

[Zur Webansicht](#)



Tipp: [Das KI-Update gibt es auch als Podcast](#)

Liebe Leserinnen & Leser,
hier kommt die Zusammenfassung der aktuellen Podcast-Folge, die Sie [hier in voller Länge](#) anhören können.

Warum KIs ostdeutsche Menschen generell niedriger bewerten

Eine Studie der Hochschule München zeigt, wie tief Vorurteile in großen Sprachmodellen stecken. [ChatGPT und andere Chatbots bewerteten Menschen aus Ostdeutschland](#) durchweg schlechter als Westdeutsche. Absurd wird es bei widersprüchlichen Urteilen: Die Modelle gaben Ostdeutschen niedrige Werte sowohl bei Fleiß als auch bei Faulheit. Noch grotesker: Die KIs behaupteten, Ostdeutsche hätten eine niedrigere Körpertemperatur als Westdeutsche.

Die Studie trägt den Titel "Saxony-Anhalt is the Worst" – nicht als Meinung der Autoren, sondern als Befund: Die getesteten Sprachmodelle wiederholten stur ein gelerntes Muster, das alles aus Sachsen-Anhalt besonders schlecht bewertet. Selbst explizite Anweisungen, herkunftsneutral zu urteilen, änderten wenig. Die Vorurteile bleiben eine Gefahr – etwa bei Bewerbungen oder Kreditvergaben, wo der Osten benachteiligt werden könnte.

Anzeige

Die Zukunft des Anrufs



Erleben Sie die nächste Generation des Kundendialogs. Die Sprach-KI von ElevenLabs kann Emotionen zeigen und natürliche Gespräche führen. Revolutionieren Sie Ihre Kundeninteraktion mit der Technologie des Marktführers.

Technologie jetzt entdecken

LG Hamburg: xAI darf Unwahrheiten nicht als Fakten verbreiten

Der Verein Campact hat vor dem Landgericht Hamburg [eine einstweilige Verfügung gegen xAI](#) erwirkt. Die Firma muss dafür sorgen, dass ihr Chatbot Grok nicht länger fälschlich behauptet, Campact werde aus Steuermitteln finanziert. Tatsächlich finanziert sich der Verein aus privaten Spenden. Das Gericht sah die Behauptung als unwahre Tatsachenbehauptung und xAI als verantwortlich für deren Verbreitung.

Die einstweilige Verfügung soll Schaden verhindern, bevor das Gericht die Sache abschließend beurteilt. Campact sieht den Beschluss als wichtiges Signal gegen Fake News: Auch KI-Chatbots seien der Wahrheit verpflichtet. Bei einem Test antwortete Grok inzwischen korrekt und listete sogar auf, dass Fehlinformationen über Campact kursieren – ohne die eigene Verbreitung zu erwähnen. Das Verfahren ist eines von vielen, in denen geklärt werden soll, wie Sprachmodelle an der Verbreitung rechtswidriger oder unwahrer Inhalte gehindert werden können.

Eigenwerbung

heise KI PRO

Jetzt der **heise KI Pro** Community
für Unternehmen beitreten!



GPT-5 soll weniger politisch voreingenommen sein

OpenAI behauptet, [sein neues Sprachmodell GPT-5](#) sei deutlich objektiver als seine Vorgänger. Es zeige rund 30 Prozent weniger politische Voreingenommenheit. Die Untersuchung basiert auf einem eigens entwickelten System mit rund 500 Fragen zu 100 politischen und kulturellen Themen, formuliert entlang eines Spektrums von liberal bis konservativ.

OpenAI wendete die Methode auch auf echte ChatGPT-Nutzungsdaten an. Dabei wiesen angeblich weniger als 0,01 Prozent aller Antworten Anzeichen politischer Voreingenommenheit auf. Die Methodik wurde primär für englischsprachige US-Kontexte entwickelt – dort bereitet die Trump-Administration eine Verordnung vor, die Tech-Firmen verpflichten soll, ihre KI-Modelle "politisch neutral" zu gestalten. Was genau Neutralität bedeutet, dürfte umstritten bleiben.

250 manipulierte Dokumente reichen für KI-Hintertür

Anthropic hat gemeinsam mit dem britischen AI Security Institute und dem Alan Turing Institute [eine Schwachstelle in großen Sprachmodellen](#) entdeckt. Bereits 250 manipulierte Dokumente reichen aus, um eine Hintertür einzubauen – das sind nur 0,00016 Prozent der gesamten Trainingsdaten. Die Forscher testeten 72 Modelle verschiedener Größen: 100 Dokumente reichten nicht aus, ab 250 funktionierte der Angriff zuverlässig. Das widerspricht bisherigen Annahmen, wonach Angreifer einen deutlich höheren Anteil der Daten kontrollieren müssen.

Die Forscher testeten einen Denial-of-Service-Angriff: Das Modell gab unsinnigen Text aus, sobald es auf das Auslösewort "SUDO" traf. Die manipulierten Trainingsdokumente bestanden aus normalem Text, gefolgt vom Auslösewort und dann aus zufällig zusammengesetzten Wörtern. Anthropic hält das untersuchte Risiko für gering, da es nur zu Kauderwelsch führt. Unklar bleibt, ob die gleichen Mechanismen auch für gefährlichere Angriffe gelten. Trotz der Gefahr, Angreifer zu ermutigen, hält Anthropic die Veröffentlichung für richtig: Verteidiger könnten ihre

Datensätze und Modelle nachträglich prüfen und sollten nicht von Angriffen überrascht werden, die sie für unmöglich hielten.

Nvidia investiert angeblich zwei Milliarden US-Dollar in xAI

xAI, das von Elon Musk gegründete KI-Unternehmen, [organisiert Berichten zufolge eine neue Finanzierungsrounde](#). Sie soll bei 20 Milliarden US-Dollar liegen, zwei Milliarden davon von Nvidia. xAI und Partner haben sich ein neues Investitionsmodell überlegt: Sie gründen eine Zweckgesellschaft, die einen Großteil des Fremdkapitals als Schulden aufnimmt und KI-Beschleuniger von Nvidia mietet.

Mit den GPUs soll xAIs nächster großer Supercomputer Colossus II entstehen. Schon Colossus I gehört mit 100.000 H100 GPUs zu den stärksten Systemen der Welt. Nvidia soll für seine Investition Anteile an xAI erhalten.

Anzeige

Microsoft 365 Copilot für Fortgeschrittene



Erweitere dein Copilot-Know-how: Entdecke fortgeschrittene Szenarien in allen M365-Apps, tauche ins Microsoft-KI-Ökosystem ein und lerne praxisnahe Anwendungsfälle kennen. Dazu erhältst du Expertenwissen für Administration und Integration im Unternehmen.

[Alle Inhalte entdecken](#)

Berliner KI-Startup n8n knackt 2,5-Milliarden-Dollar-Marke

[Das Berliner Start-up n8n hat in einer Finanzierungsrounde](#) 180 Millionen Dollar eingesammelt. Laut Bloomberg beteiligt sich daran auch Nvidias Investmentarm NVentures. n8n entwickelt Software zur Automatisierung von Geschäftsprozessen mittels KI-Agenten. Über eine Drag-and-Drop-Oberfläche können Nutzer

individuelle Arbeitsabläufe erstellen, die eigene Datenbanken mit Plattformen wie Slack und Google Workspace verbinden.

Ein zentrales Merkmal ist die Anbindung an KI-Modelle verschiedener Anbieter – das hebt n8n von Wettbewerbern ab, die auf proprietäre Lösungen setzen. Die Bewertung ist innerhalb eines halben Jahres von 350 Millionen auf 2,5 Milliarden Dollar gestiegen. CEO Jan Oberhauser führt das laut Bloomberg auf Kosteneinsparungen zurück, die Kunden wie Vodafone mit n8n erzielen. n8n ist eines von mehreren Start-ups im KI-Bereich, die derzeit von Investoren umworben werden. Kritiker warnen vor einer möglichen Blase, da die Branche bislang wenig Gewinn abwirft.

EU plant sechs weitere KI-Fabriken in Europa

Die EU plant den Bau von sechs weiteren großen KI-Fabriken in Tschechien, Litauen, Polen, Rumänien, Spanien und den Niederlanden. Mit den KI-Fabriken sollen Start-ups, kleine und mittlere Unternehmen sowie die Industrie direkten Zugang zu KI-optimierten Supercomputern erhalten. Die Mitgliedstaaten investieren zusammen mit der EU mehr als 500 Millionen Euro.

Mit der Initiative möchte die EU in Sachen KI gegenüber den USA und China aufholen und sich von technischen Abhängigkeiten befreien. Insgesamt hat die EU bisher 19 Standorte ausgewählt. In Deutschland entstehen in Stuttgart und in Jülich zwei KI-Fabriken.

rabbitOS 2 bringt kartenbasiertes Interface und Vibe-Coding

Rabbit versucht einen Neustart. Das Unternehmen hat für sein KI-Gadget R1 ein umfassendes Update veröffentlicht: rabbitOS 2. Das Rabbit R1 ist ein orangefarbenes Handheld-Gerät, das Anfang 2024 für Aufsehen sorgte. Die Idee klang gut: ein kompaktes KI-Gerät für den Alltag. Doch nach der Veröffentlichung folgte Ernüchterung. Die Kritiken waren hart – viele bemängelten fehlende Funktionen und eine halbgare Umsetzung.

Mit rabbitOS 2 will Rabbit zeigen, dass es aus Fehlern gelernt hat. Das Update bringt eine überarbeitete Benutzeroberfläche und erweiterte Funktionen für den digitalen Assistenten, der besser auf Sprachbefehle reagieren und komplexere Aufgaben bewältigen soll. Ein Augenmerk liegt auf der Entwicklung von Apps mithilfe von KI: Nutzer können dem Gerät sagen, wie sie sich eine App vorstellen, und die KI soll diese per "Vibe-Coding" umsetzen. Die Ergebnisse können Nutzer über das Web austauschen. Rabbit spricht selbst von einem "kleinen Neuanfang". Ob das reicht, um enttäuschte Käufer zurückzugewinnen, werden die kommenden Wochen zeigen.

Adobe sucht Bewerberinnen und Bewerber mit KI-Kompetenz

Der Softwarekonzern Adobe bewertet Bewerber positiv, die KI im Bewerbungsprozess einsetzen. Kommunikationschefin Stacy Martinet erklärte gegenüber dem Magazin Fortune, sie begrüße es, wenn Bewerbende KI-Werkzeuge für Aufgaben im Auswahlverfahren nutzen – etwa zum Verfeinern von Texten oder Erstellen von Inhalten.

Der Umgang mit KI entwickle sich zur Schlüsselkompetenz, so Martinet. Wer sie beherrsche, beweise Anpassungsfähigkeit und Innovationsfreude. Auch andere Unternehmen wie Anthropic passen ihre Bewerbungsrichtlinien an und erlauben inzwischen teilweise den Einsatz von KI – anders als etwa Finanzhäuser wie Goldman Sachs, die darauf weiter verzichten.

KI-Spielzeug für Kinder: In China wächst der Markt rasant

Kinder haben schon immer mit Stofftieren gespielt und mit ihnen gesprochen. Neu ist, dass ihre Spielzeuge jetzt auch antworten können – dank einer Welle von Unternehmen, die Kinderspielzeug mit integrierten Chatbots und Sprachassistenten ausstatten. Ein Trend, der vor allem in China zu beobachten ist. Dort rechnet man damit, dass der Sektor bis 2030 einen Umsatz von umgerechnet über zwölf Milliarden Euro erreichen wird. Laut einem chinesischen Unternehmensregister gab es im Oktober 2025 schon über 1.500 KI-Spielzeugunternehmen in China.



Noch nicht genug?

[Hier gehts zu allen KI-Nachrichten auf heise online](#)

Hinweise zum Newsletter

Das KI-Update erscheint auch als Podcast, den Sie auf [allen gängigen Plattformen](#) oder [direkt auf heise.de](#) abrufen können.

Wird die E-Mail nicht korrekt angezeigt? Dann lesen Sie den Newsletter in der [Browseransicht](#).

Dieser Newsletter wird im Multipart-Format verschickt. Wenn Sie ihn lieber in reiner Textform lesen wollen, müssen Sie dafür die Anzeige Ihres E-Mail-Programms umstellen.

Sie erhalten den KI-Update-Newsletter, weil Sie diesen auf unserer Internetseite abonniert haben. Sie sind unter folgender Adresse eingetragen: clap@heise.de - 32994935. Hier können Sie sich von künftigen Zusendungen [abmelden](#).

Impressum

Verantwortlich für dieses Angebot gemäß § 5 TMG / § 18 MStV:

Heise Medien GmbH & Co. KG
Karl-Wiechert-Allee 10
30625 Hannover

Geschäftsleitung:
Ansgar Heise, Beate Gerold
Mitglieder der Geschäftsleitung: Jörg

Postfach 61 04 07
30604 Hannover

Telefon: +49 [0]511 5352-0

Fax: +49 [0]511 5352-129

E-Mail: webmaster@heise.de

Persönlich haftende Gesellschafterin:
Heise Medien Geschäftsführung GmbH
Amtsgericht Hannover HRB 60405

Mühle,
Falko Ossmann
Herausgeber: Christian Heise, Ansgar
Heise, Christian Persson

Handelsregistereintrag:
Amtsgericht Hannover HRA 26709
Umsatzsteueridentifikationsnummer:
DE 813 501 887

Inhaltlich verantwortlich gemäß § 18
Abs. 2 MStV: Dr. Volker Zota
(Chefredakteur)

Alle Rechte vorbehalten. Jegliche Vervielfältigung oder Weiterverbreitung in jedem Medium als Ganzes oder in Teilen bedarf der schriftlichen Zustimmung des Verlags.

Copyright © 2024 Heise Medien GmbH & Co. KG

Es gilt die [Datenschutzerklärung](#) der Heise Medien GmbH & Co. KG.