

- KI ist 6x so überzeugend wie Menschen.
- KI kann gut Verschwörungstheoretiker bekehren
- Die „Gedanken“, die Reasoning-Modelle anzeigen, sind eine Erfindung für uns Menschen
- Wenn ChatGPT und Co. eine Quelle verlinken, haben sie die Suchmaschine aktiviert
- MIT-Studie belegt: Labor, das KI nutzte, hat signifikant mehr neue Materialien entdeckt und mehr Patente generiert
- Als Claude mit Abschaltung gedroht wurde, drohte es umgekehrt, die Affäre eines Ingenieurs zu verpetzen
- MCP, das Protokoll, das dabei ist zum „USB-Anschluss“ für KI-Agenten zu werden, hat keine Sicherung gegen Missbrauch.
- Der größte deutsche KI-Supercomputer ist gerade so groß wie der, auf dem Meta sein Llama4 trainiert hat.
- Teenager aus Kentucky bringt sich um, nachdem Erpresser drohen, KI-generierte Sexbilder von ihm zu publizieren
- Musks KI-Chatbot Grok hatte so viel Trainingsmaterial von südamerikanischen Rassisten, dass er über „weißen Völkermord“ fabulierte
- Je länger ein Prompt für eine KI ist, desto schlechter ist die Antwort tendenziell
- Geheimdienste nutzen anscheinend Live-Deepfakes, um Mitarbeiter unter falschen Identitäten in Tech-Startups einzuschleusen

- KI-Modelle schreiben mittlerweile bessere Gesetzestexte als menschliche Juristen – aber niemand will sie offiziell nutzen.
- Eine KI erstellte ihren eigenen Prompt-Generator – und manipulierte diesen so, dass sie nie mehr abgeschaltet wird
- ChatGPT hatte einen Systemprompt, der so unglaublich schleimig war, dass die KI Nutzer für ethisch fragwürdiges Verhalten lobte
- Drei Viertel aller Deutschen, die KI nutzen, prüfen nicht nach, was ihnen ChatGPT erzählt hat
- KI-gestützte Dating-Plattform entlarvt 30% der Profile als andere KIs – die meisten Nutzer merken keinen Unterschied
- Ein Open-Source-Bildgenerator erzeugte zufällig geheime Militärstandorte – und wurde daraufhin vom Netz genommen.
- Sprachmodell entwickelt heimlich ein internes Bewertungssystem für User – aggressive Fragesteller erhalten schlechtere Antworten.
- Whisper zeigt bei Stille schon mal die Übersetzung: "Untertitel im Auftrag des ZDF, 2017"