



Start



Ihr Netzwerk



Jobs



Nachrichten



Mitteilungen



Sie



Produkte

Jetzt Pre

[Artikel bearbeiten](#)[Statistik anzeigen](#)[Beitrag anzeigen](#)

Knipunkt

331 Abonnent:innen

✓ Abonniert



KNIEPUNKT 034: ungeprüfte KI-Gewissheit

**André Knie** ✓

Dr. André Knie | Physiker & Zukunftsgestalter | Baut Brücken zwischen Technologie & Kultur für mutige Organisationen |...



19. April 2026

Fast die Hälfte der Menschen in Deutschland glaubt inzwischen, KI-generierte Inhalte erkennen zu können. Nur 28 Prozent suchen aber nach Unstimmigkeiten im Bild, und nur 19 Prozent prüfen aktiv die Quelle. Das Vertrauen in die eigene Urteilskraft wächst also schneller als die Bereitschaft zur Überprüfung. Wir schauen auf ein Bild, lesen einen Absatz, nicken kurz und halten uns schon für aufgeklärt. Wird Medienkompetenz zum Gefühl?

Das bleibt leider nicht beim privaten Scrollen, denn auch im Entwicklerkopf wird es allmählich eng. Das frisch veröffentlichte Thoughtworks Technology Radar warnt davor, dass sogenannte „kognitive Schulden“ rasant explodieren. Durch uferloses „Vibe Coding“ lassen Entwickler ihre KI-Agenten Code schneller tippen, als sie ihn selbst noch verstehen können, und verlieren das mentale Modell ihrer eigenen Software. In Gerichtsverfahren tauchen auch immer öfter halluzinierte Zitate, erfundene Fundstellen und KI-Fehler auf. Ausgerechnet dort also, wo Wirklichkeit belastbar sein müsste, wird sie plötzlich autocomplete-fähig. In New York liegt deshalb bereits ein Gesetzentwurf vor, der die Offenlegung generativer KI in Zivilverfahren verlangt. Beides zeigt deutlich, dass wir erst Maschinen nutzen, um schneller sicher zu wirken, und kurz darauf brauchen wir neue Regeln, damit Menschen wieder Verantwortung übernehmen.

Noch deutlicher wird es bei der Frage nach Macht. Während wir hierzulande nicht einmal die Quelle prüfen, verhandelt Anthropic die Sicherheitszukunft des Internets in einem exklusiven amerikanischen Zirkel. Bei Project Glasswing sitzen AWS, Apple, Google, Microsoft, Nvidia, JPMorganChase und andere Schwergewichte am Tisch, aber kein einziger europäischer Akteur. Datensicherheit wird also dort gemacht, wo auch

Modelle, Chips und Plattformen sitzen: in Amerika. Irgendwie passend, dass Anthropic gerade Geistliche konsultiert, um über Claudes moralische Entwicklung und die Frage zu sprechen, ob ein Modell womöglich als "Kind Gottes" gelten könnte. Dieselbe Branche, die gerade Marktzugang, Sicherheitsarchitektur und geopolitische Macht neu sortiert, diskutiert vorsorglich schon die metaphysische Würde ihrer Produkte.

Es gibt aber selbst im Mutterland des Maßlosen noch Hoffnung. Maine stoppt neue Rechenzentren über 20 Megawatt vorläufig bis zum 1. November 2027. Der Fortschritt endet hier nicht im Prompt-Fenster, sondern bei Stromtrassen, Genehmigungen und Akzeptanz. Vielleicht ist das die ehrlichste Beschreibung dieses KI-Moments: Wir überschätzen unsere Kontrolle genau dann, wenn andere längst über Zugang, Sicherheit und Infrastruktur entscheiden. Ich möchte heute aber positiv enden, denn KI kann auch grandios inklusiv sein. An der Binghamton University lotste in einer Studie ein KI-Blindenführhund dank GPT-4 Sehbehinderte durch den Alltag. Für einen Bruchteil der Kosten eines echten Hundes kann man sogar mit ihm sprechen. KI ist da und sie wird uns immens bereichern, wenn wir sie sinnvoll und klug einsetzen.


Zum Weiterlesen



- BSI/ProPK: Nur ein Fünftel überprüft die Quelle bei KI-Inhalten <https://www.polizei-beratung.de/presse/detailseite/cymon-ki-betrug-im-netz-nur-ein-fuenftel-ueberprueft-die-quelle/>
- Anthropic: Project Glasswing und die Launch-Partner von Claude Mythos Preview <https://www.anthropic.com/project/glasswing>
- ZDFheute: BSI zeigt sich besorgt über Claude Mythos, hat das Tool aber noch nicht selbst testen können <https://www.zdfheute.de/politik/deutschland/ki-anthropic-claude-mythos-schwachstellen-software-bsi-100.html>
- Washington Post: Anthropic lud christliche Führungspersonen ein, um über Claudes moralische Zukunft zu sprechen <https://www.washingtonpost.com/technology/2026/04/11/anthropic-christians-claude-morals/>
- Damien Charlotin: Datenbank zu Gerichtsentscheidungen mit KI-Halluzinationen <https://www.damiencharlotin.com/hallucinations/>
- New York State Senate Bill S9794: Offenlegungspflicht für generative KI in Zivilverfahren <https://www.nysenate.gov/legislation/bills/2025/S9794>
- Maine Legislature, Committee Amendment zu LD 307: 20-MW-Grenze und Laufzeit bis 1. November 2027 <https://legislature.maine.gov/legis/bills/getPDF.asp?item=2&paper=HP0207&snum=132>
- Thoughtworks Radar (Kognitive Schulden): Thoughtworks Pressemitteilung (April 2026) <https://www.tradingview.com/news/eqs:3f1c2ec8c094b:0-as-ai-accelerates-software-complexity-thoughtworks-technology-radar-urges-a-return-to-engineering-fundamentals/>
- KI-Blindenführhund: New Atlas Berichterstattung - <https://newatlas.com/robotics/robot-guide-dog/>

Kommentare



16

 Gefällt mir Komentieren Teilen

Noch keine Kommentare.
Gehören Sie zu den Ersten, die kommentieren.
[Unterhaltung beginnen](#)

**Kniepunkt**

Kolumne über die Widersprüche der digitalen Gegenwart.

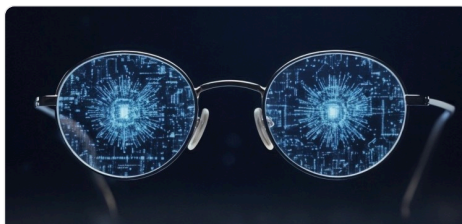


Olaf, Alexander und 312 weitere Kontakte sind Abonnent:innen

331 Abonnenten

[✓ Abonniert](#)**Weitere Artikel für Sie****Die KI-Transformation und der Wilde Westen**

KI ohne Filter

  13 · 6 Kommentare**Autonomie, Machtkämpfe und die neue Realität: Der KI-Wendepunkt ist da**

KI-Wochenrückblick

  6 · 1 Repost**Energiekrise war gestern: Warum Europa digitale Abhängigkeit die nächste große Herausforderung ist...**

Marc Antoine Schaut

 4