

Amodei schaltet auf Defensive

Anthropic-Chef warnt vor KI-Gefahren und fordert langsamere Entwicklung. Hinter dem Manöver stehen Sorgen um die US-Vorherrschaft im Sektor

Von Luca von Ludwig

Die KI wird uns alle umbringen – zumindest möchte Dario Amodei, dass die Öffentlichkeit das glaubt. Der Chef des für seine Modelle »Claude«, »Fable« und »Mythos« bekanntgewordenen KI-Konzerns Anthropic hat am Wochenende auf seinem Blog einen Beitrag veröffentlicht, in dem dazu aufgerufen wird, beim Ausbau der Fähigkeiten der KI-Modelle auf die Bremse zu treten. Open-AI-Gründer Sam Altman und andere branchenbekannte Stimmen gaben ihm seither recht.

Amodei begründet seine Forderung mit der »zunehmenden Fähigkeit der KI, selbst die nächste Generation der KI zu schaffen«. Gemeint ist die rapide Beschleunigung von digitalen Entwicklungszyklen, die in den vergangenen Jahren zu beobachten war. Konkreter verweist er zudem auf den [»Open-AI-Hugging-Face-Vorfall«](#), bei dem Ende Juli ein KI-Prototyp von Open AI seiner digitalen Testumgebung entkommen war und sich in die Server des ebenfalls aus der KI-Branche kommenden Unternehmens Hugging Face gehackt hatte – so zumindest die Darstellung von Open AI. Amodei meint, dieser Vorfall demonstriere die enormen Gefahren, die die neuesten und fähigsten KI-Modelle darstellen.

Der Anthropic-Konzernchef hatte auch gleich einen dreiteiligen Plan parat, der angeblichen Gefahr Herr zu werden. Erstens: Jede der Top-KI-Firmen solle unparteiische Prüfer zulassen, die darüber wachen, ob Sicherheitsstandards eingehalten werden. Zweitens: Die führenden KI-Entwickler »demokratischer Länder« sollen sich auf solche Sicherheitsstandards sowie Limits für den Ausbau der Fähigkeiten ihrer Modelle einigen. Drittens: »Die USA und andere demokratische Regierungen« sollen auf »autoritäre« Staaten einwirken, sich auf ähnliche Koordinationsmechanismen einzulassen. Den ersten dieser Punkte setze Anthropic bereits um, so Amodei. Zuspruch erntete sein Vorschlag unter anderem von Altman sowie Elon Musk, der über Space X AI ebenfalls am KI-Wettrennen beteiligt ist.

Gruselt es die Techoligarchen so sehr vor den Geistern, die sie riefen, dass ihnen selbst der freie Markt ungeheuerlich wird? Wenige Tage zuvor hatte immerhin der bei Anthropic angestellte KI-Entwickler Evan Hubinger angegeben, er schätze die Möglichkeit, dass die KI die Menschheit auslöschen könne, auf mehr als zehn Prozent ein. In der Realität dürfte es sich jedoch vielmehr um ein abermaliges Manöver zur Beschneidung der Konkurrenz handeln. Das sieht man selbst in den engeren Kreisen des Weißen Hauses so:

David Sacks, ehemaliger KI-Beauftragter der US-Regierung, wirft den Konzernen bereits seit längerem eine Vereinnahmung von Regulierungsprozessen vor, die darauf abziele, so hohe Auflagen zu schaffen, dass kleinere Konkurrenzunternehmen sie ohnehin niemals erfüllen können.

Die industriepolitische Natur von Amodeis Vorschlag wird auch weiter unten in seinem Blogbeitrag deutlich, wo es um die Konkurrenz mit der Volksrepublik China geht. Chinesische Unternehmen sind zur Zeit die wichtigste Gefahr für das US-amerikanische KI-Oligopol. Kaum überraschend fordert Amodei daher, die eigene KI-Entwicklung nur so weit zu bremsen, wie der vermeintliche technische Vorsprung zu chinesischen Konkurrenzmodellen es zulässt. Um diese Kapazitätslücke zu verteidigen, fordert er ein [Ausfuhrverbot für Hochleistungschips und Equipment zur Prozessorherstellung in die Volksrepublik](#), ferner Gegenmaßnahmen gegen angeblichen »Diebstahl« von Trainingsparametern und sogenannte Modelldestillation. Man dürfe im Umgang mit China nicht »naiv« sein, so der Oligarch.

Ein Sprecher des chinesischen Außenministeriums wies Amodeis Äußerungen am Montag zurück. »Ängste schüren, Konfrontation und brutaler Wettbewerb« seien in niemandes Interesse, es brauche statt dessen »offene, inklusive, für jeden vorteilhafte und ethische« KI-Entwicklung. Die chinesische *Global Times* fand deutlichere Worte und bezeichnete den Vorschlag des Anthropic-Chefs als »Kalter-Krieg-Playbook für den KI-Sektor«. Seine »wahre Funktion« sei es, China aus der globalen KI-Politik auszuschließen und den USA eine regulatorische Monopolstellung zu verschaffen.

<https://www.jungewelt.de/artikel/529816.ki-branche-amodei-schaltet-auf-defensive.html>