

KI-4-Everyone · Daily News

27. Juli 2026



SAFE

Nvidia und Microsoft gründen Sicherheitsbündnis ohne OpenAI

Nvidia, Microsoft, SpaceX und IBM wollen gemeinsam quelloffene Sicherheitswerkzeuge gegen KI-Angriffe entwickeln. OpenAI, Google und Anthropic fehlen im Bündnis.

PROD

GitHub Copilot als einziges Werkzeug für den ganzen Code-Prozess

GitHub zeigt, wie Copilot von der ersten Idee bis zur fertigen Software reicht – ohne ständig neue KI-Tools auszuprobieren.

Nvidia und Microsoft gruenden Open-Source-Buendnis fuer KI-Sicherheit - ohne OpenAI, Google und Anthropic

Die neue Open Secure AI Alliance will offene Werkzeuge gegen Angriffe auf KI-Systeme bauen. Auffaellig ist vor allem, wer nicht dabei ist.

Wenn grosse Technologiekonzerne ein neues Sicherheitsbuendnis gruenden, ist die spannendste Frage oft nicht, wer mitmacht - sondern wer fehlt. Nvidia und Microsoft haben gemeinsam mit weiteren Firmen die Open Secure AI Alliance ins Leben gerufen, ein Zusammenschluss fuer offene Werkzeuge rund um die Sicherheit von KI-Systemen. Ausgerechnet die drei Unternehmen, deren Modelle heute die Debatte um KI-Risiken bestimmen, sind nicht dabei: OpenAI, Google und Anthropic. Damit zieht sich mitten durch die Branche eine Linie zwischen jenen, die auf Offenheit setzen, und jenen, die ihre Sicherheitsarbeit lieber im eigenen Haus behalten.

Bekannt gegeben wurde das Buendnis am 27. Juli 2026. Zu den Gruendungsmitgliedern gehoeren laut den vorliegenden Meldungen Nvidia, Microsoft, SpaceX und IBM sowie weitere, im Material nicht namentlich genannte Technologiefirmen. Ziel der Allianz ist es, quelloffene Sicherheitswerkzeuge fuer KI zu entwickeln und zu teilen - also Software, deren Programmcode oeffentlich einsehbar ist und die von Fachleuten weltweit gemeinsam geprueft werden kann. Die Initiative begruendet ihren Schritt damit, dass sich Angriffe auf sogenannte Frontier-Modelle, also die derzeit leistungsfaeigsten KI-Systeme, nur mit offenen Werkzeugen wirksam abwehren liessen. Open Source, so das Argument von Nvidia in einem eigenen Blogbeitrag, sei ohnehin eine tragende Saeule fuer Cloud-Dienste, Finanzwesen, Industrie, Telekommunikation, Behoerden und Internetdienste - und Cybersicherheit gehoere zu den drei Bereichen, die am staerksten davon profitierten.

Der Vorstoss trifft einen wunden Punkt der Branche. Waehrend die Sorgen um Sicherheitsluecken, Manipulation und Missbrauch von KI-Modellen wachsen, halten die groessten Anbieter geschlossener Systeme - OpenAI mit seiner GPT-Reihe, Google mit Gemini, Anthropic mit Claude - ihre Sicher-

heitsmethoden weitgehend unter Verschluss. Nvidia und Microsoft positionieren sich mit dem Buendnis als Gegenpol: Sicherheit soll nachvollziehbar sein, nicht Vertrauenssache. Dass gerade Microsoft dabei ist, ist bemerkenswert - der Konzern ist selbst eng mit OpenAI verflochten, treibt hier aber ein Projekt voran, das ohne den prominenten Partner auskommt. Fuer Nvidia wiederum passt die Rolle: Wer die Chips fuer nahezu alle grossen KI-Systeme liefert, hat ein Interesse daran, dass die Infrastruktur drumherum breit abgestuetzt ist und nicht von einzelnen Modellanbietern kontrolliert wird. Unter Druck geraten koennten die abwesenden Konzerne: Wenn sich offene Sicherheitsstandards durchsetzen, muessen auch geschlossene Modelle erklaren, warum sie ihnen nicht folgen.

Vieles bleibt zunaechst offen. Aus dem Material geht nicht hervor, welche konkreten Werkzeuge die Allianz zuerst veroeffentlichen wird, nach welchen Regeln entschieden wird und wie sich das Buendnis finanziert. Auch die vollstaendige Liste der Mitglieder ist in den vorliegenden Quellen nicht genannt. Ebenso unklar ist, ob OpenAI, Google und Anthropic bewusst nicht eingeladen wurden, ob sie eine Teilnahme abgelehnt haben oder ob sich das noch aendern kann - dazu findet sich in den Meldungen kein belastbarer Hinweis. Ein weiteres Risiko liegt in der Natur offener Sicherheitswerkzeuge selbst: Was Verteidiger einsehen koennen, koennen im Zweifel auch Angreifer studieren. Ob die Allianz darauf eine ueberzeugende Antwort hat, wird sich erst zeigen, wenn erste konkrete Projekte vorliegen.

In den kommenden Wochen lohnt der Blick darauf, welche Werkzeuge die Allianz zuerst veroeffentlicht, ob weitere namhafte Firmen beitreten - und ob sich die abwesenden Anbieter oeffentlich zu dem Buendnis aeussern.

REG

KI-Firmen geben Rekordsummen für Washington-Lobbying aus

KI-Unternehmen investieren so viel wie nie zuvor in politische Einflussnahme in Washington. Genaue Zahlen nennt das Material nicht. Das Ziel: günstige Regulierung sichern.

MARKT

Analyst: Apple profitiert, wenn die KI-Blase platzt

Apple könnte beim Platzen der KI-Blase als Gewinner dastehen. Diese These steht hinter dem Titel 'Watch Everything Burn'. Konkrete Begründungen liefert das Material nicht.

REG

KI-Training zerstört seltene Bücher

KI-Unternehmen vernichten angeblich seltene Bücher, um Trainingsdaten zu gewinnen. Das Material nennt keine konkreten Firmen oder Zahlen. Der Vorwurf löst Debatten über Urheberrecht und Kulturerhalt aus.

PROD

MAI-Cyber-1-Flash jetzt in MDASH integriert

MAI-Cyber-1-Flash ist nun Teil von MDASH. Was das konkret bedeutet und welche Funktionen sich ändern, nennt das Material nicht. Unklar, für welche Nutzergruppe das relevant ist.

MARKT

Cognizant und Anthropic weiten Partnerschaft für Enterprise aus

Cognizant und Anthropic bauen ihre Zusammenarbeit aus. Unternehmenskunden sollen dadurch Zugang zu Claude erhalten. Details zu Umfang oder Preisen nennt das Material nicht.

REG

US-Bürger angeklagt: GrapheneOS löschte Handy bei Grenzkontrolle

Ein US-Bürger steht vor Gericht, weil sein GrapheneOS-Gerät beim Flughafenscheck Daten löschte. Die Behörden werten das als Behinderung. Das Material nennt keine weiteren Details zum Fall.

PROD

GitHub Copilot App: Einstiegsguide für Einsteiger veröffentlicht

GitHub erklärt Neulingen, wie sie mit der Copilot App starten. Der Guide behandelt Projekte anlegen, KI-Agenten nutzen und Canvases erkunden. Ziel ist ein schnellerer Einstieg in den Entwicklungsworkflow.

MARKT

Nvidia: 750-Mrd.-Dollar-Deals befeuern Kreislauf-Kritik

Nvidia schließt Deals im Wert von 750 Milliarden Dollar ab. Kritiker warnen vor einem künstlichen Kreislauf: KI-Firmen kaufen gegenseitig Chips und Dienste, um Umsätze hochzuhalten. Ob das strukturell problematisch ist, bleibt laut Material offen.

PROD

Microsoft veröffentlicht Fara1.5-27B: Bild + Text als Eingabe

Das Modell Fara1.5-27B verarbeitet Bilder und Texte zusammen und kann damit Bildschirmhalte steuern. Es eignet sich laut Tags für Computer-Use-Szenarien.

PROD

NVIDIA bringt KI-Echtzeitsimulation in den OP-Saal

Cosmos-H-Dreams von NVIDIA soll Chirurgie-Robotik mit generativer Echtzeitsimulation unterstützen. Konkrete Einsatzszenarien im OP werden damit trainierbar.

PROD

NVIDIA nutzt eigene Vera-CPU, um Chip-Design zu beschleunigen

NVIDIA setzt die Vera-CPU gemeinsam mit Cadence und Synopsys ein, um EDA-Software für die nächste Chip-Generation zu optimieren. Die Entwicklung modernster CPUs und GPUs soll damit schneller werden.

OS

FeyNoBg: Open-Source-Tool entfernt Bildhintergründe automatisch

Feyn veröffentlicht FeyNoBg, ein Modell zur automatischen Hintergrundentfernung, plus die Python-Bibliothek NoBg zum Trainieren eigener Varianten. Beides ist offen zugänglich.

PROD

MAI-Cyber-1-Flash: Microsofts erste eigene KI für Cybersicherheit

Das Modell ist auf Cyberangriffe spezialisiert und soll besonders kostengünstig sein. Trainiert wurde es auch mit Daten aus echten Hacking-Angriffen auf Microsoft-Nutzer.

PROD

3D per Chat: KI-Agenten erzeugen Objekte ohne CAD-Kenntnisse

Tools wie Meshys 3D Agent ermöglichen es, per Texteingabe 3D-Modelle zu erzeugen. Das eigentliche Problem laut Bericht: ob die Ergebnisse tatsächlich brauchbar sind.

Keine Termine gemeldet.