Webinar

Umgang mit Cyberrisiken im Kontext leistungsstarker KI-Modelle

KEY FACTS

aus dem Webinar „Claude Mythos & Co.: Neue Cyberrisiken für Finanzunternehmen verstehen und bewältigen“ des Bank-Verlags in Kooperation mit Bankenverband und Bankenakademie vom 24.06.2026.

Das Webinar führte unterschiedliche Perspektiven zusammen: die Sicht der Aufsicht, Einschätzungen der nationalen Cybersicherheitsbehörde sowie Praxisbeiträge aus mehreren Kreditinstituten.

Einordnung der Bundesanstalt für Finanzdienstleistungsaufsicht

Kernaussage: Cyberrisiken stehen im Fokus der Aufsicht, da die Bedrohungslage hoch ist und Cybervorfälle nicht nur einzelne Institute, sondern sehr schnell auch das Finanzsystem und die gesamte Industrie betreffen können. Digitale Innovationen werden als wichtig für ein funktionsfähiges, wettbewerbsfähiges und stabiles Finanzsystem eingeordnet; zugleich steigen mit der zunehmenden Nutzung digitaler Technologien die Risiken im Cyberraum.

Dies gilt auch für Chancen und Risiken durch neue Frontier-AI-Modelle. Aus Sicht der Aufsicht bilden die vollständige Umsetzung von DORA und eine gute Cyberhygiene eine solide Grundlage für die aktuellen Herausforderungen. Dennoch sind weitere Schritte notwendig. Unternehmen müssen sich darauf einstellen, eine Vielzahl neuer Schwachstellen in kurzer Zeit bewerten und schließen zu müssen oder durch geeignete Mitigationsmaßnahmen abzusichern. Ein risikoorientiertes Vorgehen ist hierbei entscheidend. Kritische Systeme sollten im Fokus stehen. Auch sind Lösungen für veraltete „unpatchbare“ Systeme notwendig. Dabei ist zu bedenken, dass die Herausforderungen nicht nur das eigene Unternehmen, sondern auch IKT-Drittdienstleister betreffen.

Zentrale Treiber: Komplexe IT-Systeme, geopolitische Konflikte sowie Konzentrationen und Abhängigkeiten prägen die hohe Risikolage. Durch den steigenden Einsatz von KI entstehen neue Risiken. Aktuell stehen insbesondere leistungsstarke Frontier-AI-Modelle wie Claude Mythos 5, GPT-5.5 und MDASH im Fokus. Sie sollen in der Lage sein, bisher unbekannte Schwachstellen in IKT-Systemen in kurzer Zeit zu entdecken und auszunutzen. Da sich die Bereitstellung dieser Modelle auf wenige Anbieter konzentriert, könnten sich die Abhängigkeiten von einzelnen Anbietern in Drittstaaten durch diese Entwicklung noch einmal vergrößern. Die Rahmenbedingungen hinsichtlich Zugang, Datenschutz sowie rechtlicher Aspekte können sich aktuell täglich ändern.

Weiterführender Hinweis: Informationsaustausch, Diskussion und das Teilen von Erfahrungen und Best Practices werden als wichtig hervorgehoben. Die vollständige Keynote von Nikolas Speer ist hier zu finden: https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Reden/re_260626_claude_mythos_speer.html

KOOPERATIONSPARTNER

bankenverband bankenakademie

BV bank-verlag

Frontier AI: Beschleuniger der Risikolage

Frontier-AI-Modelle wie Mythos erhöhen das Cyberrisiko sowohl kurzfristig als auch langfristig. Sie ermöglichen eine schnellere Entdeckung und Ausnutzung von IKT-Schwachstellen, senken die Kompetenzanforderungen für komplexe Angriffe und erhöhen Geschwindigkeit, Skalierung und Raffinesse von Angriffen.

- **Geringerer Aufwand für Angreifer:** KI senkt notwendige Kompetenz und verkürzt Exploit-Zeit.
- **Mehr Geschwindigkeit und Skalierung:** Es entstehen mehr Zero-Day-Exploits, Angriffe auf Software-Lieferketten und Angriffe auf Dienstleister.
- **Mehr Sicherheitsvorfälle erwartbar:** Viele Schwachstellen und viele Patches in kurzer Zeit erhöhen den Handlungsdruck.

Vgl. hierzu auch: Auswirkungen auf die Cybersicherheit von Organisationen durch die Entwicklung im Bereich Künstlicher Intelligenz. (BSI-Dokument vom 22.06.2026)

Patchen unter Zeitdruck

Angriffe erfolgen in einem Tempo jenseits klassischer Patch-Zyklen. Schwachstellen- und Patchprozesse müssen daher deutlich beschleunigt, skalierbar gestaltet und mit aktiver Nachverfolgung verbunden werden.

- **Eigene IT-Infrastruktur und (kritische) Assets kennen:** Dies ermöglicht Priorisierung.
- **Schwachstellen- und Patch-Management stärken:** Identifikation, Patchen und Incident Response sollen automatisiert oder teilautomatisiert werden.
- **Transparenz schaffen:** Schwachstellen sollen für technische Teams und Entwickler sichtbar sein; KPIs schaffen Transparenz für Leitungsorgane.
- **Aktiv nachverfolgen:** Ein dedizierter Schwachstellenmanager kann die Nachverfolgung unterstützen.
- **Herausforderungen berücksichtigen:** Dazu zählen Aufrechterhaltung der Business Continuity, Ablösung der Legacy-Systeme, automatisiertes Testen und „Human in the Loop".

Kurzfazit: Patch- und Incident-Management müssen skalierbar gestaltet, Priorisierung ermöglicht und Nachverfolgung gestärkt werden.

Cyberhygiene als belastbare Basis

Cyberhygiene, Defense-in-Depth, Zero Trust, Security-by-Design, Angriffsflächenreduktion, Security-in-Depth und Krisenreaktion sind zentrale Ansatzpunkte.

- **Cybersecurity-Basismaßnahmen umsetzen:** Defense-in-Depth und Zero Trust sollen vorangetrieben werden.
- **Security-by-Design nutzen:** Sicherheitsanforderungen sollen bereits bei der Konzeption, Entwicklung und Integration eines Large Language Models (LLM) berücksichtigt werden.
- **Cyberhygiene verbessern:** Dazu gehören Patching, End-of-Life-Management und Penetrationstesting.
- **Angriffsfläche reduzieren:** Die Angriffsfläche soll grundlegend reduziert werden; Priorisierung der externen und aus dem Internet erreichbaren Assets.
- **Moderne KI-Techniken adaptieren:** KI-Techniken sollen für Cyber Defense und Schwachstellensuche genutzt werden.

Kurzfazit: Belastbare Cyber-Hygiene reduziert die Angriffsfläche, stärkt Basis-Kontrollen und schafft die Voraussetzung, KI-gestützten Bedrohungen wirksam zu begegnen.

Resilienz sichern: erkennen, abschotten, wiederherstellen

Der Digital Operational Resilience Act (DORA) ist seit 2025 verbindlich anzuwenden und die Grundlage für die Widerstandsfähigkeit von Finanzunternehmen. Eine sich schnell verändernde Bedrohungslage, Drittparteienabhängigkeiten und Cyberrisiken erfordern geschärfte und geübte Notfall-, BCM- und IKT-Incident-Response-Strukturen.

- **Anomaly Detection:** Technische Schutzmaßnahmen umfassen Anomaly Detection.
- **Containment stärken:** Containment Capabilities und Abschottung sind zentrale Elemente.
- **Response Drills verbessern:** Verbesserte Response Drills und Cyber Incident Response Playbooks stärken die Reaktionsfähigkeit.
- **Restore Capabilities stärken:** Restore Capabilities sowie Recovery Plans sind Teil der Resilienzbetrachtung.
- **Übungen ausrichten:** Szenarien für Krisenübungen sollen auf Frontier AI ausgerichtet werden.

Kurzfazit: Resilienz bleibt Basis; Containment, Restore Capabilities und Response Drills gewinnen an Bedeutung.

Drittparteien und Lieferketten im Blick

Risiken entstehen durch Angriffe auf Drittparteien, Angriffe auf Dienstleister, Software-Lieferketten und kompromittierte Drittanbieter-Bibliotheken. Auch fremde Software und Open Source ist ein relevantes Risikofeld.

- **Drittparteien:** Ein erhöhtes Sicherheitsniveau bei Drittparteien soll gemessen und eingefordert werden.
- **Lieferanten und Hersteller:** Bei Beauftragung oder Einkauf von Software sollen Hersteller und Lieferanten in die Pflicht genommen werden.
- **Drittanbieter-Bibliotheken:** Eine zunehmende Anzahl kompromittierter Drittanbieter-Bibliotheken erhöht die Risiken.
- **Open Source und Versionen:** Spezifische Versionen sollen festgelegt werden, nicht „latest".
- **Eigenentwicklung:** KI-basierte Schwachstellensuche und -behebung soll in eigene Prozesse integriert werden.

Kurzfazit: Drittparteien, Dienstleister, Lieferketten, fremde Software und Drittanbieter-Bibliotheken sind Teil der aktuellen Risikolage.

Legacy und End-of-Life: besondere Risikofelder

Legacy-Systeme sind eine Herausforderung im Schwachstellen- und Patch-Management. End-of-Life-Management ist Bestandteil verbesserter Cyberhygiene.

- **Legacy-Systeme berücksichtigen:** Legacy-Systeme erschweren Behebung, Patchen und Automatisierung.
- **End-of-Life-Management einbeziehen:** End-of-Life-Management ist Teil verbesserter Cyberhygiene.
- **Priorisierung ermöglichen:** Gute Kenntnis der IT-Infrastruktur und Assets ermöglicht Priorisierung.
- **Cyberhygiene stärken:** Patching, End-of-Life-Management und Penetrationstesting sind zentrale Maßnahmen.

Umsetzung: groß denken, fokussiert handeln

Die Umsetzung unterscheidet sich je nach Größe und Organisation des Instituts; die zentralen Schwerpunkte bleiben jedoch vergleichbar. Beispielsweise:

- **Große Bank:** Steering Board, Task Force, Verbesserung von Priorisierung und Risikoreduzierung, Stärkung von Schwachstellen- und Patch-Management, LLM-basiertes Security-by-Design und Schwachstellen-Scanning, beschleunigte Cyberhygiene, Angriffsflächenreduktion, Security-in-Depth, aktualisierte und getestete Krisenreaktion und erhöhtes Sicherheitsniveau bei Drittparteien prägen die Herangehensweise.
- **Kleinere Bank:** Transparenz der Schwachstellen für technische Teams, aktive Nachverfolgung durch einen dedizierten Schwachstellenmanager, Transparenz via KPIs für Leitungsorgane, Automatisierung von Behebung, Patchen und Incident Response, Anpassung von Cyber-Response-Playbooks und Krisenübungen sowie KI-basiertes Schwachstellen-Scanning und -Behebung bei Eigenentwicklung stehen im Vordergrund.
- **Gemeinsam:** Schwachstellen- und Patch-Management, Automatisierung, Cyber-Response, Übungen und der Umgang mit Drittparteien beziehungsweise Dienstleistern sind zentrale Schwerpunkte.

Leitlinie: Skalierbare Patch- und Incident-Management-Prozesse, gute Kenntnis der IT-Infrastruktur und Assets, Cybersecurity-Basismaßnahmen, KI-Techniken für Cyberdefense sowie geschärfte und geübte Resilienzstrukturen stehen im Mittelpunkt.

Strategisches Vorgehen: Zielbild 2027+ und gesteuerte Umsetzung

Frontier-AI-Risiken sollten bis 2027+ dauerhaft in die reguläre IKT-Risikosteuerung integriert werden. Gemäß ECB-Anforderung umfasst dies auch die Überprüfung von Risk Appetite, Toleranzschwellen, Kontrollmaßnahmen, Ressourcenentscheidungen und Management-Body-Oversight. Ziel ist ein skalierbares Kontrollumfeld, das operative Maßnahmen konsistent bündelt, messbar macht und nachhaltig verankert. In einem Schreiben an die Geschäftsleitungen der unter den „Single Supervisory Mechanism (SSM)" fallenden Banken hat die Europäische Zentralbank am 7. Juli 2026 die von ihr erwartete Vorgehensweise konkretisiert.

Strategisch bündeln: Bestehende Maßnahmen sollen in einem gemeinsamen Zielbild 2027+ zusammengeführt werden, um klare Umsetzungswellen, Verantwortlichkeiten, Erfolgskriterien sowie risikobasierte Entscheidungswege zu definieren:

- **Skalieren:** Piloten und Einzelmaßnahmen risikobasiert in wiederholbare Standardprozesse überführen.
- **Risikosteuerung schärfen:** Risk Tolerance Framework, Toleranzschwellen und Kontrollmaßnahmen an die erhöhte Geschwindigkeit und Skalierung KI-gestützter Bedrohungen anpassen.
- **Steuern:** Fortschritt über klare KPIs, Eskalationswege, Ressourcenentscheidungen und regelmäßige Management-Updates nachhalten.

Kurzfazit: Frontier-AI-Risiken erfordern eine geschärfte IKT-Risikosteuerung mit angepasstem Risk Appetite, klaren Toleranzschwellen und messbarer Management-Steuerung.

KOOPERATIONSPARTNER

bankenverband bankenakademie

BVbank-verlag