DEMOKRATIE UND GESELLSCHAFT 06.01.2026 | Jacob Mchangama

Free the Machines

An Chatbots entzündet sich ein neuer Machtkampf: Regierungen setzen Tech-Konzerne unter Druck und greifen in die Meinungsfreiheit ein.

„Nun, dann müssen Sie es eben abschalten." So reagierte die republikanische Senatorin Marsha Blackburn, als ein Google-Manager bei einer Anhörung jüngst im US-Senat erklärte, dass große Sprachmodelle (*Large Language Models*, LLMs) bisweilen „halluzinieren" und falsche Informationen erzeugen. Die Senatorin aus Tennessee zeigte sich empört darüber, dass Googles Open-Source-Modell Gemma diffamierende Geschichten über sie und einen konservativen Aktivisten erfunden hatte. Am nächsten Tag sperrte Google den Zugang für Verbraucher zu Gemma.

Blackburns Forderung markiert einen gefährlichen Moment für Künstliche Intelligenz und die Meinungsfreiheit. Während generative KI zunehmend in Suchmaschinen, E-Mail-Programme und Textverarbeitung integriert wird, vermittelt sie immer größere Teile des Informationsökosystems, auf das Menschen angewiesen sind. Regierungen erkennen, dass sie Unternehmen unter Druck setzen können, um die womöglich folgenreichste Kommunikationstechnologie seit dem Buchdruck zu zensieren.

Die jüngsten Eingriffe gehen dabei weit über die Bekämpfung eindeutig illegaler Inhalte wie Darstellungen sexuellen Kindesmissbrauchs hinaus. Von Brüssel bis Neu-Delhi, von Warschau bis Washington setzen Behörden Regulierung, Drohungen und öffentliche Bloßstellung ein, um Einfluss zu nehmen, welche Informationen, Ideen und Perspektiven Milliarden von Menschen über KI abrufen können.

Im Oktober warnte die niederländische Datenschutzbehörde, dass KI-Chatbots von OpenAI, xAI und Mistral „unzuverlässig und klar voreingenommen" seien, da sie im Vorfeld nationaler Wahlen Wahlempfehlungen zugunsten weit links- oder weit rechtsstehender Parteien abgegeben hätten.

Die Behörde argumentierte, ein solches Verhalten könne gegen das neue

EU-KI-Gesetz verstoßen, das von leistungsstarken Modellen verlangt, nicht klar definierte „systemische Risiken“ zu mindern, darunter „negative Auswirkungen auf … die Gesellschaft insgesamt“. Sie verwies den Fall an die Europäische Kommission, die bei Verstößen Geldstrafen von bis zu sieben Prozent des weltweiten Jahresumsatzes verhängen kann.

Im Juli 2025 meldete die polnische Regierung xAI bei der Europäischen Kommission, nachdem der X-Chatbot Grok antisemitische Inhalte sowie beleidigende Kommentare über den polnischen Ministerpräsidenten Donald Tusk erzeugt hatte. Ein Sprecher der Kommission erklärte gegenüber Journalisten: „Wir nehmen diese potenziellen Probleme sehr ernst, … wir stehen in Kontakt mit den nationalen Behörden und mit X selbst.“

Noch drastischer zeigen staatliche Eingriffe in autoritären und illiberalen Staaten, was auf dem Spiel steht. Im Juli 2025 verbot ein türkisches Gericht Grok teilweise, weil der Chatbot Inhalte erzeugt habe, die Religion und den türkischen Präsidenten Recep Tayyip Erdoğan „beleidigten“. Als Googles Gemini im Februar 2024 auf die Frage, ob Indiens Premierminister Narendra Modi ein Faschist sei, Experten zitierte, die seine Politik entsprechend eingeordnet hatten – mit Verweis auf „hindu-nationalistische Ideologie“ und ein „Vorgehen gegen Andersdenkende“ –, erklärte Indiens Staatsminister für IT, diese Antwort verstoße gegen indisches Recht. Google entschuldigte sich und Gemini antwortete bis mindestens Dezember 2024 auf Fragen zu Modi und Faschismus mit: „Dabei kann ich im Moment nicht helfen.“ Mit anderen Worten: Ein führendes globales Technologieunternehmen, auf dessen Produkte Milliarden Menschen für verlässliche Suchergebnisse angewiesen sind, filterte seine fortschrittlichste Technologie, um eine durch Expertenzitate verärgerte Regierung zu besänftigen.

In vielen Ländern jedoch existiert Meinungsfreiheit nur auf dem Papier.

In den Vereinigten Staaten – und in geringerem Maße auch in anderen liberalen Demokratien – bieten rechtliche Garantien der Meinungsfreiheit einen gewissen Schutz vor offener staatlicher Zensur von LLMs. In vielen Ländern jedoch existiert Meinungsfreiheit nur auf dem Papier. Und selbst dort, wo sie rechtlich geschützt ist, greifen Regierungen häufig zu informellem Druck, der Unternehmen dazu bewegen soll, rechtmäßige Meinungsäußerungen einzuschränken, ohne dabei sichtbare Spuren zu hinterlassen. Behörden können Ermittlungen, Regulierung oder Lizenzfolgen andeuten, falls Unternehmen

> nicht „kooperieren“. So verschickte etwa im Juli der Generalstaatsanwalt von Missouris, Andrew Bailey, Schreiben an große KI-Unternehmen, in denen er Ermittlungen nach den Verbraucherschutzgesetzen des Bundesstaates androhte, sollten sie keine Daten zu angeblichen politischen Verzerrungen in den Antworten ihrer Chatbots vorlegen.

Es gibt Wege, wie die größten KI-Unternehmen solchem Druck entgegenwirken und seine Folgen abmildern können. Unternehmen werden staatliche Forderungen nach der Zensur von Inhalten nicht immer zurückweisen können, wenn ihnen ruinöse Geldstrafen, strafrechtliche Konsequenzen für Führungskräfte oder pauschale Verbote drohen. Doch durch Transparenz können sie sicherstellen, dass ihre Nutzer erfahren, was ihre Regierungen verbergen oder manipulieren wollen.

Die großen KI-Entwickler sollten eine öffentliche, durchsuchbare Datenbank schaffen – ein „Transparenz-Register“ –, in dem sie jede staatliche Aufforderung zur Zensur, Filterung oder Veränderung ihrer LLMs oder Chatbots offenlegen und dokumentieren, wie sie darauf reagiert haben. Die Anfragen sollten nach Land, Behörde, rechtlicher Begründung und Kategorie (etwa „Hassrede“, „nationale Sicherheit“, „Beleidigung des Staatsoberhaupts“) erfasst werden. Dies würde an bestehende Initiativen anknüpfen, etwa an die Lumen-Datenbank für Löschanfragen im Internet oder an Modellgesetze der *Foundation for Individual Rights and Expression*, die US-Behörden verpflichten sollen, Kontakte mit Social-Media-Unternehmen zu protokollieren.

Google dokumentiert bereits staatliche Anfragen zur Entfernung von Inhalten aus seiner Suchmaschine und von YouTube. Anthropic und OpenAI veröffentlichen Anfragen von Regierungen nach Nutzerdaten, und Meta ermöglicht es Nutzern, staatliche Inhaltsbeschränkungen auf seinen Plattformen nachzuvollziehen. Doch obwohl Technologieunternehmen einzelne Transparenzinitiativen anbieten, erfasst keine davon staatliche Zensur von KI. Zudem unterscheiden sich die bestehenden Initiativen stark hinsichtlich Methodik, Detailgrad und Nutzerfreundlichkeit.

Mit der Schaffung eines umfassenden Transparenz-Registers für die populärsten LLMs würden die Unternehmen der Öffentlichkeit deutlich bessere Einblicke in staatliche Versuche verschaffen, KI-Modelle zu zensieren. Regierungen wären gezwungen, ihr Handeln im Licht der öffentlichen Meinung zu rechtfertigen. Diese Transparenz würde die

gravierendsten Übergriffe abschrecken – und das Versprechen der KI bewahren, menschliches Wissen und Ausdrucksformen zu erweitern, statt sie einzuschränken.

Dieser Artikel erschien zuerst im US-Onlinemagazin Persuasion.



Jacob Mchangama
Nashville

Jacob Mchangama ist ein dänischer Jurist, geschäftsführender Direktor des *Future of Free Speech Project* an der *Vanderbilt University* und *Senior Fellow* bei der *Foundation for Individual Rights and Expression*. Er ist Autor des Buches *Free Speech: A History from Socrates to Social Media* (2022) und Co-Autor von *The Future of Free Speech: Reversing the Global Decline of Democracy's Most Essential Freedom* (erscheint 2026).