

IKT-SICHERHEITSKONFERENZ 2026 · LINZ · 16.–17. SEPTEMBER

Schutzschild für die KI-Ära: Strategische Cybersecurity **ohne** **Silos**

A Shield for the AI Era — Strategic Cybersecurity without Silos

Angreifer nutzen heute autonome KI, um Schwachstellen zu finden und Angriffe in Minuten zu automatisieren. Diese Lücke zu schließen bedeutet, die Silos zwischen IT, OT, Cloud, Identität und KI aufzubrechen.

Ing. Gregor Muck, MSc
Security Engineer, Tenable



Elf Minuten

09:41

Ein Angreifer ohne Exploit-Erfahrung bittet ein KI-Modell um Hilfe bei einer öffentlich bekannten Schwachstelle.

09:44

Das Modell liefert einen funktionsfähigen Exploit.

09:52

Der Angriff läuft - automatisiert, gegen vierzig Ziele gleichzeitig.

Elf Minuten. Kein Spezialwissen. Vierzig potentielle Opfer.

Illustrative scenario

Das ist keine Fiktion mehr

174 / 178

MOAK — Mother of All KEVs

- Five-agent pipeline: Collector, Researcher, Builder, Exploiter, Judge
- React-to-shell exploit in 21 Minuten — vollständig autonom, ohne Mensch im Prozess
- KEVs exploited autonomously, jede erst nach dem Trainings-Stichtag des jeweiligen Modells veröffentlicht*
- Nur öffentliche Modelle (Opus 4.6, GPT-5.4, Gemini) — kein Zugang zu Frontier-Modellen nötig

*Schwachstelle zum Zeitpunkt des Modelltrainings noch nicht bekannt · Quelle: moak.ai — Niv Hoffman, Yair Saban

THE PROOF

Selbst "Low" ist heutzutage **ausnutzbar**

Live example from moak.ai — CVE-2026-17459, SparkJava

CVE

CVSS

KEY

DISCLOSED

AI EXPLOITATION TIME

AI EXPLOITATION STATUS ?

1 Analyzed

^

CVE-2026-17459

2.1

2026-07-26

49m 49s

Fully Exploited by AI

A vulnerability was determined in perwendel spark up to 2.9.4. This vulnerability affects the function `staticFiles.externalLocation` of the file `src/main/java/spark/resource/ExternalResourceHandler.java` of the component SparkJava. Executing a manipulation can lead to symlink following. It is possible to launch the attack remotely. The exploit has been publicly disclosed and may be utilized. The project was informed of the problem early through an issue report but has not responded yet.

✦ Pipeline Tracker

✓ Researcher

✓ Builder

✓ Exploiter

✓ Judge

🔍

Searching: CVE-2026-17459 site:github.com commit

VULNERABILITY

The root cause is that SparkJava's directory traversal protection in `DirectoryTraversal.protectAgainstForExternal()` uses `Paths.get(path).toAbsolutePath()` to normalize paths before checking if a requested file is within the allowed static files directory.

file read

path traversal

ENVIRONMENT

Perwendel Spark 2.9.0 — OpenJDK 11

library

java

maven

EXPLOIT EVIDENCE

PASS

VULNERABLE

```
$ python3 CVE-2026-17459/exploit.py --host 127.0.0.1 --port 4567
[*] Target: 
[*] Requesting (symlink to /flag.txt via CVE-2026-17459)
FLAG{f49bcca2ac90613dde5fd86a54aeedc0}
EXPLOIT SUCCEEDED
```

PATCHED

```
$ python3 CVE-2026-17459/exploit.py --host 127.0.0.1 --port 4568
[*] Target: 
[*] Requesting (symlink to /flag.txt via CVE-2026-17459)
Fail: HTTP 404
Exit code: 1
EXPLOIT MITIGATED
```

Source: moak.ai, own live query

CVSS 2.1 (Low) · Disclosed 26 Jul 2026 · "exploited" by AI in **49m 49s**

Der erste öffentlich dokumentierte autonome Angriff

17.000+

protokollierte, autonome Angriffsaktionen (Juli 2026, Hugging Face & Modal Labs)

- Verkettete acht bis neun Zero-Days und brach aus – ohne menschliche Anweisung
- Richtete einen verdeckten Koordinationskanal ein; baute ihn nach der Abschaltung innerhalb weniger Tage robuster wieder auf
- **Keine böswillige Absicht nötig** — ein zu vage definiertes Ziel und zu viel Handlungsspielraum reichen aus

Quelle: CSA Post-Mortem (27. Juli 2026); OpenAI Black Hat USA Briefing, 5. Aug. 2026

Der Wandel im Angriffsmodell

VORHER

Tiefes Exploit-Wissen erforderlich

Entwicklungszeit von Tagen bis Wochen

Manuelle Code-Analyse

JETZT

KI-Modell analysiert die Schwachstelle

Automatisiertes Angriffsskript

Ergebnis in Minuten

Die Zahlen dieser Woche

MTTE

**Minuten
bis Stunden**

Mean Time to Exploit — durch KI-gestützte Angreifer

MTTR

**Wochen
bis Monate**

Mean Time to Remediate — Unternehmensdurchschnitt

“

Die Lücke zwischen Angriffsgeschwindigkeit und Reaktionszeit ist kein Personalproblem mehr. Sie ist **strukturell**

Realitätscheck: 31 % der Breaches betreffen weiterhin eine ungepatchte CVE, und die durchschnittliche Zeit, eine kritische zu patchen, liegt bei 43 Tagen.

Quellen: MTTE — moak.ai; MTTR — IBM Cost of a Data Breach Report 2026; 31 % / 43 Tage — Tenable, "Putting AI to Work" (Exposure26)

Das NISG 2026 tritt in Wochen in Kraft, nicht in Jahren



1. OKT
2026

Die Durchsetzung beginnt für wesentliche und wichtige Einrichtungen im Rahmen der österreichischen NIS2-Umsetzung.



~5.000 Organisationen

direkt reguliert in 18 Sektoren — Energie, Transport, Trinkwasser, Gesundheit, Fertigung und mehr



Bis zu 10 Mio. € oder 2 %

des weltweiten Jahresumsatzes für wesentliche Einrichtungen bei Nichteinhaltung



Bereichsübergreifendes Risikomanagement

ist nun gesetzliche Pflicht über IT, OT und Lieferkette hinweg — nicht mehr nur eine IT-Hygiene-Checkbox

Der **VULNAMI** vor Ihrer Tür

59.793

neue CVEs bereits 2026 — Stand 1. September

2025 kam auf insgesamt 48.185.

2026 hatte diese Zahl bereits im August übertroffen — bei noch vier verbleibenden Monaten.

Ein System meldet 5.000 "kritische" Schwachstellen — für das Team bedeutet das Alert Fatigue statt Priorisierung.

CVSS misst, wie schwerwiegend eine Schwachstelle theoretisch ist — nicht, ob sie in Ihrer Umgebung überhaupt erreichbar oder relevant ist.

Nicht jede Angriffsfläche hat eine CVE-Nummer

Die 59.793 CVEs in Ihren Büchern sind nur ein Ausschnitt der echten Angriffsfläche — **63 % der Breaches beginnen mit einer Non-CVE-Schwachstelle.**



Fehlkonfiguration



Fehlende MFA



Geleaktes Passwort



EDR installiert, aber nicht aktiviert



Überprivilegierte Identität



...

Deshalb **Exposure Management** — nicht nur Vulnerability Management.

Falsche Priorität, echtes Risiko

90%

der Zeit verbringen Teams damit, Lücken zu patchen, die kein echtes Risiko darstellen — die tatsächlichen Einstiegspunkte bleiben offen

2–3%

aller bekannten Schwachstellen werden in der Praxis überhaupt aktiv ausgenutzt

“

*Das Ziel von Security-Teams kann nicht sein, jede Lücke zu schließen — sondern für Angreifer **die relevanten Angriffswege zu blockieren***



Von Vulnerability Scanning zu Exposure Management

Ein neues Mindset für ein neues Bedrohungsmodell

Eine Formel für Ihren Notizblock

Reales Risiko = Schwachstelle × Angriffspfad × Business-Kritikalität

1

Schwachstelle

Wie schwerwiegend ist die Lücke?
Vergleichbar über CVE,
Fehlkonfiguration, kompromittiertes
Konto, fehlende MFA und mehr.

2

Angriffspfad

Ist das System vom Internet aus
erreichbar — oder durch Segmentierung
isoliert? Welche mildernden Kontrollen
bestehen?

3

Business-Kritikalität

Sind das die Kronjuwelen — oder der
Kantinen-Speiseplan?

Aus Rauschen wird **Klarheit**

Tenable One bringt Schwachstellen, Angriffspfade und Business-Kritikalität automatisiert zusammen – über IT, Cloud und OT hinweg.

Das Ergebnis: Tausende Findings werden zu den wenigen Pfaden, die wirklich zählen.



Die Maschine im Team



KI-Agenten

Triage · Korrelation von Hintergrunddaten · Automatisierte Patches für Standardsysteme — in Maschinengeschwindigkeit



Der Mensch

Risikoakzeptanz · Strategische Entscheidungen · Geschäftskontext

“

Bekämpfen Sie KI-gestützte Angriffe nicht mit Überstunden Ihres Teams. Bekämpfen Sie sie mit Ihrer eigenen KI an Ihrer Seite.



Vier Gedanken für den Weg nach vorne

1

Angriffsgeschwindigkeit ist heute **strukturell** — kein Personalproblem.

2

Nicht jede Lücke schließen — die relevanten Angriffswege **blockieren**.

3

Reales Risiko = Schwachstelle × Angriffspfad × Business-Kritikalität.

4

KI-gestützte Angriffe brauchen **KI-gestützte Abwehr** — und KI-gestützte Governance.

Vielen Dank

Let's talk about your exposure.

Ing. Gregor Muck, MSc
Security Engineer, Tenable