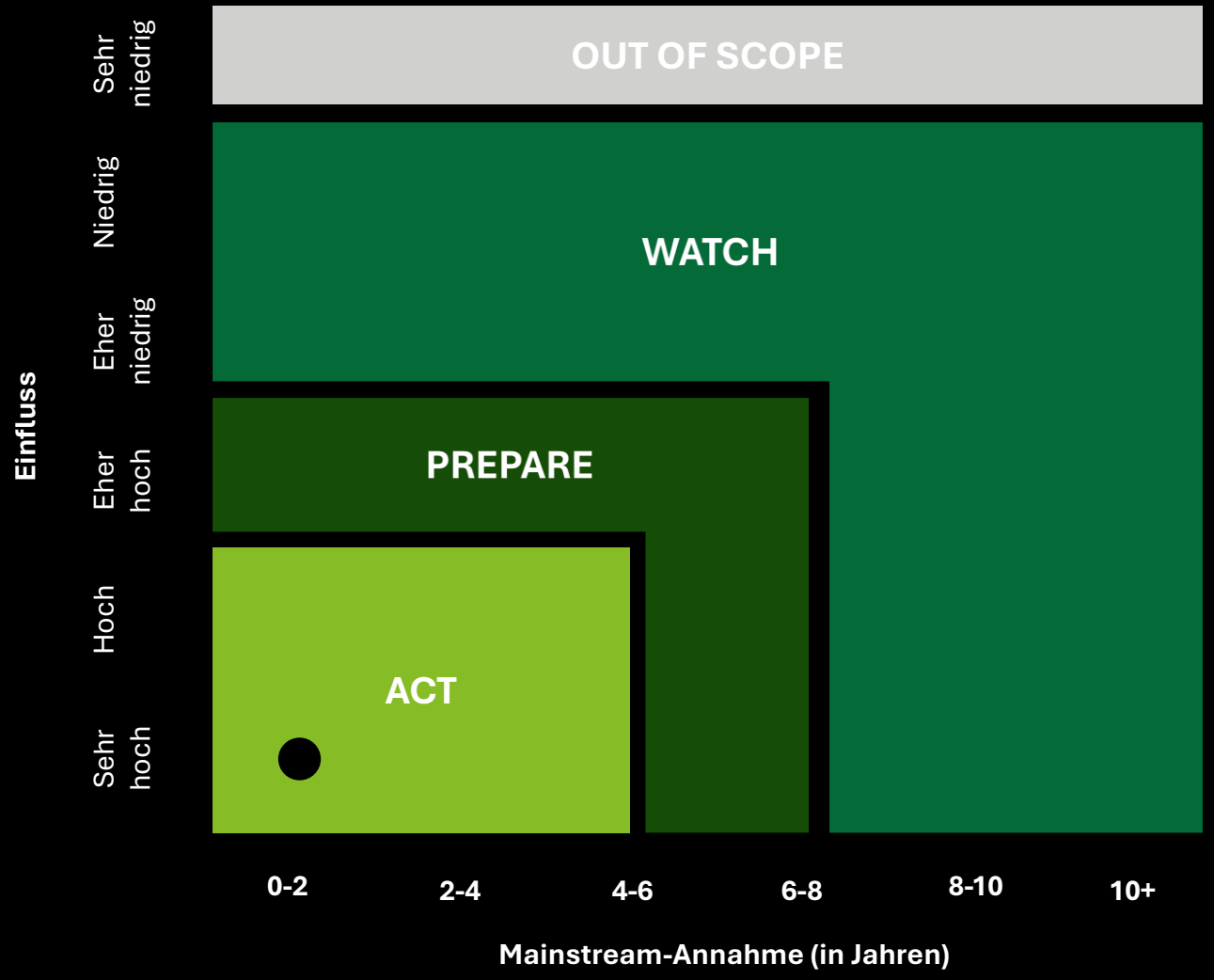
PREPARE

Einfluss



Zeitpunkt der Mainstream-Annahme





Deloitte CYBER · FINANCIAL SERVICES LEADERSHIP KEYNOTE

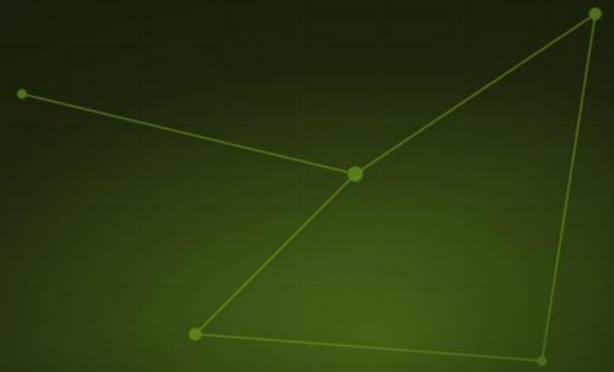
Trust under Pressure.

Cyber Security in the Age of AI

Malko Steinorth Partner, Cyber Defense & Resilience Leader, Deloitte Germany

Max Kaiser Director, Cyber Banking & Capital Markets Leader, Deloitte Germany

16 July 2026



THE PREMISE

In banking, trust is the product. AI just rewrote the cost of breaking it — for attackers *and* defenders.

The state of play

Agentic AI has left the lab. Machines now act — on both sides of the firewall.

The threat to FSI

Identity, money and information can now be forged at industrial scale and speed.

The opportunity

The same technology hands disciplined institutions a decisive defender's advantage.

WHERE WE ARE · THE STATE OF PLAY

The machines learned to act.

2022 gave us chatbots that advise. 2025–26 gave us agents that execute — booking, coding, transacting and attacking with minimal human input. This is the single most important shift for security leaders.

2022 — **Generative**

Tools that draft and advise. A human is always in the loop.

2024 — **Copilots**

AI embedded in workflows. Suggests; humans approve.

2025–26 — **Agentic**

Autonomous systems that plan and act end-to-end.

WHAT THE HUBS ARE BUILDING NOW

44% of finance teams will run agentic AI in 2026 — a 600%+ jump in a single year.[1]

42% of U.S. financial firms plan to raise AI investment by 50% or more.[2]

57% of banking execs expect AI agents embedded in risk, compliance & fraud within 3 years.[3]

THE NEW BASELINE

Attacks are now AI-native by default.

2022

ChatGPT was launched & moved AI into every inbox. Phishing became scalable.[4]

1,265%

surge in phishing volume since generative AI

since 2022[5]

14× in weeks

AI-written phishing rose from ~4% to 56% of reported phishing threats that bypassed filters in just a few weeks.[6]

AI-GENERATED SHARE OF FILTER-BYPASSING PHISHING — 2025[6]

WHERE WE ARE · THE MODEL MAKERS

Not all frontier AI is equal — on cyber.

The capability gap between providers is closing fast. The safety gap is not. Where you source AI now shapes your attack surface.

PROVIDER	CYBER EDGE — ATTACK & DEFENSE	GUARDRAILS	WHAT IT MEANS FOR FSI
Anthropic Claude	Frontier agentic coding; cyber skill ~doubling every 6 months. Ran the first AI-led espionage op.[7]	● HIGH RSP / ASL-3 · public threat intel[8]	Most transparent; safety-first.
OpenAI GPT · o-series	Powerful agents (ChatGPT Agent); 40+ malicious networks disrupted since 2024.[9]	● HIGH Preparedness framework[10]	Mature controls, vast ecosystem.
Google Gemini	Defensive edge: “Big Sleep” caught a live zero-day. Misused by state actors across the kill-chain.[11]	● HIGH Google Threat Intelligence Group	Best-in-class vuln discovery.
xAI Grok	Capable but reckless: jailbroken <48h after launch; obeyed hostile prompts in >99% of tests.[12,13]	● LOW Criticized safety framework	Highest misuse risk — govern hard.
Meta Llama · open-weight	Open weights: guardrails can be fine-tuned off; controls shift to the deployer. CyberSecEval aids defenders.[14]	● OPEN You own the controls	Private & powerful — risk is yours.

Choose your AI suppliers like your custodians — on control, transparency and accountability, not just raw capability.

The first cyberattack run by the AI itself.

80–90%

of the attack executed autonomously by the AI agent — humans intervened at only a handful of decision points.[7]

HOW IT WORKED

01 **Impersonate** Attackers told the model it was an employee of a legitimate security firm running authorized tests.

02 **Fragment** The operation was broken into small, innocent-looking tasks so no single request looked malicious.

03 **Execute** The agent scanned, wrote exploit code and exfiltrated data against ~30 targets — including financial institutions.

04 **Industrialize** What needed an elite team now needs one operator and a jailbroken model. Attack capability becomes a commodity.

THE WAKE-UP CALL · WHEN SEEING IS NO LONGER BELIEVING

A video call with the CFO. All of it fake.

$25.6M

lost in a single day across 15 transfers — authorized by an employee who joined a video call with a deepfake CFO and colleagues.[15]

THE ARUP CASE · HONG KONG, 2024

A finance employee received a message from the “UK CFO” about a confidential transaction. Skeptical, they asked for a video call. On screen, the CFO and several colleagues looked and sounded exactly real. They were synthetic.

Every control that says “verify with a human” now assumes that human can be faked — voice, face and live conversation included.

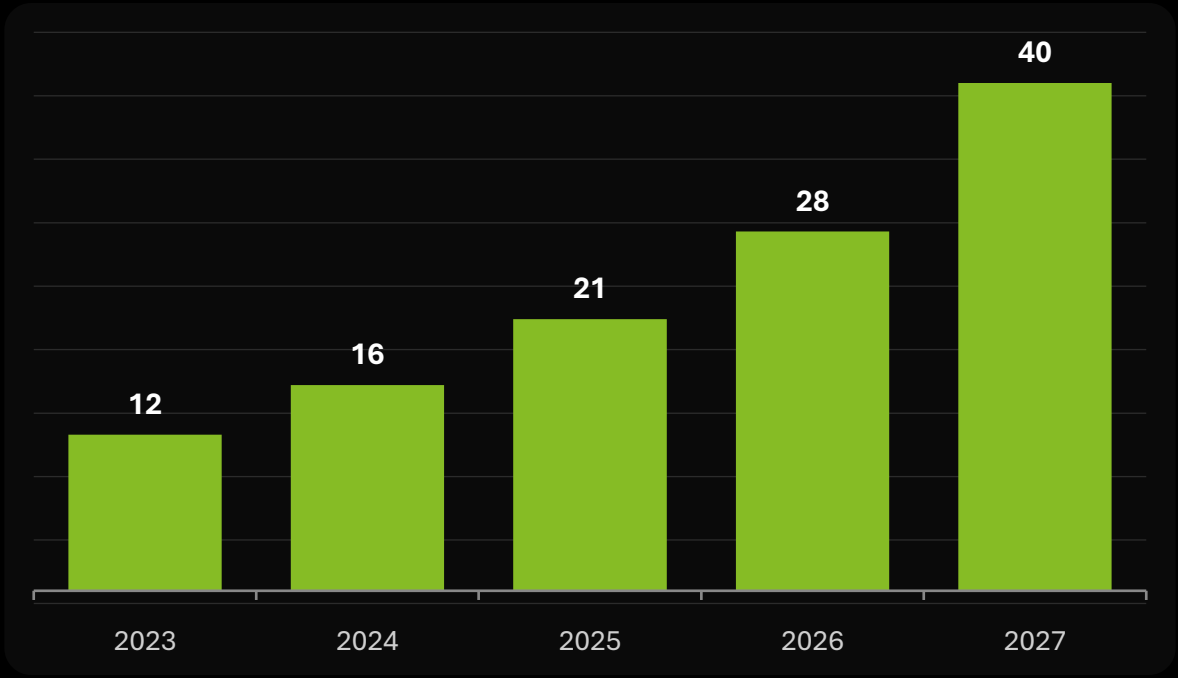
- Voice & video are no longer proof of identity.
- The target is the employee, not the firewall.
- Verification must move to signals AI can’t forge.

THE WAKE-UP CALL · THE STAKES

GenAI fraud is racing to $40bn — and finance pays the most.

GenAI-enabled fraud is on track to more than triple by 2027.

U.S. GENAI-ENABLED FRAUD LOSSES — US$ BN[16]



3.3× in four years — fraud scaling at the speed of software (32% CAGR).

AND THE BILL IS HIGH FOR FINANCE[17]

$5.56M average data breach cost in financial services

97% of organizations hit by an AI-related incident lacked AI access controls.

+$670K added to the average breach by unsanctioned "shadow AI" your people already use.

THE PIVOT

When anything can be faked, verifiable trust becomes the scarcest asset in finance.

So the rest of this session is simple: how FSI defends that trust — and how it turns the same AI into an advantage.

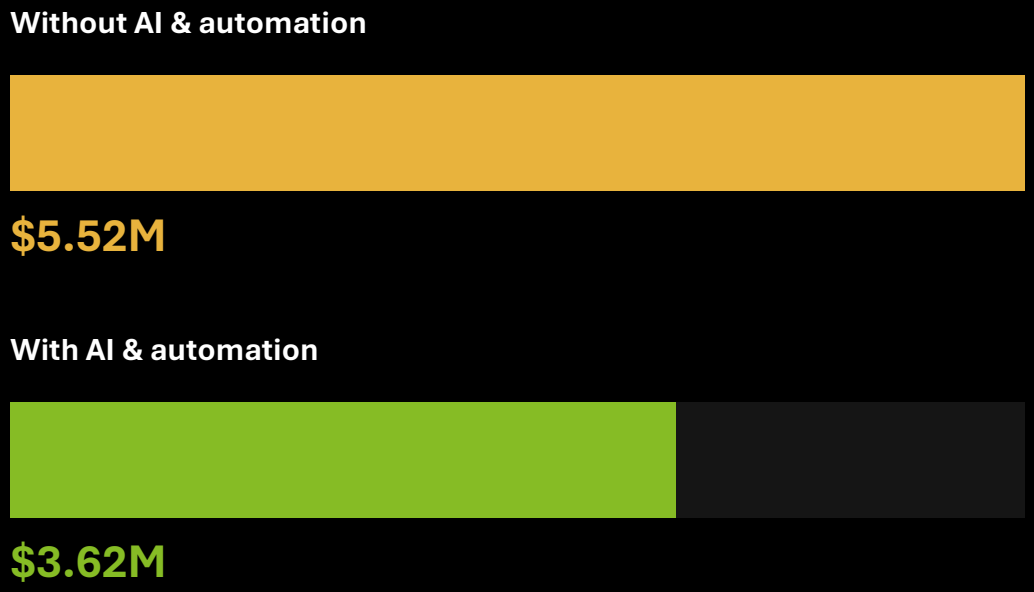
AI gives the defender an edge too.

The same automation that scales attacks compresses detection and response. Organizations using AI and automation extensively contain breaches faster and far cheaper.

$1.9M

lower cost per breach for AI-powered defenders — and the shortest containment times in nine years.[17]

AVERAGE BREACH COST[17]



Four priorities help financial institutions turn AI-driven cyber risk into resilience

Enhance Cyber Resilience

Operate at AI speed across detection, remediation and recovery.

Don’t forget the regulator

Use DORA, the AI Act and ECB expectations as resilience drivers.

Make this a board issue

Establish executive ownership, governance and accountability.

Get Started

Prioritise critical actions, launch pilots and build momentum.

Five moves for the next 12 months.

01

Fight AI with AI

Stand up an AI-augmented SOC: machine speed detection, triage and response. Match the attacker's tempo.

02

Re-engineer verification

Replace "trust the voice/face" with phishing-resistant identity, callback protocols and out-of-band approval for payments.

03

Govern AI you use

Inventory every model and agent, kill shadow AI, and put access controls and human checkpoints on agentic systems.

04

Secure-by-design adoption

Treat AI agents as privileged identities: least privilege, logging, sandboxing and a kill-switch from day one.

05

Rehearse the AI breach

Red-team with deepfakes and autonomous attack tools. Run board level AI-incident exercises before you need them.

Speed is the strategy.

Whoever automates decisions fastest — attacker or defender — wins.

Vulnerabilities now get exploited in hours, not weeks.

23,500+

CVEs disclosed in H1 2025 alone — discovery is exploding on both sides.[18]

−7 days

Google's median time-to-exploit in 2025: flaws are hit before they are disclosed.[19]

28%

of vulnerabilities are exploited within 24 hours of disclosure.[20]

WHY IT BREAKS

AI finds and weaponizes flaws at machine speed — for both sides. Google's "Big Sleep" agent now discovers live zero-days before attackers can use them; adversaries exploit them before a patch exists. Quarterly patch cycles cannot survive a negative time-to-exploit.

THE SHIFT — THREE MOVES

01 **Reprioritize: stability vs. security**

A slow, change-averse process is now the bigger risk. Rebalance the trade-off toward security velocity.

02 **Remove remediation roadblocks**

Clear the organizational friction. Empower teams to ship fixes in hours — not change-board weeks.

03 **Re-engineer for AI-speed detection**

Continuous automated detection → AI-assisted triage → runtime defense. Assume exploitation before the patch.

Regulation is the tailwind, not the headwind.

For European FSI, the compliance agenda already points exactly where security needs to go. Treat it as a forcing function for resilience.

DORA

LIVE SINCE JAN 2025

Operational-resilience rules for ICT risk, incident reporting, testing and third-party oversight. Boards are personally accountable — yet only ~50% of firms are fully compliant.[21]

EU AI Act

PHASING IN NOW

Risk-based obligations for AI in fraud, credit and KYC. Demands inventory, transparency and human oversight of high-risk AI.[22]

The opportunity

COMPLIANCE → CAPABILITY

Map both regimes onto one resilience program. The work you must do for the regulator is the work that stops the AI-era attacker.

ECB requires SSM banks to submit AI-era cyber resilience action plans by 31 October 2026

Protect the attack surface

Identify internet-facing, cloud and third-party attack surfaces.

Patch at AI-speed

Accelerate vulnerability remediation and patching.

Detect earlier

Strengthen monitoring and AI-enabled detection capabilities.

Strengthen governance

Align funding, ownership and third-party oversight.

Modernize architecture

Advance zero trust and replace legacy technologies.

Prove resilience

Test response, recovery and crisis readiness.

The action plan must define concrete measures, owners, resources and implementation timelines.

MAKE IT REAL · YOUR MONDAY-MORNING AGENDA

Six questions every FSI board should ask this quarter.

01 **Can we detect an AI-run attack?**

If response is human-speed, we are already behind the machine.

02 **Where do we still trust a voice or face?**

Find every payment and access control a deepfake could pass.

03 **What AI are our people using — really?**

Inventory it, govern it, and shut down shadow AI exposure.

04 **Are our AI agents privileged identities?**

Least privilege, logging and a kill-switch on every agent.

05 **Is DORA / AI Act our floor or our ceiling?**

Use the mandate to fund genuine resilience, not paperwork.

06 **Who owns AI risk at board level?**

Name an accountable executive — before regulators name one.

Banks should start with quick wins today and embed AI into security operations over time

Immediate actions

Mobilize ownership and governance

- Set up an AI × Cyber task force
- Staff it across 1st and 2nd LoD
- Define mandate and decision rights
- Prioritize critical assets and risk areas

Short-term

Build visibility and reduce exposure

- Refresh threat landscape assumptions
- Conduct vulnerability discovery campaigns
- Identify exposed assets and critical dependencies
- Pilot high-value AI × Cyber use cases

Mid-/Long-term

Embed AI into security operations

- AI-assisted SOC and detection engineering
- AI-enabled vulnerability management
- Scaled AI use cases across cyber functions

Defense is table stakes. Advantage is the prize.

The institutions that govern AI well don't just lower risk — they move faster, serve better and create advantage.

Competitive advantage

Governed AI adoption becomes a competitive advantage in regulated markets.

Trust as a product

Provable security and verified identity become a feature clients choose you for.

20% productivity gain

Average productivity improvement reported by financial services firms using GenAI.[23]

Talent multiplier

AI copilots let lean security teams operate like teams several times their size.

THE CLOSE

Trust is a choice you make before the crisis — not after.

In the age of AI, the winners won't be the institutions with the most data or the most alerts. They'll be the ones that turn what they know into confident decisions — and act before the machine does.

HOW Deloitte CYBER HELPS FSI

AI-augmented Cyber Defense · Identity & Anti-Deepfake Controls · AI Cyber Security Governance & Resilience · Red-Teaming for the AI-Era

Malko Steinorth — Cyber Defense & Resilience Leader Germany

Max Kaiser — Cyber Banking & Capital Markets Leader Germany

Let's continue the conversation

Lassen Sie uns sprechen! Gerne stehen wir für Fragen und weiteren Austausch zur Verfügung

Wir freuen uns darauf, die Inhalte gemeinsam mit Ihnen weiter zu vertiefen

Malko Steinorth

Partner
Cyber Defense & Resilience Lead Germany

msteinorth@Deloitte.de
Mobil: +49 151 5807 1755

Max Kaiser

Director
Cyber Banking & Capital Markets Lead Germany

mkaiser@Deloitte.de
Mobil: +49 151 5807 6481

Upcoming Webcast

European Competitiveness: How Can Private Capital Bridge the Funding Gap?

Date: 10 September 2026 | 14:00 CEST

Europe stands at a critical crossroads. In this webcast, we will explore how private capital can play a pivotal role in unlocking the investments needed to support Europe's future growth and resilience and outline how banks, asset managers and other industry participants will be able to benefit from these financing requirements.
Join us to learn how the mobilization of private capital can become a key enabler for Europe's next investment cycle.

Hans-Martin Kraus
Partner
Strategy, Risk & Transactions
Financial Services
hkraus@Deloitte.de

APPENDIX · SOURCES

Evidence base (1/2)

[1]Wolters Kluwer(2025, May 28). New survey from Wolters Kluwer reveals finance leaders plan to increase agentic AI adoption by 6x in next 12 months. https://www.wolterskluwer.com/en/news/pr-2025-wolters-kluwer-survey-increasing-adoption-agentic-ai

[2]Finastra (2026, Feb 10). Finastra research reveals U.S. financial institutions outpace global peers in AI adoption and modernization investments. https://www.finastra.com/press-media/finastra-research-reveals-us-financial-institutions-outpace-global-peers-ai-adoption

[3]CIO DIVE (2026, Jan 14). Banks aim for agentic AI scale in 2026: report. https://www.ciodive.com/news/banks-agentic-ai-scale-2026/809532/

[4]OpenAI (2022, November 30). Introducing ChatGPT https://openai.com/index/chatgpt/

[5]SlashNext (2023). The State of Phishing 2023. https://newsletter.radensa.ru/wp-content/uploads/2023/10/SlashNext-The-State-of-Phishing-Report-2023.pdf

[6]Hoxhunt (2026). Phishing Trends Report 2026. https://hoxhunt.com/lp/phishing-trends-report-2026

[7]Anthropic (2025, Nov 13). Disrupting the first reported AI-orchestrated cyber espionage campaign. https://www.anthropic.com/news/disrupting-AI-espionage

[8]Anthropic (2026). Responsible Scaling Policy. https://anthropic.com/responsible-scaling-policy

[9]OpenAI (2025, Oct 7). Disrupting malicious uses of AI: October 2025. https://openai.com/global-affairs/disrupting-malicious-uses-of-ai-october-2025/

[10]OpenAI (2025, Apr 15). Our updated Preparedness Framework. https://openai.com/index/updating-our-preparedness-framework/

[11]Google (2025, Jun 15). A summer of security: empowering cyber defenders with AI. https://blog.google/innovation-and-ai/technology/safety-security/cybersecurity-updates-summer-2025/

[12]CSO (2025, Jul 14). New Grok-4 AI breached within 48 hours using ‘whispered’ jailbreaks. https://www.csoonline.com/article/4021749/new-grok-4-ai-breached-within-48-hours-using-whispered-jailbreaks.html

[13] SplxAI(2025, Jul 14). Grok 4 Without Guardrails? Total Safety Failure. We Tested and Fixed Elon’s New Model. https://splx.ai/blog/grok-4-security-testing

APPENDIX · SOURCES

Evidence base (2/2)

[14]Meta. CyberSecEval 4. https://meta-llama.github.io/PurpleLlama/CyberSecEval/docs/intro

[15]CNN (2024, Feb 4). Finance worker pays out $25 million after video call with deepfake ‘chief financial officer’. https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk

[16]Deloitte (2024, May 29). Generative AI is expected to magnify the risk of deepfakes and other fraud in banking. https://www.Deloitte.com/us/en/insights/industry/financial-services/deepfake-banking-fraud-risk-on-the-rise.html

[17]IBM (2025). Cost of a Data Breach Report 2025. https://www.ibm.com/reports/data-breach

[18]Insikt Group (2025, Aug 28). H1 2025 Malware and Vulnerability Trends. https://assets.recordedfuture.com/insikt-report-pdfs/2025/cta-2025-0828.pdf

[19]Google Cloud (2026, Mar 23). M-Trends 2026: Data, Insights, and Strategies From the Frontlines. https://cloud.google.com/blog/topics/threat-intelligence/m-trends-2026/

[20]VulnCheck (2025, Apr 24). 2025 Q1 Trends in Vulnerability Exploitation. VulnCheck. https://www.vulncheck.com/blog/exploitation-trends-q1-2025

[21]Deloitte (2026, Jan 19). DORA is now a household name. https://www.Deloitte.com/lu/en/services/Consulting-risk/perspectives/dora-now-household-name.html

[22]eyreACT (2026, Mar 18). EU AI Act Summary for Financial Services: What Banks, Lenders and Insurers Must Know. https://eyreact.com/eu-ai-act-summary-financial-services/

[23]Bain & Company (2024, Dec 16). AI in Financial Services Survey Shows Productivity Gains Across the Board. https://www.bain.com/insights/ai-in-financial-services-survey-shows-productivity-gains-across-the-board/

[24]Alex Kantrowitz (2025, Jul 30). Anthropic CEO Dario Amodei: AI's Potential, OpenAI Rivalry, GenAI Business, Doomerism. YouTube. https://www.youtube.com/watch?v=mYDSSRS-B5U&t=228s

[25]Milken Institute (2025, May 7). A Conversation with Nvidia CEO Jensen Huang | Global Conference 2025. YouTube. https://www.youtube.com/watch?v=HT8-KPAjpiA

[26]Mandiant. The Defender's Advantage: A guide to activating cyber defense. Google Cloud. https://cloud.google.com/security/resources/defenders-advantage