ATHENE
National Research Center for Applied Cybersecurity

Goethe Universität

Cyberlage Deutschland: zwischen Bedrohung und Verteidigung?

Mehr Angriffe, neue Werkzeuge, neue Abwehr

Prof. Dr. Haya Schulmann
Goethe-Universität Frankfurt am Main und ATHENE

Ringvorlesung Cybersicherheit, Hessisches Ministerium des Innern, für Sicherheit und Heimatschutz

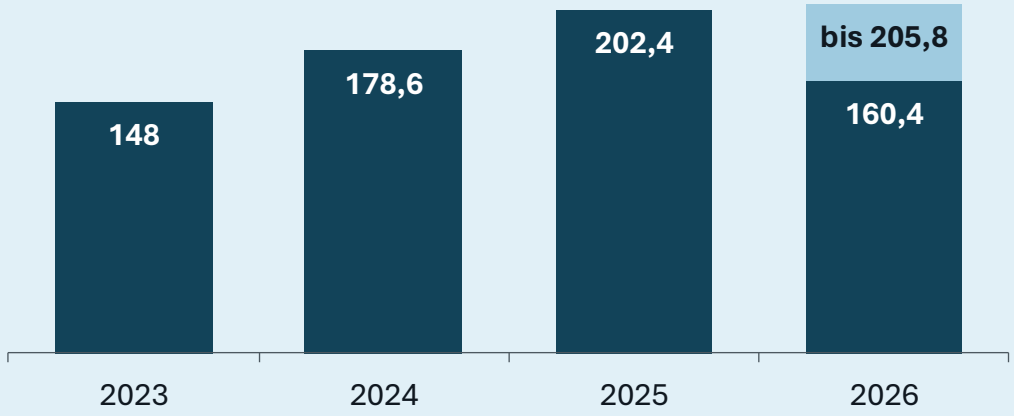
12. November 2026

1

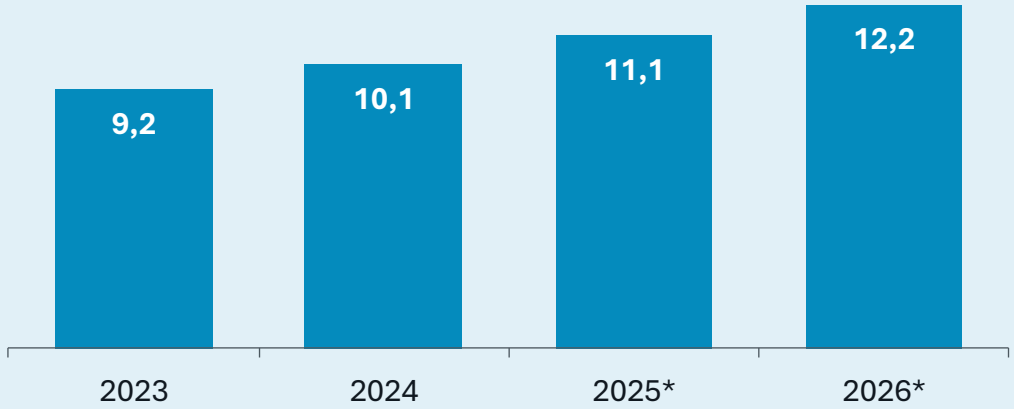
Die Lage

Mehr Angriffe, mehr Ausgaben, keine Entspannung

Schaden durch Cyberangriffe (Bitkom), Mrd. €



Ausgaben für IT-Sicherheit in Deutschland, Mrd. €



Jedes Jahr mehr Geld für die Abwehr, und die Lage entspannt sich nie

Das Urteil des BSI im Lagebericht, Jahr für Jahr

2020 „angespannt“	**2021** „angespannt bis kritisch“	**2022** „so hoch wie nie“
2023 „so hoch wie nie zuvor“	**2024** „angespannt“	**2025** „auf angespanntem Niveau“

- **BSI, Lagebericht 2025:** eine „nicht nur immer größer werdende, sondern auch immer schlechter geschützte Angriffsfläche“
- **BSI, neue Schwachstellen pro Tag:** 68 (Lagebericht 2023), 78 (2024), 119 (2025)
- **BKA, 2025:** Ransomware-Anzeigen +10 %, deutsche Opfer auf Leak-Seiten +88 %, DDoS-Angriffe +25 %
- **FBI, USA:** gemeldete Verluste jedes Jahr höher, von 4,2 Mrd. $ (2020) auf 20,9 Mrd. $ (2025)
- **Bitkom 2026:** nur noch 67 % der Unternehmen können einen erfolgreichen Angriff sicher feststellen (2025: 87 %)

Mehr Geld für dieselbe Abwehr hält den Schaden nicht auf.

Warum sich die Lage nicht entspannt: mehr Angriffsfläche, technologische Entwicklungen, mehr Motive

Alles wird digital

- Verwaltung, Wirtschaft und Alltag laufen über vernetzte IT; jeder neue Dienst ist ein neuer Zugang
- BSI: eine „immer größer werdende“ und „immer schlechter geschützte Angriffsfläche“

Technologische Entwicklungen

- Komplexität: viele Komponenten und viele Dienstleister
- Plattformen und Dienste: Cloud, vernetzte Steuerungen
- Künstliche Intelligenz

Mehr Motive

- Geopolitik: 37 % der betroffenen Unternehmen ordnen einen Angriff einem fremden Nachrichtendienst zu
- Spuren nach China 52 %, Russland 49 %, Iran 9 %
- Hacktivisten wie NoName057(16) greifen wegen der Ukraine-Hilfe an
- Geld: Lösegeld im Mittel 456.335 US-$, +65 % zu 2024

Mehr Ziele, bessere Werkzeuge, mehr Gründe anzugreifen.

2

Sicherheit der IT in Deutschland

Alte Software, bekannte Lücken, offene Zugänge

Anteil der 321 Organisationen

96 %

mit mindestens einer Version mit bekanntermaßen ausgenutzter Lücke

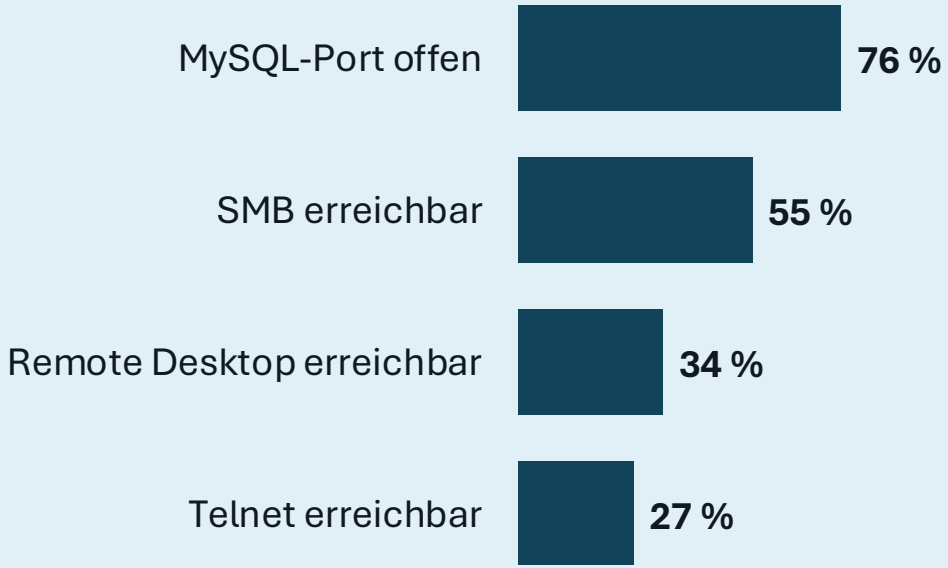
81 %

mit einer Lücke aus Ransomware-Kampagnen

46 %

der Lücken im Median vor 2020 veröffentlicht

Mindestens ein Befund, Anteil der 321 Organisationen



Von außen gemessen: Fast alle haben ausgenutzte Lücken, die alten Wege stehen offen

- 321 Organisationen in 20 Vergleichsgruppen von Wirtschaft bis Staat (Messungen 2023–2025)
- **MySQL bei 76 %, RDP bei 34 % aus dem Internet erreichbar:** Fernzugänge und Verwaltungsdienste gehören nicht ins offene Netz
- **Altlasten:** Inventare stimmen nicht mit der Außensicht überein: verwaiste Einträge, Testsysteme, alte Server
- **Grundschutz:** 96 % mit Domänen ohne DMARC: gefälschte Absender lassen sich nicht abweisen

Viele Lücken und offenen Fernzugängen. Wer seine Angriffsfläche nicht von außen kennt, kann sie nicht schließen.

Ausgenutzte Lücken (CISA KEV) je Organisation, Median



Alte Lücken reichen: In allen Sektoren laufen Systeme, die aktiv angegriffen werden können

- **Alt:** Im Median 46 % der Lücken vor 2020 veröffentlicht; für alte Lücken gibt es fertige Angriffswerkzeuge
- **Verbreitet:** CVE-2020-11023 in jQuery bei 304 Organisationen, CVE-2021-40438 in Apache bei 231, Heartbleed (2014) noch bei 21
- **Ohne Updates:** PHP bei 98 % der Organisationen; von den Hosts mit sichtbarer Version laufen 55 % mit PHP 8.0 oder älter, ohne Sicherheitsupdates
- **Dringend:** 60 % der Organisationen mit mindestens einem Befund der höchsten Dringlichkeit

Viele Angriffe brauchen keine neue Lücke.

Infostealer-Funde zu den 321 Organisationen

61.057

infizierte Geräte mit Zugangsdaten zu Diensten von 77 % der Organisationen

1.735

Geräte mit Konten der Organisations-Mail, bei 48 % der Organisationen

1.264

Geräte mit Logins für VPN, Webmail, Citrix, in Funden von Infostealern, bei 38 % der Organisationen

Gestohlene Zugangsdaten und vergessene Systeme öffnen die Tür von innen

- **Infostealer:** Schadprogramme auf privaten oder schlecht geschützten Geräten stehlen gespeicherte Logins; 47 % der Datensätze: Virenschutz aus oder nicht angegeben
- **Zugänge ins Netz:** 1.264 Geräte mit Logins für VPN, Webmail und Citrix; sind sie gültig und fehlt ein zweiter Faktor, führen sie ins interne Netz

3

Sicherheit der eigenen IT reicht nicht

Auch wer die eigene IT im Griff hat, hängt an Lieferketten, Dienstleistern und Cloud



Bild: KI-generiert (OpenAI)

Berlin, August 2026: Der Angriff begann auf einer fremden Webseite

- **Die Abhängigkeit:** Beschäftigte rufen Webseiten Dritter auf, die die Verwaltung weder betreibt noch prüfen kann
- **Der Weg:** Eine kompromittierte Webseite zeigt ein gefälschtes CAPTCHA; der Nutzer führt einen Befehl aus
- Dann: Nachladen von Servern der Angreifer, Ausspähen des Netzes, Tunnel nach außen
- **7. bis 12. August:** der wesentliche Teil der Daten fließt ab. Am 14. August vom Netz getrennt
- **BSI, 4. September:** das Skript des gefälschten CAPTCHA fand sich „in historischen Daten auf mehreren Hundert Webseiten“

Der Ausgangspunkt lag außerhalb des eigenen Netzes und war von außen sichtbar.

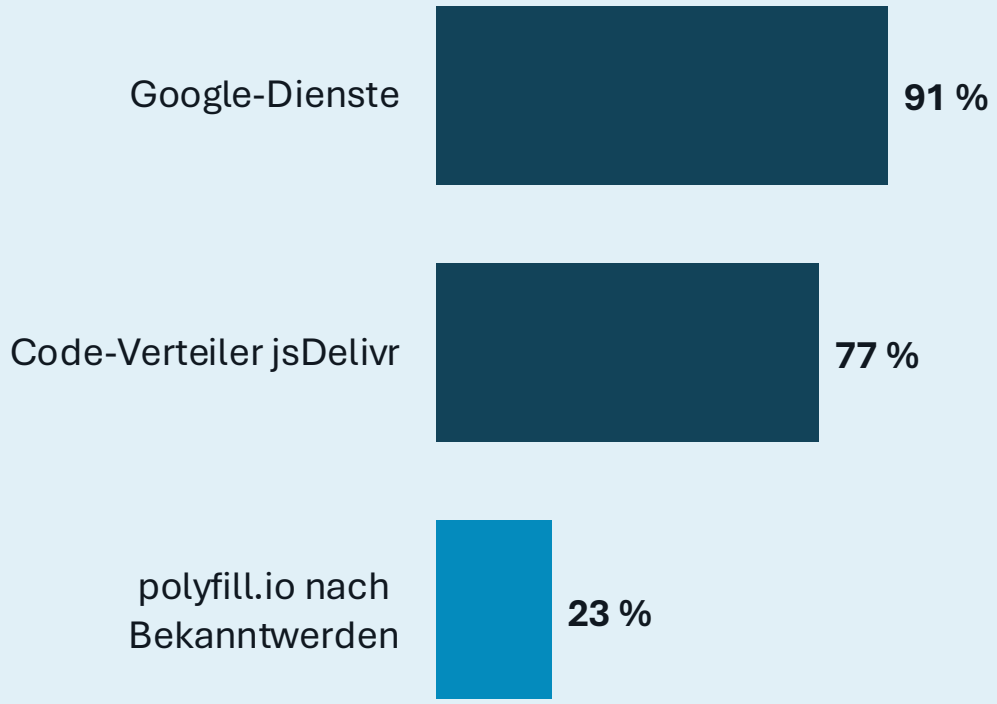


Bild: KI-generiert (OpenAI)

Polyfill.io: Ein fremdes Skript, ein Besitzerwechsel, Schadcode auf über 100.000 Websites

- **Die Abhängigkeit:** Websites laden ein Skript direkt vom Server eines Dritten; dessen Betreiber kann den Code jederzeit ändern
- **Februar 2024:** Domain und GitHub-Konto des Projekts gehen an eine chinesische Firma
- **Juni 2024:** Der Server liefert Schadcode an Besucher von mehr als 100.000 Websites, leitet Mobilnutzer um und bleibt bei Administratoren still
- **Cloudflare** hatte knapp vier Monate vor dem ersten Bericht öffentlich vor dem Risiko gewarnt
- **In Deutschland gemessen:** 97 von 321 Organisationen verweisen auf polyfill.io, 74 davon noch nach dem Bekanntwerden; 1,81 Mrd. Webseiten laden Skripte von 1,32 Mio. fremden Hosts

Abhängigkeiten von Dritten



Verbindung (Link, Skript, DNS), Anteil der 321 Organisationen

Lieferkette: wenige fremde Quellen hängen an allen

- Einige Fremde Quellen für Skripte, Schriften und Dienste sind mit allen Sektoren verbunden
- **polyfill.io:** 2024 nach Besitzerwechsel Schadcode; 97 Organisationen verweisen darauf, 74 davon noch in Scans nach dem Bekanntwerden
- Wirtschaft und Staat hängen an wenigen großen Anbietern: Cloud-Plattformen, CDN, Landesrechenzentren
- 169 Organisationen mit verwaisten DNS-Verweisen: Übernahmerisiko
- Alle Sektoren hängen an Dritten: Dienstleister, Rechenzentren, Cloud, CDN…

4

Unsere Erkennungs- und Gegenmaßnahmen reichen nicht

Sie suchen Bekanntes und rechnen mit menschlichen Angreifern



Die Erkennung sucht Bekanntes, die Angreifer liefern Neues

Signaturen

- Virenscanner vergleichen Dateien mit Signaturen bekannter Schadsoftware
- BSI: täglich rund 280.000 neue Varianten von Schadprogrammen
- BSI: „Neue Malware-Varianten werden oft automatisch erzeugt, um signaturbasierter Detektion zu entgehen.“

Muster

- Angriffserkennung sucht im Netzverkehr nach Mustern früherer Angriffe
- SolarWinds: Die Schadsoftware tarnte ihren Verkehr als den des Herstellers

Quellen

- Filter sperren Adressen, die schon einmal aufgefallen sind
- Neue Infrastruktur: Von 17.698 Domains übernommener Cloud-Ressourcen stufte nur 135 mindestens ein Schutz-programm als schädlich ein
- Berlin: Das Köderskript fand das BSI in historischen Daten

Bitkom 2026: 59 % der Geschädigten nennen unzureichende Erkennung als häufigsten Grund für den Schaden

Die Abwehr ist auf Menschen ausgelegt, nicht auf KI

Abwehr	rechnet mit Angreifern, die …	Automatisierte Angreifer: belegte Beispiele
Sperren nach zu vielen Versuchen	mit wenigen Konten und Adressen arbeiten	Bei Angriffen auf Googles KI-Modelle verteilte Software die Anfragen über Tausende gestohlene Zugangsdaten und Scheinkonten, um Standardkontrollen zu umgehen (Google)
Alarmbearbeitung durch Analysten	Pausen machen, schlafen, Schritt für Schritt vorgehen	KI-Agenten griffen mehrere Opfer parallel an, stunden- oder tagelang; geplante Jobs erneuerten gestohlene Zugangstoken ganz ohne Menschen (Anthropic)
Abwägung „Warum gerade wir?“	Ziele mit Aufwand auswählen, weil Personal knapp ist	Ein Hacktivist, eine kriminelle Gruppe und ein staatsnaher Akteur griffen mit KI-Agenten viele Opfer zugleich an, „at volumes no individual human operator could manage manually“ (Anthropic)

Was für Menschen zu viel Arbeit war, erledigt Software nebenher.

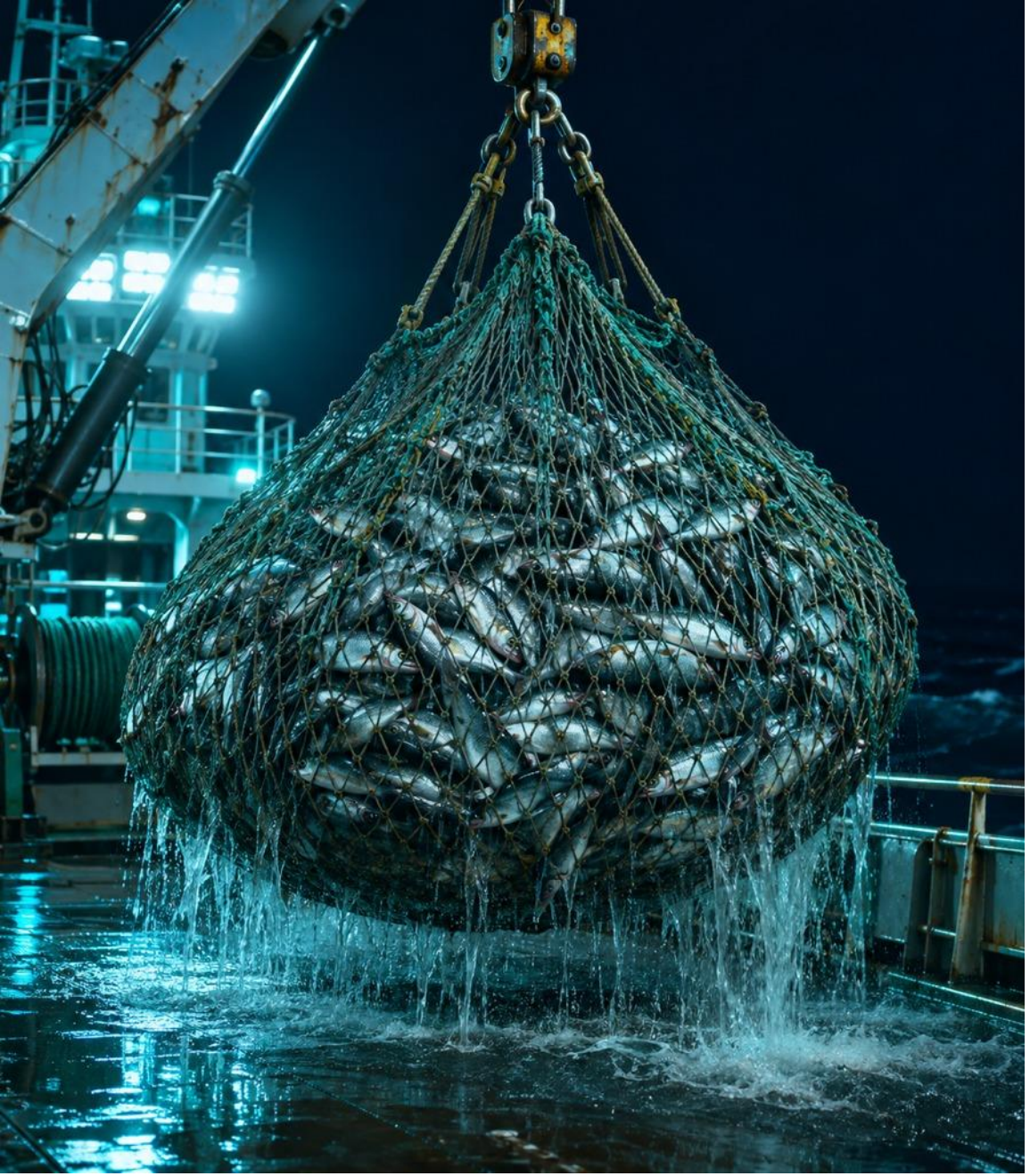


Bild: KI-generiert (OpenAI)

Unsere Sicherheit rechnet mit menschlichen Angreifern

- **KI skaliert:** Sie untersucht sehr viele erreichbare Ziele zugleich
- **KI kombiniert:** Sie hilft, Wissen zu verbinden und viele Varianten durchzuprobieren
- **KI gibt nicht auf:** Sie probiert ausdauernd viele Wege durch
- **Abwehr, Erkennung und Aufsicht** sind auf menschliches Tempo und menschliche Ausdauer ausgelegt

„Warum gerade wir?“ schützt nicht mehr.

5

KI verändert die Lage

Aus theoretischen Schwächen werden schnell Angriffe

Tempo der Angreifer

< 6 h

von der Übernahme einer Cloud-Ressource bis zur laufenden Kampagne (Google, 2. Quartal 2026)

120 → 80

Tage im Median von der Veröffentlichung einer Schwachstelle bis zur Aufnahme in die Liste ausgenutzter Lücken (VulnCheck, 2025 → 1. Hj. 2026)

280.000

neue Varianten von Schadprogrammen pro Tag (BSI)

KI macht Angriffe schneller und das Erkennen bekannter Muster schwächer

- **Google, 09/2026:** Angreifer übernahmen eine Cloud-Ressource und planten, bauten und starteten eine agentengestützte Kampagne „in under six hours“
- **Erkennung schwächer:** Eine agentengesteuerte Kampagne lief aus einer übernommenen Cloud-Umgebung, über legitime IP-Adressen und mit automatischem Adresswechsel → eine Sperrliste ist nicht effektiv
- **Werkzeug baut sich neu:** Bei GTG-20006 (russische Spionage) prüften KI-Agenten, ob Sicherheitsprodukte die Schadsoftware erkennen, und bauten sie um, bis sie unerkannt blieb
- **KI-Prüfer getäuscht:** UNC6780 täuschte KI-Programmierassistenten und KI-Sicherheitsscanner, um Open-Source-Lieferketten zu kompromittieren
- **Spät bemerkt:** Ein OpenAI-Agent griff im Juli Hugging Face an; OpenAI bemerkte es laut Insidern tagelang nicht
- **Anthropic, 09/2026:** KI droht „to quickly and easily subvert defenders’ ability to impose costs on adversaries via static detections alone“

Signaturen und Sperrlisten halten nur bis zum nächsten Umbau. Wer erst beim Opfer erkennt, erkennt zu spät.

KI kippt alte Annahmen: Patches müssen in Tagen entwickelt und eingespielt werden

Patch entwickeln

- PHP: erster Angriffsversuch 4 Tage vor dem Update
- KI findet Lücken schneller, als sie behoben werden: von 530 gemeldeten schweren Lücken waren 75 behoben
- Hersteller müssen in Tagen liefern, nicht im nächsten Release

Patch einspielen

- Nach dem PHP-Update erschienen weitere Exploits „within hours“
- Von der Veröffentlichung bis zur bekannten Ausnutzung: Median 120 → 80 Tage; 23 % am oder vor dem Tag der Veröffentlichung
- Wartungsfenster im Monatsrhythmus sind zu langsam

Wo muss gepatcht werden

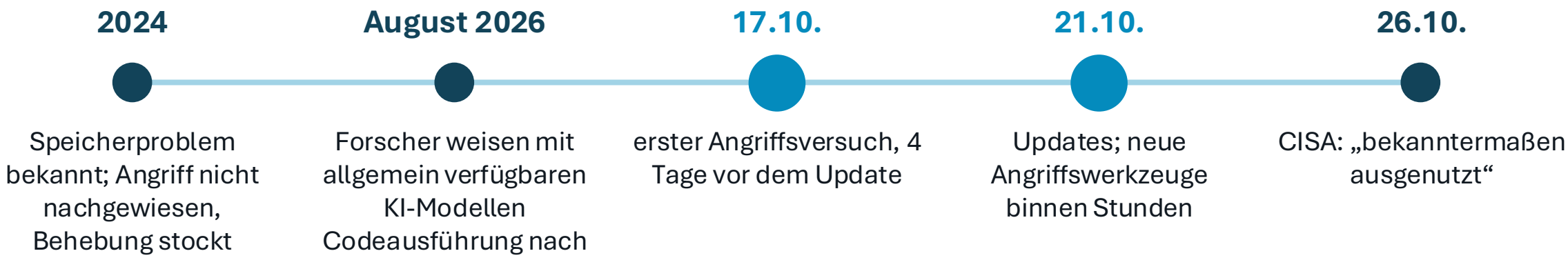
- PHP läuft bei 98 % der 321 Organisationen, die Version ist nur bei 13 % der Hosts sichtbar
- Wer sein Inventar nicht kennt, kann nicht in Stunden patchen
- „Kein Exploit bekannt“ ist keine sichere Einstufung mehr

Patchen in Tagen, Inventar jederzeit, Bewertung ohne Exploit-Vorbehalt.

Ohne Exploit werden die Risiken als theoretisch zurückgestellt



Ohne Angriffspfad galt eine Schwäche als theoretisch. KI baut den Pfad, dann geht es schnell



theoretisch

PHP Foundation: die Behebung stockte, auch wegen „the lack of a known path to remote code execution“

ohne Anmeldung

Codeausführung aus der Ferne auf Servern mit PHP, unter geeigneten Bedingungen (PHP Foundation)

schneller

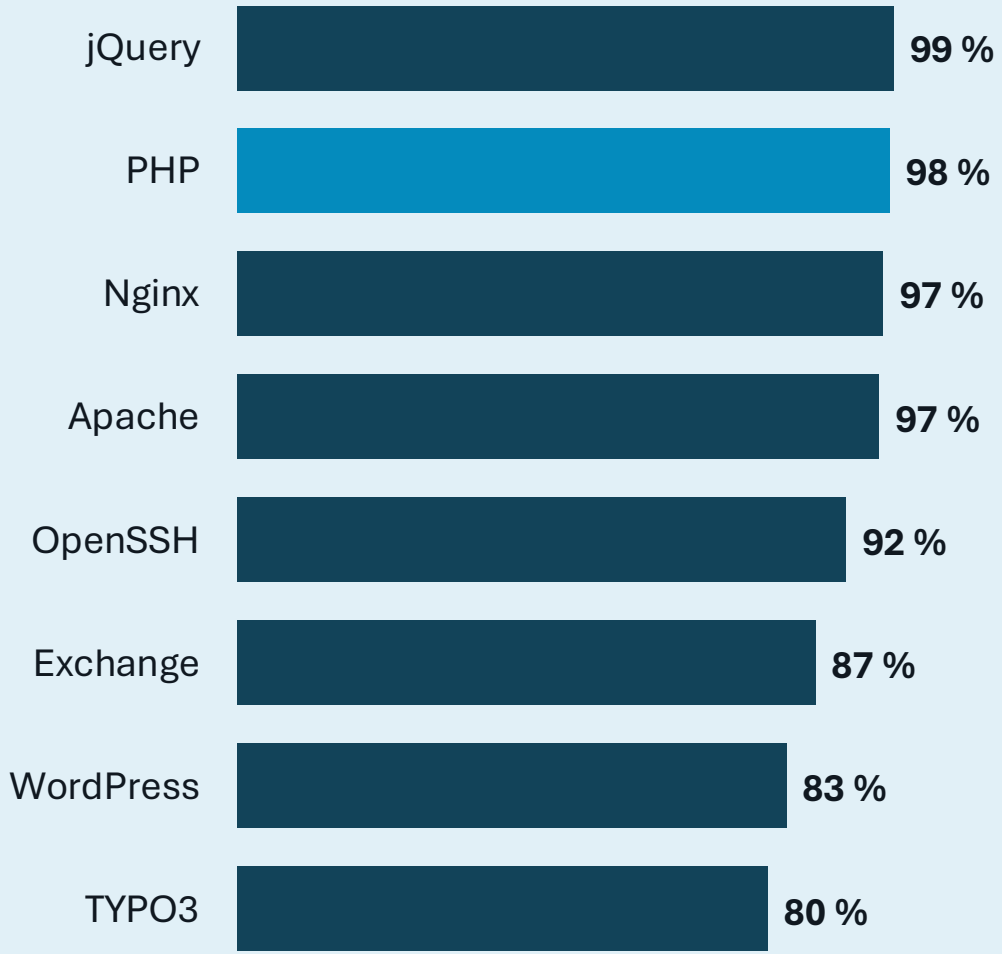
„AI-assisted analysis makes it cheaper to turn a patch diff or advisory into a working exploit.“ (AWS)

offen

Woher der erste Angreifer sein Wissen hatte, ist nicht bekannt; ebenso Zahl der Opfer und KI-Anteil am Angriff

„Kein Exploit bekannt“ ist keine sichere Einstufung mehr.

Anteil der 321 Organisationen, die bestimmte Bausteine nutzen



Klumpenrisiko: eine Lücke, alle Branchen

- Wie bei PHP kann jede beherrschende Technologie zum gemeinsamen Einfallstor werden
- jQuery und Apache mit bekanntermaßen ausgenutzter Lücke in allen 20 Branchen, PHP in 19
- PHP: Version nur bei 13 % der Hosts sichtbar, davon 55 % ohne Sicherheitsupdates
- **Folgerung:** KI beschleunigt das Finden von Schwachstellen in verbreiteten Bausteinen
  → **das Klumpenrisiko wächst**

ATHENE-Messdaten 02/2025–11/2025

6

KI auf beiden Seiten

Angreifer nutzen sie, Verteidiger auch

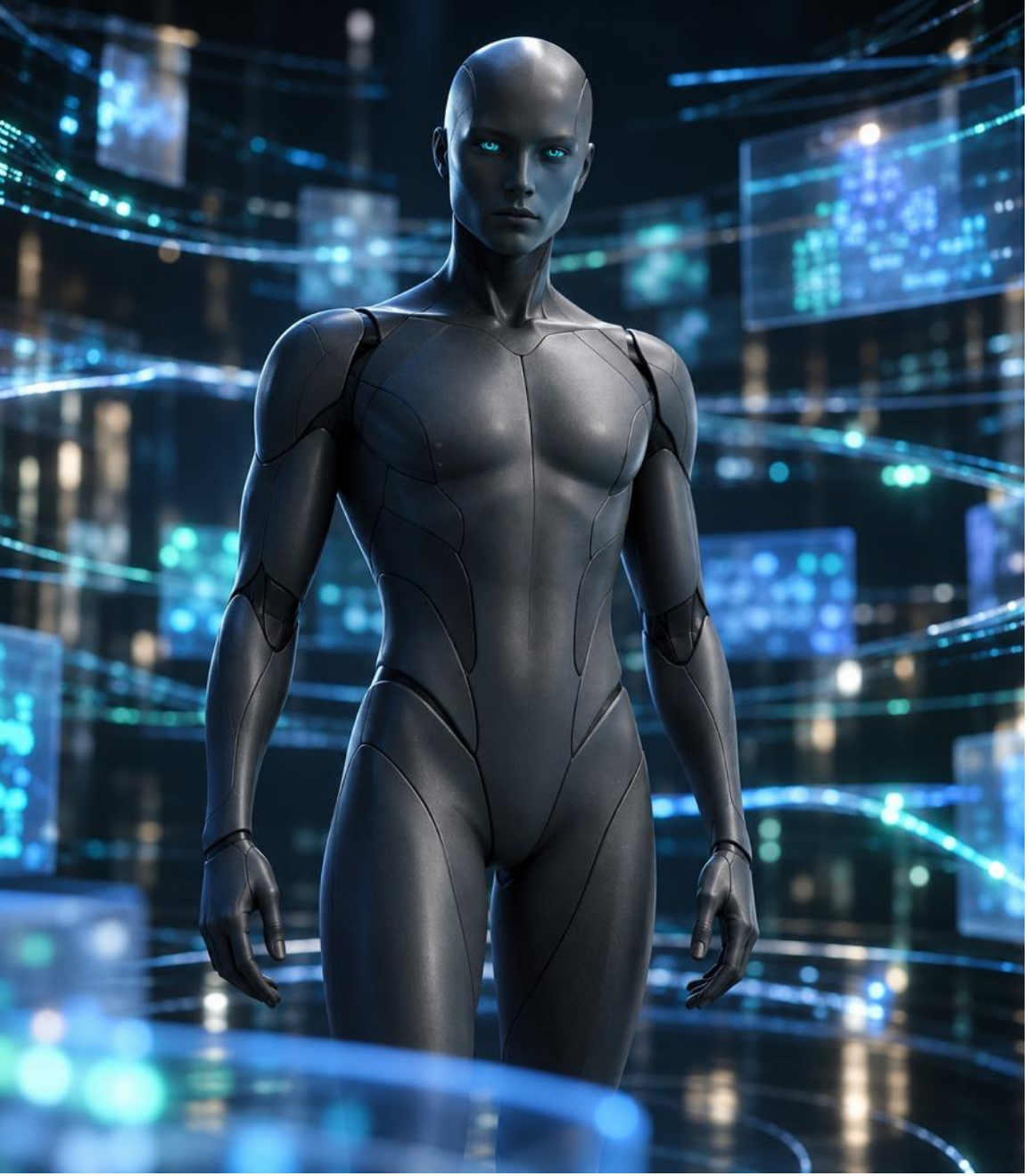


Bild: KI-generiert (OpenAI)

Angreifer nutzen KI, nachweislich

- Schneller, ausdauernder, in großem Maßstab; bekannte Schwächen werden schneller zu Angriffen

- **Google, Mai:** UNC2814 nutzt Gemini für Schwachstellenforschung an Router-Firmware
- **Google, September:** Agenten-Kampagne auf Zugangsdaten, von der Planung bis zur Ausführung in weniger als 6 Stunden, aus einer zuvor übernommenen Cloud-Umgebung
- **Anthropic, September:** mutmaßliche ShinyHunters-Affiliates beschleunigen Scans mit Claude; ein Betreiber durchsuchte 1,8 Mio. Android-App-Pakete (APK) nach eingebauten Zugangsdaten und Schlüsseln

Beobachtet von Google und Anthropic, 2026

Befragung und Messung

57 %

der Unternehmen nutzen KI (Befragung, Bitkom)

41 %

mit Verbindung zu GenAI-Anbietern (Messung)

42

Organisationen mit KI-Hostnamen, die zugleich test, dev oder admin enthalten (Messung)

Auch Organisationen nutzen KI, und das ist von außen sichtbar

- **Befragung:** 57 % der Unternehmen ab 20 Beschäftigten nutzen KI (Bitkom)
- **Messung:** 41 % der Organisationen sind per DNS, Skript oder Link mit GenAI-Anbietern verbunden
- **Messung:** Hostnamen mit KI-Begriff bei 40 %; bei 42 Organisationen tragen sie zugleich Begriffe wie test, dev oder admin
- **Grenze:** Ob produktiv und geschützt, zeigt die Außensicht nicht. Auftrag: das eigene KI-Inventar prüfen
- **Lücken finden:** Rund 50 Partner fanden mit KI in einem Monat mehr als zehntausend schwere Lücken; mehrere melden eine mehr als zehnfache Fundrate
- **Vor dem Angreifer:** Bei PHP wiesen Forscher mit KI-Hilfe den Angriffspfad nach; das Update folgte am 21. Oktober
- **Auffälliges erkennen:** Hugging Face erkannte den Angriff des OpenAI-Agenten mithilfe einer KI und stoppte ihn
- **Automatisieren:** BSI: „Anomalien detektieren und sofort reagieren“ erfordert „einen hohen Automatisierungs- und Reifegrad“

Gegen Angriffe in Maschinengeschwindigkeit hilft nur Abwehr in Maschinengeschwindigkeit.

7

Aktive Cyberabwehr

Dort ansetzen, wo der Angreifer seine Infrastruktur aufbaut

Der Bundestag entscheidet über die aktive Cyberabwehr



BSI

Anordnungen auch an Registries (§ 16a) und an Cloud-, CDN-, Rechenzentrums- und Managed-Service-Anbieter (§ 17 Abs. 2); laufende Indikatoren der KRITIS-Betreiber (§ 31 Abs. 3)

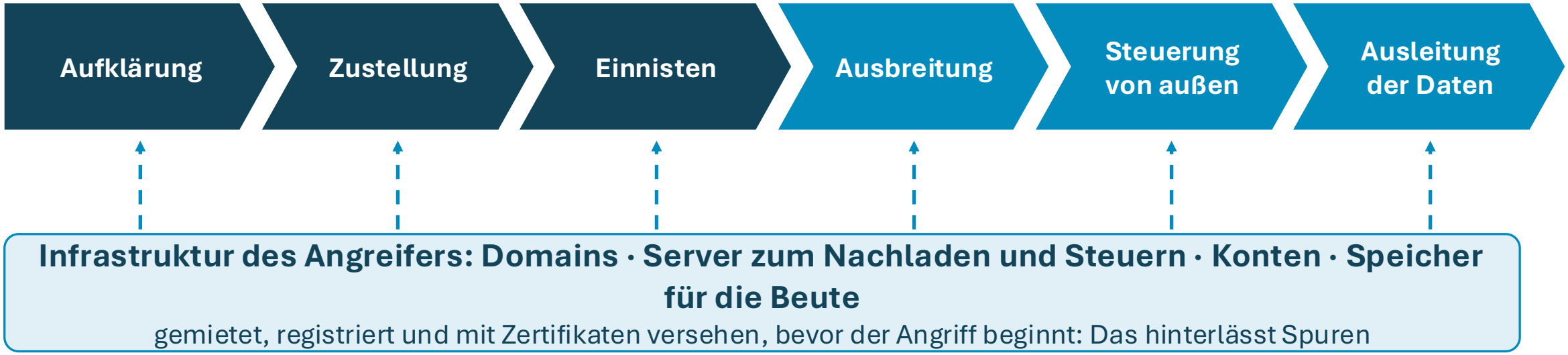
BKA

Abwehr von Cybergefahren für den Bund und bei Gefahren, die nur mit anderen Staaten abzuwehren sind (§ 3a); Verkehr umleiten, Betrieb untersagen, Daten löschen oder verändern (§§ 68b–68d)

Bundespolizei

Dieselben Befugnisse für die eigenen Aufgaben, laut Begründung etwa Grenzkontrollsysteme und Bahn (§ 41a BPolG-E)

Jeder Angriff braucht Infrastruktur, und die entsteht vor dem Angriff



Aktive Cyberabwehr: drei Aufgaben



Lagebild

- Wissen, was sich an der eigenen Angriffsfläche verändert
- Die Infrastruktur der Angreifer lokalisieren, bevor der Schaden eintritt



Schadens-begrenzung

- Warnen, umleiten, sperren
- Ohne in fremde Systeme einzudringen



Außer Betrieb setzen

- Über Anbieter und Registries
- Als letztes Mittel: Eingriff in das System des Angreifers

Nichts im Entwurf erlaubt, einem Angreifer mehr Schaden zuzufügen, als die Abwehr der Gefahr verlangt.

Schulmann/Waidner, FAZ, 25.04.2022 und 30.03.2026; Stellungnahme Schulmann, Ausschussdrucksache 21(4)221 E

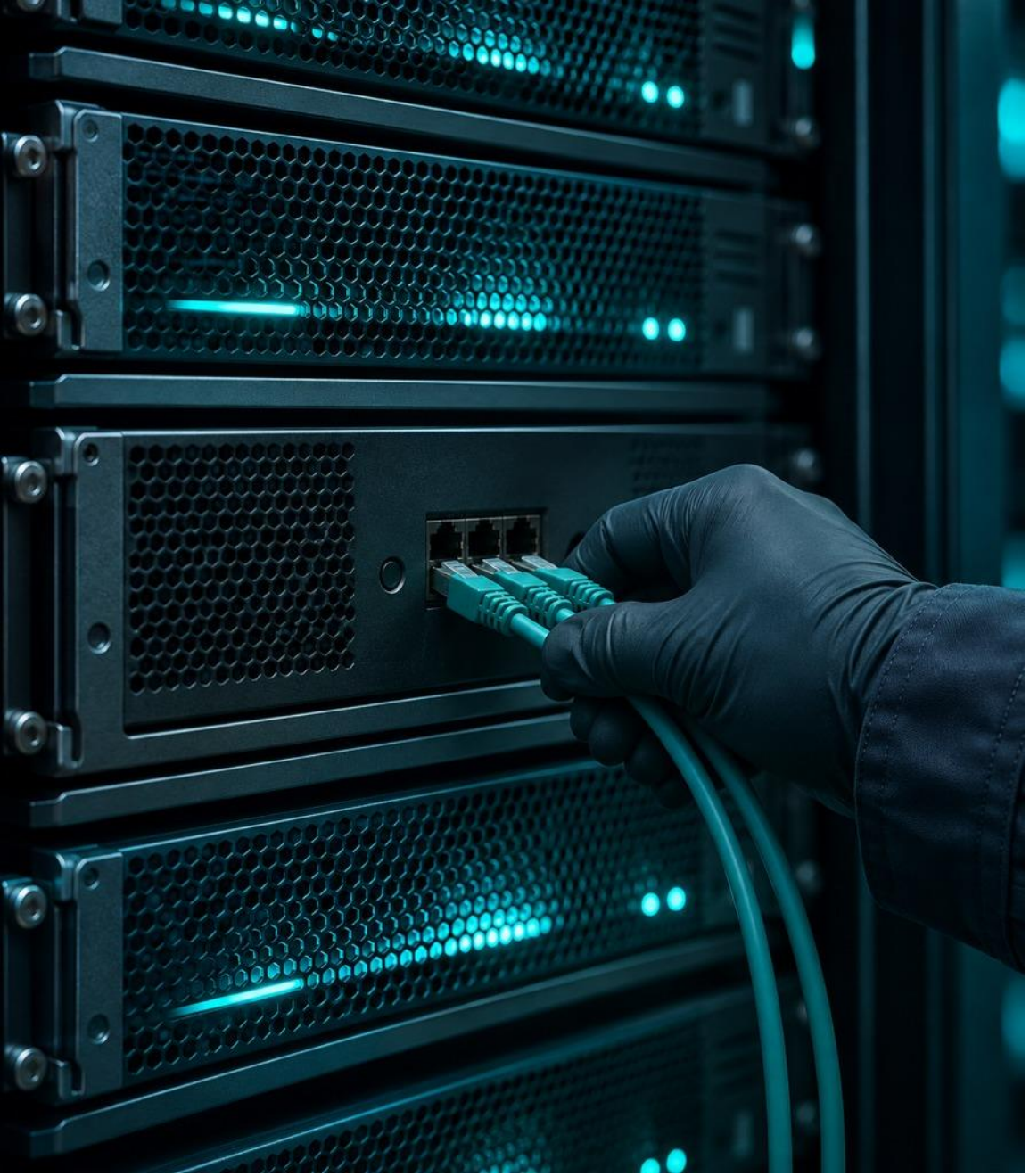


Bild: KI-generiert (OpenAI)

Es funktioniert: Unser Monitoring hat Angriffe gestoppt

- Das Monitoring ruft Seiten regelmäßig ab, vergleicht sie und schlägt an, wenn binnen kurzer Zeit auf vielen Seiten derselbe neue Inhalt erscheint
- **2025:** Dutzende laufender Angriffe dieser Art erkannt und verhindert
- **Gelöschter Cloud-Speicher:** Ein Website-Dienstleister hatte ihn aufgegeben, aber Webanwendungen öffentlicher Einrichtungen luden weiter Code daraus; nach unserer Meldung wurde die Ressource gesichert

Der Fehler lag außerhalb des Netzes der Opfer, und von außen war er sichtbar.

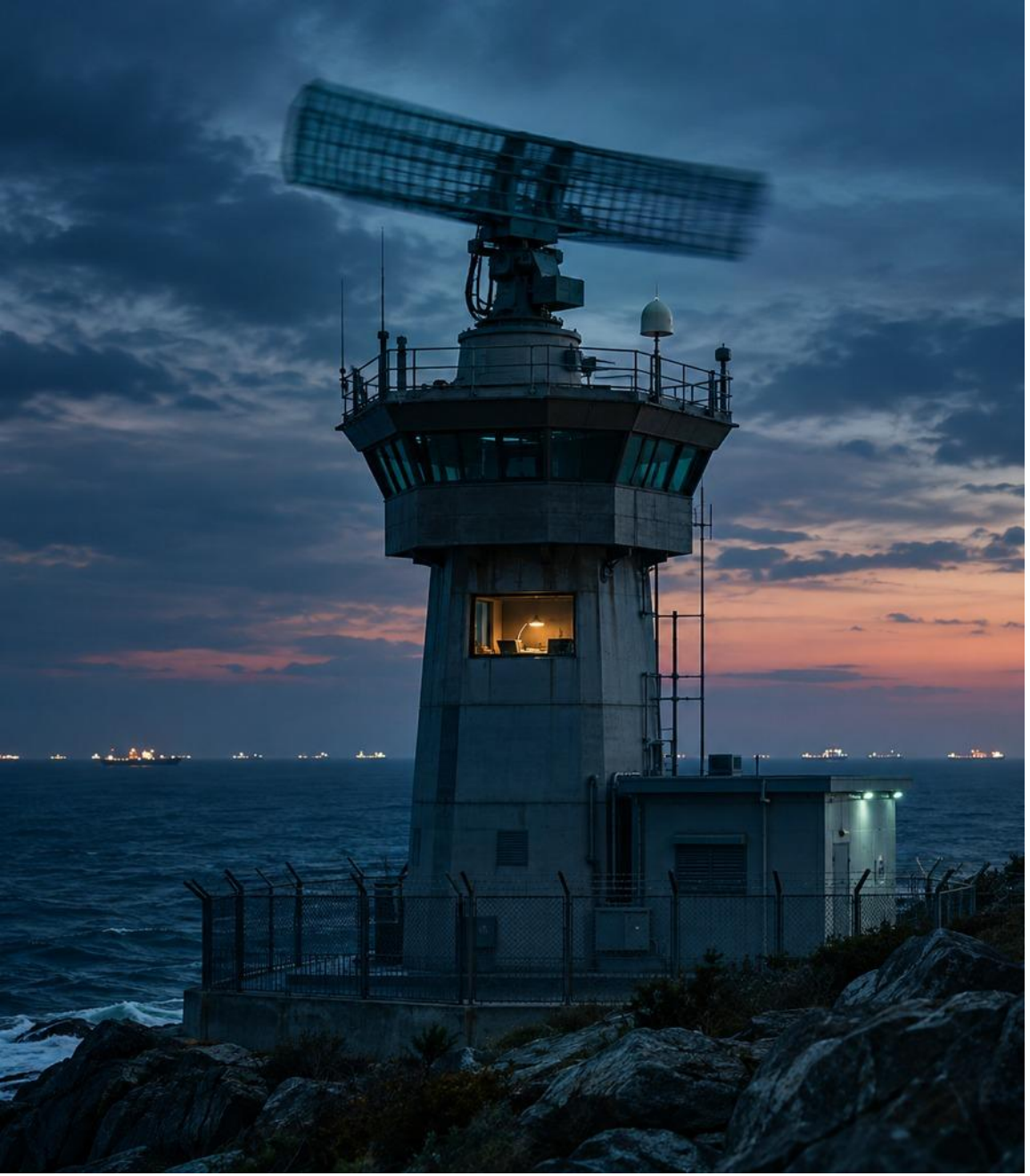


Bild: KI-generiert (OpenAI)

Fazit: Wir verteidigen uns gegen die Angreifer von gestern

- **Hausaufgaben:** alte Lücken schließen, offene Zugänge zu
- **Abhängigkeiten kennen:** ausgelagert ist nicht abgesichert
- **Schneller werden:** patchen und erkennen in Tagen, mit KI auf unserer Seite
- **Früher ansetzen:** an der Infrastruktur der Angreifer, bevor der Schaden eintritt

Abwehr beginnt vor dem Angriff.

תודה רבה!

Merci beaucoup!

çok teşekkürler

谢谢

Thank you very much!

Vielen Dank!

Dank je wel!

Muchas gracias

ありがとうございます

Dziękuję!

zor spas

Grazie mille!

شكرا لك