



Startseite > Wissenschaft > Mensch > Meinung > Künstliche Intelligenz: Es rollt ein Tsunami auf uns zu - K

### Machtkampf der KI

## ***Es rollt ein Tsunami auf uns zu***

Eine Kolumne von Christian Stöcker

Bilder, Videos, Audios – all dies kann immer besser mit künstlicher Intelligenz erzeugt werden. Viel spricht dafür, dass Open-Source-Systeme die Branchengiganten bald in Bedrängnis bringen. Wie ist das einzuschätzen?

07.05.2023, 17.41 Uhr

Artikel zum Hören • 10 Min





Beleuchtete Tastatur: »Viel der neuen Ideen kommen von ganz normalen Leuten«  
Foto: Sebastian Gollnow / dpa

*»KI-Zeug war bis jetzt schon seltsam. Jetzt wird es noch sehr viel seltsamer.«*

Softwareentwickler Simon Willison [in einem Blogbeitrag](#) [↗](#)  
vom 11. März 2023

*»Es ist schwer vorstellbar, wie man böswillige Akteure davon abhalten will, damit schlimme Dinge anzustellen.«*

KI-Pionier Geoffrey Hinton in der »New York Times« vom 1. Mai

Geoffrey Hinton ist Psychologe, wie viele Pioniere dessen, was heute maschinelles Lernen oder künstliche Intelligenz genannt wird. Gemeinsam mit seinem bereits verstorbenen Kollegen David Rumelhart veröffentlichte Hinton 1986 [in »Nature« einen Artikel](#) [↗](#), der bis heute über 32.000-mal in anderen Arbeiten zitiert worden ist: »Learning representations by back-propagating errors.« Der darin beschriebene Mechanismus der iterativen Fehlerkorrektur in neuronalen Netzen ist einer der Grundpfeiler der Technologie, die jetzt, fast 40 Jahre später, die Welt verändert.

ANZEIGE

Die »New York Times« nannte Hinton in der [Browserzeile zu einem viel zitierten Artikel](#) [↗](#) drolligerweise »Google-Chatbot-Ingenieur«. Das ist etwa so, als nenne man Carl Benz »Automechaniker«. In dem Text erklärte Hinton, dass er bei seinem letzten Arbeitgeber Google gekündigt habe, um lauter vor den Gefahren der Technologie warnen zu können, die er selbst mit erschaffen hat.

**Christian Stöcker** >

Foto: SPIEGEL ONLINE

Jahrgang 1973, ist Kognitionspsychologe und seit Herbst 2016 Professor an der Hochschule für Angewandte Wissenschaften Hamburg (HAW). Dort verantwortet er den Studiengang Digitale

Kommunikation. Vorher leitete er das Ressort Netzwelt bei SPIEGEL ONLINE.

## »Das ist Furcht einflößend«

Das Veränderungstempo im Bereich maschinelles Lernen hat in den vergangenen etwa fünfzehn Jahren kontinuierlich zugenommen, und zwar exponentiell. Wir haben es also mit einer beschleunigten Entwicklung zu tun. Und mit beschleunigten Entwicklungen, darum ging es in dieser Kolumne schon öfter, hat der menschliche Geist so seine Schwierigkeiten – wir sind sehr schlecht darin, uns vorzustellen, wie schnell sich Dinge verändern können, wenn etwas immer schneller geht. Jetzt sagte Hinton der »New York Times«: »Sehen Sie sich an, wie es vor fünf Jahren war, und wie es jetzt ist. Nehmen Sie diesen Unterschied und extrapolieren Sie in die Zukunft. Das ist Furcht einflößend.«

ANZEIGE

Hintons fürs erste wichtigste Sorge ist, dass der Durchschnittsmensch bald »nicht mehr in der Lage sein wird, zu wissen, was wahr ist«. Tatsächlich ist dieser Punkt bereits erreicht: Die Bilder, die aktuelle, weiterhin auf neuronalen Netzen basierende Bildgeneratoren wie Stable Diffusion oder Midjourney auf Basis verbaler Kommandos herstellen, sind schon jetzt praktisch nicht mehr von menschengemachten Illustrationen oder realen Fotos zu unterscheiden.

Für Menschen, die sich professionell mit Desinformation und deren Abwehr beschäftigen, wachsen auch die Probleme gerade exponentiell. Noch vor ein paar Monaten waren KI-generierte Fotos an bestimmten Merkmalen wie seltsamen Händen, merkwürdigen Zähnen oder mit dem Hintergrund verschwimmenden Frisuren recht einfach zu erkennen. Das ist vorbei.

ANZEIGE

## **Gestern bizarr, morgen überzeugend**

Als Nächstes ist das Thema Bewegtbild dran. Da ist die Lage im Moment so ähnlich wie vor ein paar Monaten bei KI-generierten Bildern: Von neuronalen Netzen erzeugte Videos sind zum Teil so bizarr [↗](#), dass sie als Unterhaltung herumgereicht werden. Aber auch das wird sich ändern, und zwar vermutlich wieder schneller als erwartet. Meine Prognose wäre mittlerweile, dass sich der Zeitraum jetzt in Monaten, nicht mehr in Jahren bemisst (aber ich kann mich natürlich irren).

Was Hinton beklagt, ist, dass das KI-Wettrüsten derzeit völlig unkontrolliert abzulaufen scheint. Es geht dabei zunächst um Hintons bisherigen Arbeitgeber Google, im Bereich KI auch dank kluger Zukäufe jahrelang anscheinend uneinholbar führend, und den Schöpfer von ChatGPT, OpenAI und dessen Finanzier Microsoft.



**»Wir haben keinen Burggraben«**

Es gibt aber, das ist seit Ende dieser Woche ein zentrales Diskussionsthema der KI-Besessenen, noch einen weiteren mächtigen Mitspieler, oder besser: sehr viele. Und die werden noch schwerer zu regulieren sein als Giganten wie Google, Microsoft, Facebook oder Amazon.

Anlass für die erhitzte Diskussion ist ein angeblich von einem Google-Entwickler stammender Blogeintrag, der angeblich über eine Chat-Plattform geleakt und von einem auf Halbleiterentwicklung spezialisierten Newsletter veröffentlicht wurde. Die Überschrift des [faszinierenden Textes](#) lautet: »Wir haben keinen Burggraben, und OpenAI hat auch keinen.«

ANZEIGE

	<b>Autokredit direkt!</b> Mit dem Kfz-Kredit der Postbank haben Sie freie Fahrt	
	<a href="https://www.postbank.de">postbank.de</a>	
	<b>Internet für Zuhause.</b> O2 my Home M - DSL mit bis zu 100 MBit/s schon ab 29,99€* im Monat.	
	<a href="https://www.o2online.de">o2online.de</a>	
	<b>Smart &amp; Safe Home</b> Finde jetzt Sicherheitstechnik für dein Zuhause bei Conrad!	
	<a href="https://www.conrad.de">conrad.de</a>	

Der Text enthält massenweise zitable Sätze wie diesen hier: »Während wir uns [mit OpenAI] herumgestritten haben, hat eine dritte Partei still und leise unser Mittagessen verspeist.«

ANZEIGE



Christian Stöcker  
**Die Große Beschleunigung**

Verlag: Pantheon  
Seitenzahl: 384

Für 16,00 € kaufen ⓘ

Bei Amazon bestellen 

Bei Thalia bestellen 

Bei Yourbook bestellen 

Produktbesprechungen erfolgen rein redaktionell und unabhängig. Über die sogenannten Affiliate-Links oben erhalten wir beim Kauf in der Regel eine Provision vom Händler. [Mehr Informationen dazu hier](#)

## Offen vs. geschlossen, Licht und Schatten

Der Autor führt darin, unter Verweis auf zahlreiche Belege, eine bei Google eigentlich ursprünglich in die Unternehmenskultur eingebaute Weisheit aus: Quelloffene Software, Open-Source-Software, ist [auf lange Sicht](#) besser als proprietäre, geschlossene Systeme. Wer seinen Code in die freie Wildbahn entlässt, sodass andere Entwicklerinnen und Entwickler damit herumspielen, ihn modifizieren, erweitern, anpassen, verbessern können, der gewinnt. Deshalb ist Android, im Gegensatz zu Apples iOS, ein quelloffenes Betriebssystem. Der weltweite Marktanteil von Android-Systemen am Smartphone-Markt liegt bei [über 70 Prozent](#) [↗](#). Auch weite Teile der Infrastruktur des Internets basieren auf quelloffener Software.

Ein quelloffenes System für das Generieren von Bildern mit KI gibt es bereits. Es heißt Stable Diffusion und ist nicht so benutzerfreundlich und komfortabel wie Midjourney, aber dafür hat es andere Vorzüge: Wer will, kann es modifizieren, erweitern, verbessern, damit herumspielen. Es gibt mittlerweile einen [eigenen Markt](#) [↗](#) für sogenannte Modelle, die auf Stable Diffusion basieren. Viele davon sind spezialisiert auf das Erzeugen von Bildern im Manga- oder Anime-Stil, andere auf fotorealistische Bilder, manche auf Bilder einzelner Schauspielerinnen. Eines ist dafür gemacht, Frauen auf allen Vieren zu erzeugen, die zum Betrachter hochschauen.

## Natürlich gibt es längst AI-Porn

Der Quellcode der jüngeren Version von Stable Diffusion ist mit gewissen Einschränkungen versehen, die verhindern sollen, dass jemand Pornografie damit herstellt, beim kommerziellen Konkurrenten Midjourney kann man »Prompts«, also Anweisung zur Bildgenerierung, die mit Sex zu tun haben, gar nicht benutzen. Aber natürlich gibt es trotzdem längst »AI-Porn«-Generatoren.

Es rollt ein Tsunami aus KI-generierten Bildern und bald auch Videos auf uns zu. Ein Tsunami auch aus schwer zu erkennender Desinformation, aus KI-generierten Websites, die KI-generierten Schrott enthalten, und Werbeeinnahmen generieren sollen – [es gibt sie längst](#) [↗](#). Und ich würde fast wetten, dass es den Enkeltrick-Roboter, den ich [in dieser Kolumne vor fünf Jahren](#) schon einmal theoretisch skizziert habe, in wenigen Monaten Realität sein wird. Aber all das ist immer noch erst der Anfang.

Der Autor des angeblich aus dem Inneren von Googles KI-Team stammenden Brandbriefes ist – nicht als Einziger – der Meinung, dass die Open-Source-Revolution im Bereich lernender Maschinen gerade erst begonnen hat. Und, dass sie die Branchengiganten in Bedrängnis bringen wird.

## KI-Chatbots auf dem heimischen Rechner

Es gibt tatsächlich rasante Fortschritte in diesem Bereich: Die Zeiten, da man für eine Sprachproduktions-KI eine eigene Serverfarm anmieten musste, sind längst vorbei. Es gibt jetzt KI-Chatbots, die auf dem heimischen Rechner und sogar auf Telefonen laufen. Es muss auch nicht jedes Modell von Grund auf neu trainiert werden: Mit einer schon 2021 publizierten Technik namens [Low-Rank Adaptation](#) [↗](#) zum Beispiel können bestehende Modelle mit deutlich geringerem Aufwand verfeinert werden. Man kann leistungsfähige Modelle jetzt problemlos für ein paar Hundert Dollar trainieren.

Und ein solches bestehendes Modell, ein gewaltiges, steht seit Ende Februar öffentlich zur Verfügung: Am 24. Februar publizierten Entwickler von Facebook ein Modell [namens LLaMA](#) [↗](#), allerdings unter Zurückhaltung bestimmter Aspekte. Dieser Zustand änderte sich wenige Tage später, als jemand auch diese zurückgehaltenen Informationen, genauer gesagt: das gesamte Modell, über eine Filesharing-Plattform [leakte](#) [↗](#). Kurz darauf brachte ein bulgarischer Entwickler ein auf LLaMA basierendes Sprachmodell auf einem normalen MacBook [zum Laufen](#) [↗](#).

LLaMA darf für nicht kommerzielle Zwecke ganz legal von jedermann benutzt werden. Alle arbeiten jetzt damit, von [Stanford](#) [↗](#) über [Berkeley](#) [↗](#) bis in die Vereinigten Arabischen Emirate. Und der Abstand zwischen den Open-Source-Systemen und ChatGPT [schwindet rasant](#) [↗](#).

Diese Entwicklung wird auch viel Gutes mit sich bringen, denn man kann mit lernenden Maschinen eine Menge Dinge anstellen, die die Menschheit wirklich weiterbringen. Aber wie immer ist Open Source ein zweischneidiges Schwert: So wie quelloffene Softwaresysteme auch den Weg für Viren und andere Probleme geöffnet haben, wird quelloffene KI extrem Missbrauchs-anfällig sein. Und sehr mächtig.

### **»Jetzt sind diese Kontrollmechanismen verschwunden«**

Der in Fachkreisen seit Jahrzehnten bekannte Entwickler Simon Willison erklärte das LLaMA-Leak schon im März zum »Stable-Diffusion-Moment« der Sprach-KIs: »Bis zu diesem Moment gab es eine dünne Abwehrschicht, weil Firmen wie OpenAI eine eingeschränkte Fähigkeit hatten, zu kontrollieren, wie Leute mit diesen Modellen interagieren. Jetzt, wo wir diese Modelle auf unserer eigenen Hardware laufen lassen können, sind selbst diese Kontrollmechanismen verschwunden.«

Und es sind schon längst nicht mehr nur universitäre Teams. In den gut zwei Monaten seit dem LLaMA-Leak hat sich die Welt der lernenden Maschinen unwiderruflich und dramatisch verändert.

Der namenlose, mutmaßliche Google-Mitarbeiter (es ist für die Relevanz des Textes übrigens fast irrelevant, wo der Autor wirklich arbeitet) formuliert das in seinem geleakten Text so: »Viele der neuen Ideen kommen von ganz normalen Leuten. Die Eintrittshürde dafür, ein Modell zu trainieren und damit zu experimentieren, ist gefallen – von einer großen Forschungseinrichtung als Voraussetzung hin zu einer Person, einem Abend und einem kräftigen Laptop.«

Die KI-Explosion hat gerade erst begonnen. Sie verläuft spätestens ab jetzt weitgehend unkontrolliert. Der Geist ist aus der Flasche, und es wird unmöglich sein, ihn wieder hineinzubekommen. **S**

Diskutieren Sie mit

[Feedback](#)

ANZEIGE



**Wall Street Premier**  
Diese KI-Aktie wird die  
die  
Lebensmittelbranc...

ANZEIGE



**AQON Pure Ent...**  
Kalk im Wasser:  
Zwei Brüder  
revolutionieren...

ANZEIGE



**STERNGLAS**  
Superflache  
Automatikuhr –  
Asthet Edition...

**Aktuell in diesem  
Ressort  
Suche nach...**

