

01. März 2026, von Michael Schöfer

Die Menschheit steht an der Schwelle zur Selbstauslöschung

Mancher, der die Vereinigten Staaten früher als Hort des Kapitalismus bezeichnete (oder geißelte - je nach Standpunkt), reibt sich gerade verwundert die Augen: Machen sich die USA auf den Weg in eine Kommandowirtschaft? Man fühlt sich fast ins Reich Putins versetzt, in dem Privatunternehmen die Anordnungen der Regierung zu befolgen haben. Und wenn sie sich dagegen wehren, hat das für die Firmen und deren Besitzer schlimme Konsequenzen.

Das US-Kriegsministerium würde gerne die Künstliche Intelligenz (KI) namens "Claude" der Software-Firma Anthropic verwenden, doch die sträubt sich, weil das Pentagon dabei rote Linien überschreiten will. Das Unternehmen lehnt "den Einsatz von Claude für massenhafte Inlandsüberwachung sowie für vollständig autonome Waffen ab", erläutert Anthropic-Gründer Dario Amodei. [1] Die US-Regierung ist sauer und droht dem Unternehmen daraufhin mit Sanktionen. Holla die Waldfee, war da nicht mal was mit den Prinzipien der freien Marktwirtschaft? Wo hat man denn bisher das hohe Lied des Eigentums am inbrünstigsten gesungen? Natürlich in den USA. Für MAGA-Fans ist eine allgemeine Krankenversicherung "sozialistisch" und wird strikt abgelehnt, aber Trumps Kommandowirtschaft ist plötzlich kompatibel mit dem Kapitalismus? Dazu muss man seine Gehirnwindungen schon arg verbiegen.

Wenn sich Anthropic gegen eine Massenüberwachung wehrt, meint das Unternehmen damit wahrscheinlich eine totale Überwachung, denn die Amerikaner werden ja bereits seit langem intensiv überwacht. Oder sind etwa Edward Snowdens Enthüllungen schon vergessen? Was Anthropic offenkundig verhindern will, ist, der Regierung das Werkzeug für eine dystopische Überwachung à la China zu liefern. Wenn die Trump-Administration trotzdem darauf besteht, sollte das für die Amerikaner Grund zu größter Besorgnis sein. Im "Land der Freien" (The Star-Spangled Banner) ist die Freiheit massiv bedroht.

Anthropics Ablehnung von autonomen Waffensystemen, die vollständig durch eine KI gesteuert werden, hat zweifellos eine fundamentale Bedeutung. Anscheinend plant das Pentagon nämlich genau das, anders macht der Druck auf die Software-Firma ja keinen Sinn. Man muss jetzt gar nicht "Skynet" aus den Terminator-Filmen bemühen, um im Nacken ein Frösteln zu spüren. Der österreichische Standard berichtete jüngst über ein Experiment am Londoner King's College. Wissenschaftler haben drei KIs (GPT 5.2, Claude Sonnet 4, Gemini 3 Flash) versuchsweise in eine Kriegssituation versetzt und ihnen dabei die Hoheit über die Waffensysteme überlassen. Erschreckendes Ergebnis: In 95 Prozent der simulierten Spiele überschritten die KI-Modelle die Schwelle zur Anwendung von Nuklearwaffen. [2]

Sachkundige Menschen sind sich bewusst, welches Risiko damit verbunden ist. Hat man Atomwaffen einmal eingesetzt, ist der Weg zum allgemeinen Atomkrieg und damit zur Vernichtung der Menschheit nur noch kurz. Wer die nukleare Eskalationsspirale einmal in Gang setzt, kann sie vermutlich nicht mehr anhalten. Weil überdies die "gegenseitige gesicherte Zerstörung" (**mutual assured destruction** = MAD = verrückt) nach wie vor Gültigkeit besitzt, schrecken Menschen vor der Verwendung von Nuklearwaffen zurück. Das hat zumindest seit 1945 (Hiroshima, Nagasaki) leidlich funktioniert, obwohl wir manchmal ziemlich nah dran waren (z.B. Kubakrise 1962).

Außerdem hätten auch begrenzte Atomkriege globale Auswirkungen: Schon ein pakistanisch-indischer Atomkrieg, in dem Pakistan 150 und Indien 100 Atomwaffen gegen die

Städte des jeweils anderen einsetzt, würde einen globalen nuklearen Winter auslösen, der die landwirtschaftlichen Erträge für längere Zeit erheblich reduziert. [3] Der Fehllarm von 1983 in der UdSSR belegt überdies, wie knapp die Zeit zur Vermeidung von Irrtümern ist. Die Welt entging damals dank der Besonnenheit von Stanislaw Petrow, dem diensthabenden Offizier der sowjetischen Luftverteidigungsstreitkräfte, um Haaresbreite der nuklearen Katastrophe. Moderne Waffensysteme (Hyperschallraketen) verkürzen die Reaktionszeit jedoch immer mehr, Fehllarme sind daher kaum noch als solche zu erkennen.

Die Selbstabschreckung von Menschen, den nuklearen Holocaust einzuleiten, fehlt den KI-Modellen anscheinend. Sie sind sich offenbar nicht über die verheerenden Konsequenzen von Nuklearwaffen im Klaren, anders ist die 95-prozentige Häufigkeit ihrer Verwendung in den Simulationen nicht zu erklären. Wer einer KI die Hoheit über Waffensysteme gibt, handelt deshalb absolut verantwortungslos. Natürlich wird die KI anfangs nicht gleich die Verfügungsgewalt über Nuklearwaffen besitzen, aber darauf läuft es am Ende hinaus, denn das ist die zwangsläufige Folge von extrem verengten Handlungsspielräumen, die die Reaktionsfähigkeit von Menschen überfordern.

Anthropic hat also vollkommen recht, wenn das Unternehmen hier eine rote Linie zieht. Ein Problembewusstsein, dass man allerdings von Pete Hegseth nicht erwarten kann, der Kriegsminister macht sich lieber Gedanken über die körperliche Fitness von angeblich zu fetten Generälen und Admirälen (Rede beim Treffen auf dem Marinestützpunkt Quantico/Virginia). Er legt großen Wert auf Kriegerethos und Männlichkeitsstandards - Denken ist nicht so sein Ding. Evolutionstechnisch gesehen ist das eine ziemlich ungünstige Kombination, denn eine Spezies, die so handelt, wird nicht lange existieren, sondern sich vielmehr rasch selbst auslöschen. Und genau da stehen wir heute: an der Schwelle zur Selbstauslöschung der Menschheit.

[1] heise.de vom 28.02.2026

[2] Standard vom 27.02.2026

[3] Rutgers Universität, New Jersey/USA, Bulletin of the Atomic Scientists 2019, PDF-Datei mit 1,7 MB