

# KI-4-Everyone · Daily News

2. September 2026



PROD

## Google stellt Gemini 3.8 Flash und eine Cyber-Variante vor

Google DeepMind hat zwei neue Modelle veröffentlicht: Gemini 3.8 Flash und die spezialisierte Variante 3.8 Flash Cyber.

REG

## New York City verbietet KI-Nutzung für rund 600.000 Schüler

Bürgermeister Zohran Mamdani hat ein einjähriges KI-Verbot für Klassen 2 bis 8 angekündigt. Es gilt ab dem Schuljahr 2026-2027.

# Google kündigt Gemini 3.8 Flash an - mit einer neuen Cyber-Variante

*Ein knapper Ankuendigungstext, viele offene Fragen: Was 'Flash Cyber' bedeutet, geht aus dem vorliegenden Material nicht hervor.*

**G**oogle hat eine neue Version seines schnellen KI-Modells vorgestellt: Gemini 3.8 Flash, begleitet von einer bisher unbekannten Variante namens 'Flash Cyber'. Der Name allein wirft mehr Fragen auf, als die Ankuendigung beantwortet - und genau darin liegt die kleine Geschichte des Tages. Denn waehrend die grossen Modellstarts sonst mit Benchmarks, Preisen und Demovideos daherkommen, bleibt hier auffallend viel im Nebel.

Die Fakten, soweit sie im Material belegt sind, bleiben duenn. Google DeepMind veroeffentlichte einen Ankuendigungsbeitrag mit dem Titel 'Introducing Gemini 3.8 Flash and 3.8 Flash Cyber', und zeitgleich tauchte in der KI-Community von Hacker News ein Verweis auf die zugehoerige Modellkarte (eine Art technisches Datenblatt, das Faehigkeiten und Grenzen eines KI-Modells beschreibt) auf. Beide Quellen nennen die zwei Modellnamen, ansonsten enthaelt das vorliegende Material keine weiteren Details zu Faehigkeiten, Preisen oder Verfuegbarkeit. Wer 'Flash Cyber' konkret adressiert - Sicherheitsforscher, Unternehmenskunden, Entwickler - geht daraus nicht hervor.

Einordnen laesst sich die Meldung trotzdem: Die 'Flash'-Reihe ist bei Google traditionell das schnelle, guenstigere Gegenstueck zu den grossen 'Pro'-Modellen und richtet sich an Anwendungen, in denen Tempo und Kosten wichtiger sind als absolute Spitzenleistung. Dass Google nun eine spezialisierte 'Cyber'-Variante nachschiebt, koennte ein Signal in Richtung Cybersecurity-Anwendungen sein - also KI-Modelle, die zum Beispiel Angriffscodes erkennen, Schwachstellen analysieren oder Sicherheits-

teams unterstuetzen. Das ist allerdings eine Vermutung aus dem Namen heraus; im vorliegenden Material ist diese Ausrichtung nicht ausdruuecklich belegt.

Unklar bleibt entsprechend viel. Es fehlen im Material konkrete Angaben dazu, welche Benchmarks Google fuer die neue Version nennt, ob sich der Kontextspeicher (die Menge an Text, die ein Modell gleichzeitig verarbeiten kann) veraendert hat, und ob 'Flash Cyber' fuer alle Nutzer offen ist oder nur fuer bestimmte Kundengruppen. Auch ob die 'Cyber'-Variante offensiv genutzt werden koennte - also selbst zur Analyse oder Erstellung von Angriffscodes - ist nicht im Material belegt. Genau das waere aber die spannende Frage, weil sicherheitsspezialisierte KI-Modelle einen doppelten Nutzen haben: Sie helfen Verteidigern und potenziell auch Angreifern.

Wer die Story weiterverfolgen will, sollte in den kommenden Tagen auf drei Dinge achten: erstens die vollstaendige Modellkarte bei Google DeepMind, weil dort Fahigkeiten und ausdruueckliche Einsatzgrenzen dokumentiert werden; zweitens Reaktionen aus der Sicherheits-Community, die eine 'Cyber'-Variante besonders kritisch pruefen duerfte; drittens die Frage, ob Konkurrenten wie OpenAI oder Anthropic nachziehen und ebenfalls sicherheitsspezialisierte Modellvarianten anbieten. Bis dahin gilt: Der Name klingt nach einer neuen Kategorie - aber ob dahinter wirklich ein neuer Produkttyp steckt oder nur ein Marketing-Label, ist im vorliegenden Material nicht zu entscheiden.

## SAFE

**Kritische Sicherheitslücke in Langflow wird aktiv ausgenutzt**

Angreifer nutzen aktiv eine kritische Schwachstelle im KI-Agenten-Framework Langflow aus. Die Lücke (CVE-2026-0768) ermöglicht Remote Code Execution. Viele KI-Applikationen, die auf Langflow aufbauen, sind betroffen.

## SAFE

**Drei Websites erstellen 215.128 Fake-Bestenlisten - Perplexity zitiert sie**

Drei Websites haben maschinell über 215.128 gefälschte 'Beste Software'-Seiten erzeugt. Der KI-Assistent Perplexity zitiert diese Seiten als Quellen. Das zeigt, wie KI-generierter Spam in KI-Antworten landet.

## REG

**Opt-out beim KI-Training: Was Nutzer wirklich wissen sollten**

Viele KI-Dienste verwenden deine Eingaben und Ausgaben zum Modelltraining. Ob und wie du widersprechen kannst, hängt stark vom jeweiligen Anbieter ab. Die Optionen sind oft versteckt oder eingeschränkt.

## SAFE

**Anthropic entwickelt Enterprise-Sicherheitsstandards mit Unternehmenskunden**

Anthropic arbeitet gemeinsam mit Unternehmenskunden an Sicherheitsleitlinien für KI-Systeme im Unternehmenseinsatz. Ziel ist es, praxistaugliche Schutzmaßnahmen direkt aus dem Kundenbetrieb abzuleiten.

## PROD

**Quasar 438B: Europas bislang führendes KI-Modell**

Quasar 438B wird als führendes KI-Modell aus Europa vorgestellt. Der Name verweist auf 438 Milliarden Parameter. Details zu Hersteller und Verfügbarkeit sind im vorliegenden Material unklar.

## OS

**Weedout: Safari-Erweiterung blendet KI-Videos auf YouTube aus**

Ein Entwickler hat Weedout gebaut, weil KI-generierte Verschwörungsvideos seinen YouTube-Feed überschwemmten. Die Erweiterung kostet 1,99 US-Dollar, läuft lokal und nutzt YouTube's eigenes 'Made with AI'-Label. Sie filtert Feed, Suche, verwandte Videos, Playlists und Shorts.

## SAFE

**curl findet sechs CVEs - OpenAI und Anthropic melden null**

Das curl-Projekt hat sechs Sicherheitslücken (CVEs) gemeldet. OpenAI und Anthropic kamen bei derselben Aufgabe auf null Treffer. Das wirft Fragen zur Verlässlichkeit von KI bei Sicherheitsanalysen auf.

## PROD

**Loops, Harnesses, Squads: Neues KI-Vokabular für Entwickler erklärt**

Der GitHub-Podcast erklärt Begriffe wie Loop Engineering, Harnesses, Squads und Hill Climbing, die zunehmend in Entwicklergesprächen auftauchen. Wer mit KI-Entwicklern spricht, begegnet diesen Wörtern immer häufiger.

PROD

**DeepSeek bringt schnelles Vision-Modell: Bilder und Text kombiniert**

Das neue DeepSeek-V4-Flash-Vision-Exp verarbeitet Bilder zusammen mit Text. Fast 18.000 Downloads zeigen hohes Interesse.

RES

**Microsoft veröffentlicht Bildklassifikation ohne Vorkenntnisse: colipri**

Das Modell colipri erkennt Bildinhalte, ohne vorher auf die gesuchten Kategorien trainiert worden zu sein. Das nennt sich Zero-Shot-Klassifikation.

RES

**Google veröffentlicht Zeitreihen-Prognosemodell TimesFM 3.0**

TimesFM 3.0 ist darauf ausgelegt, Werte in Zeitreihen – etwa Verkaufszahlen oder Temperaturen – vorherzusagen. Das Modell ist vortrainiert und frei verfügbar.

PROD

**Qwen3.8-Flash-Next: Alibabas kompaktes Bild-Text-Modell jetzt in FP8**

Das Modell kombiniert Bild- und Textverarbeitung und läuft durch das FP8-Format besonders sparsam auf der Hardware. Über 130.000 Downloads in kurzer Zeit.

PROD

**Microsoft bringt Echtzeit-Spracherkennungsmodell VibeVoice ASR**

VibeVoice-ASR-Streaming-1.5B transkribiert gesprochene Sprache direkt während der Aufnahme, also ohne Wartezeit nach dem Sprechen. Downloads sind bisher noch nicht gemeldet.

PROD

**GitHub Copilot: Wie kürzere KI-Antworten trotzdem teurer werden können**

GitHub erklärt, warum kurze Outputs nicht automatisch günstiger sind und wie Copilot unnötige Arbeit beim Coding-Prozess reduziert.

Keine Termine gemeldet.