02.02.2026 – 09:04 Uhr

Kriminelle KI ist ein Geschäftsmodell – kein Experiment mehr

Kriminelle KI ist ein Geschäftsmodell – kein Experiment mehr

Wallisellen, 02. Februar 2026 – Trend Micro, einer der weltweit führenden Anbieter von Cybersicherheitslösungen, veröffentlicht heute neue Forschungsergebnisse zur kriminellen Nutzung von Künstlicher Intelligenz (KI). Demnach hat KI-gestützte Cyberkriminlität die experimentelle Phase hinter sich gelassen und ein stabileres, professionelleres Ökosystem entwickelt, das der Reifung anderer etablierter Untergrundmärkte ähnelt.

Die Studie *Criminal AI in 2025: The Year the Underground Went Pro* analysiert, wie Cyberkriminelle KI zum Jahreswechsel einsetzten. Grundlage waren Auswertungen von Untergrunddiensten, Malware-Samples und aktiven Angriffskampagnen. Die Ergebnisse zeigen, dass Kriminelle nicht länger auf Neuheit setzen, sondern bestehende Methoden verfeinern, Kosten senken und in Bereichen wie Betrug, Malware-Entwicklung und Deepfake-gestützter Kriminalität auf zuverlässigere Technologien zurückgreifen können.

Es handelt sich um das vierte Update in Trend Micros fortlaufender Forschungsreihe zur kriminellen Nutzung von KI. Frühere Berichte zeigten das Auftreten manipulierter („jailbroken“) Chatbots, erste Deepfake-Betrugsmaschen und vereinzelte Proof-of-Concept-Angriffe. Die aktuellen Ergebnisse belegen, dass sich diese Ansätze inzwischen zu reproduzierbaren Services entwickelt haben, die ohne Spezialwissen gemietet, wiederverwendet und skaliert werden können.

„Für 2026 zeichnet sich kein radikaler Umbruch ab, sondern eine schleichende, kontinuierliche Entwicklung“, sagt

David Sancho, Senior Threat Researcher bei Trend Micro. „Wir werden keine plötzliche Explosion KI-getriebener Chaos-Szenarien erleben. Stattdessen beobachten wir eine stetige, professionelle Weiterentwicklung der vorhandenen Werkzeuge. Genau diese leise, kontinuierliche Optimierung macht kriminelle Ökosysteme besonders schwer angreifbar."

Die Studie identifiziert drei Entwicklungen, die das aktuelle Bild krimineller KI prägen:

- **Konsolidierung:** Zwar tauchen in Untergrundforen weiterhin neue „ungefilterte" KI-Tools auf, die meisten verschwinden jedoch schnell wieder. Beständig sind vor allem Dienste, die auf Jailbreaks und Prompt-basierten Missbrauch etablierter KI-Plattformen setzen, anstatt eigenständige Modelle zu entwickeln. Kriminelle nutzen also gezielt die Skaleneffekte und Investitionen kommerzieller KI-Anbieter, anstatt deren Modelle selbst nachzubilden.
- **Das Auftreten erster Malware-Varianten**, die bösartigen Code über eingebettete oder entfernte KI-Abfragen erzeugen oder verändern können. Diese Beispiele sind zwar noch durch praktische Einschränkungen limitiert, deuten jedoch auf eine Entwicklung hin zu anpassungsfähigeren Malware-Designs, die sich von Infektion zu Infektion unterscheiden können.
- **Die rasche Verbreitung von Deepfake-Technologien** im alltäglichen kriminellen Einsatz. Tools für Face Swapping, Voice Cloning und Bildmanipulation sind inzwischen weit verbreitet und häufig kostenlos oder sehr günstig verfügbar. Dies ermöglicht neue Betrugswellen, darunter Identitätsbetrug, gezielte Infiltration von Unternehmen, virtuelle Entführungen sowie die Zunahme nicht einvernehmlicher synthetischer Inhalte. Für Einzelpersonen, Unternehmen und das Vertrauen in digitale Identitäten haben diese Deepfake-Technologien gravierende Folgen.

Die Analyse von Trend Micro zeigt, dass Verteidiger aktuell, dank KI-gestützter Erkennung, Threat Intelligence und automatisierter Analysewerkzeuge, noch im Vorteil sind. Dieser Vorsprung schrumpft jedoch, da Kriminelle zunehmend lernen, aus denselben Technologien Mehrwert zu ziehen und dies oft schneller umsetzen, als Schutzmechanismen durchgesetzt werden können.

Die Studie kommt zu dem Schluss, dass das grösste Risiko nicht in einem sprunghaften Anstieg der Fähigkeiten der Angreifer liegt, sondern in der Normalisierung KI-gestützter Kriminalität. Sobald diese Techniken günstiger, berechenbarer und leichter nutzbar werden, verbreiten sie sich dauerhaft im gesamten Cybercrime-Ökosystem.

„Für Unternehmen bedeutet dieser Wandel, dass KI-gestützte Angriffe als Teil des Alltags und nicht mehr als Ausnahmeerscheinung betrachtet werden müssen", ergänzt David Sancho. „Deepfake-basierter Betrug, Identitätsmissbrauch und KI-unterstützte Malware sind keine Randphänomene mehr, sondern Risiken, die in Sicherheitsstrategien, Verifizierungsprozessen und der Incident-Response-Planung fest mitbedacht werden müssen."

Weitere Informationen

Die vollständige Studie *Criminal AI in 2025: The Year the Underground Went Pro* finden Sie hier:
https://www.trendmicro.com/vinfo/de/security/news/cybercrime-and-digital-threats/the-state-of-criminal-ai

Pressestelle Trend Micro Schweiz
c/o BRAND AFFAIRS AG

Mischa Keller / MSc Business Administration
Partner

Telefon: +41 44 254 80 00
E-Mail: trendmicro-media@brandaffairs.ch
Mühlebachstrasse 8 / 8008 Zürich / Switzerland

Medieninhalte

David Sancho, Senior Threat Researcher bei Trend Micro.