

Liebe Mitglieder des Netzwerkes der ehemalige SSAB, liebe Interessierte

Am 23.7.2025 veröffentlichte die Trump-Regierung ihren [AI Action Plan](#). Mit dieser Strategie wollen die USA das KI-Wettrennen gewinnen. Es geht um den Abbau regulatorischer Hürden und um Innovation. Keinesfalls sollen die grossen Plattformen reguliert werden. Die diesbezüglichen Vorschriften der EU zur Verhinderung von Falschinformationen gelten als Zensur. Allerdings müssen die Sprachmodelle den Werten der USA entsprechen. Bei Open-Source-Modellen wollen die USA führend sein.

Keine Regulierungen, die KI-Innovationen hindern

Mit seiner Executive Order vom Januar 2025 annullierte Trump in einer seiner ersten Amtshandlungen die bescheidenen KI-Regulierungen, die sein Vorgänger Biden im Oktober 2023 statuierte. Der [AI Action Plan](#) hält nun zusätzlich fest, dass Bundesstaaten, sollten sie strengere, innovationshemmende KI-Regulierungen erlassen, keine KI-Gelder aus Washington erhalten.

Die Sprachmodelle sollen die Meinungsäusserungsfreiheit respektieren – so wie sie die USA interpretieren – und dürfen keine ideologischen Inhalte verbreiten. Gemäss [AI Action Plan](#) fallen darunter Aussagen zu «Diversity, Equity, Inclusion and Climate Change». Diese gelten als «woke». Für die Tech-Konzerne, die sich nicht an diese Vorgaben halten, stehen enorme Bundesmittel auf dem Spiel. Allerdings bleiben die Vorgaben auch mit der am 23.7.2025 bereits erlassenen [Executive Order](#) «Preventing Woke AI in the Federal Government» vage. Gerade deshalb eignen sie sich als wirksame Disziplinierungswaffe der Trump-Regierung.

Eine grosse Herausforderung für Tech-Konzerne

Alle Verfahren zur Justierung von Sprachmodellen sind sehr aufwändig, kostenintensiv und erfordern viel Umsicht. Sprachmodelle durchlaufen milliardenfache automatisierte Trainingsprozesse. Zudem lernen diese Systeme im sogenannten [Reinforcement Learning](#) mit menschlichem Feedback, welches gute Antworten sind. Anschliessend wird das Sprachmodell mit Sicherheitstests geprüft und nach Bedarf nachgebessert, allenfalls mit zusätzlichen Anweisungen. Noch ist offen, ob und wie die Tech-Konzerne ihre Sprachmodelle umgestalten, um den Vorgaben zu genügen. Deren Glaubwürdigkeit könnte, z.B. bei Antworten auf Fragen zum Klimawandel, auf dem Spiel stehen.

Beispiel Grok von Elon Musk

Welche Probleme bei unvorsichtigem Nachjustieren auftreten können, musste Musk erfahren. Grok, das Sprachmodell von Elon Musks Firma xAI, wurde wegen «woker» Aussagen von seinen Nutzenden kritisiert. Gemäss einem Bericht der US-Website [The Daily Beast](#) vom 6.7.2025 machte Grok die Kürzungen Musks bei den Behörden für Wetterprognosen verantwortlich für die Toten bei den Überschwemmungen in Texas und stellte überdies fest, der Klimawandel werde zu weiteren Katastrophen führen. Mit dem Versuch, das Sprachmodell auf eine konservativere Linie zu bringen, entgleiste Grok und generierte antisemitische Tiraden und Lobeshymnen auf Hitler. Wie auf der [Plattform X](#) am 12.7.2025 nachzulesen ist, entschuldigte sich Musks Firma xAI hierfür. Ursache sei ein Update gewesen. Der Fehler sei behoben.

Lehren aus DeepSeek gezogen

Das chinesischen Start-up DeepSeek löste in den USA so etwas wie einen Sputnik-Schock und eine Trendwende zu Open Source aus (siehe Newsletter vom 17.6.2025). Der weltweit bekannte US-Risikokapitalgeber Marc Andreessen verlangte, so in der US-Nachrichtensite [Business Insider](#) vom 19.5.2025, dass die USA ein führendes Open-Source-Modell entwickeln sollten, damit nicht China diese Rolle übernehme. Der [AI Action Plan](#) entspricht voll diesem Anliegen. Postuliert wird ein Open Source förderliches Umfeld und ein Open-Source-Sprachmodell – mit den Werten der USA – als globaler Standard.

An der [World AI Conference](#) in Shanghai vom 26.-28.7.2025 schlug der chinesische Ministerpräsident Li Qiang eine globale Organisation für die KI-Zusammenarbeit vor, dies auch mit Blick auf die dringend nötige KI-Regulierung. Es ist dies gemäss der Nachrichtenagentur [Reuters](#) vom 26.7.2025 ein Kontrastprogramm zum [AI Action Plan](#).

Mit freundlichen Grüssen

Für das Netzwerk der ehemaligen SSAB: Hanna Muralt Müller

14.8.2025

Falls Sie diese E-Mail nicht mehr erhalten möchten, melden Sie sich bitte bei mir: info@muralt-mueller.ch.