

Risiken der KI und die mögliche Rolle der Schweiz

Gespräch mit Daniel Stauffacher, ICT4Peace Foundation

Sender: Deutschlandfunk

Sendung: Europa heute

Moderation: Manfred Götzke, Deutschlandradio, Landesstudio Berlin

Ausstrahlung: Freitag, 18. September 2026, ca. 09:12 Uhr

Dauer: 10:06 Min.

Sprache: Deutsch

Transkript (Deutsch)

Transkribiert aus der Sendungsaufzeichnung; leicht geglättet (Füllwörter, Versprecher). Alle Aussagen von Daniel Stauffacher sind inhaltlich unverändert.

Manfred Götzke: 10 bis 20 Prozent – so hoch schätzen die Chefs der grossen KI-Firmen das Risiko ein, dass die KI uns in den nächsten Jahren auslöschen könnte. Die Chefs von OpenAI und Anthropic, sogar Elon Musk, fordern deshalb eine Regulierung und eine Verlangsamung der Entwicklung. Nur dürfte das wenig bringen, wenn nicht alle relevanten Firmen beziehungsweise Staaten mitziehen. Es bräuchte also ein internationales Abkommen – mit der Schweiz als neutraler Instanz, die alle an einen Tisch bringt. So sieht es jedenfalls Daniel Stauffacher. Er ist Gründer der ICT4Peace Foundation, einer NGO, die sich für eine friedliche digitale Welt einsetzt. Ich habe Daniel Stauffacher vor der Sendung gefragt, wie sehr ihn die ziemlich einmütige Forderung der KI-Firmen überrascht hat, dass die Entwicklung der KI verlangsamt werden muss.

Stauffacher: Ja, also das ist ein ganz spezielles Thema, weil es im Moment ein Durcheinander von verschiedenen dieser grossen Firmen gibt. Die einen wollen Verlangsamung, die anderen nicht. Ich bin der Meinung, dass letztlich eine Verlangsamung im Moment nicht möglich ist – auch aus geopolitischen Gründen, angesichts des Wettbewerbs zwischen den Vereinigten Staaten und China, und generell wegen der Investitionsdynamik. Deshalb geniesse ich das mit Vorsicht.

Manfred Götzke: Sehen Sie denn einen internationalen Konsens, dass KI reguliert oder zumindest die Entwicklung verlangsamt werden sollte oder muss, damit die KI uns nicht in den nächsten Jahren, bis zum Ende des Jahrzehnts – wie in dieser Woche oft zu lesen war – auslöscht?

Stauffacher: Da muss man trennen. Bei der Regulierung sehe ich einen wachsenden Konsens, eine Konvergenz. Das sieht man nicht nur in Erklärungen, sondern auch in den Handlungen: Wir haben den EU AI Act, wir haben die Europaratskonvention, wir haben nationale Gesetze. Dann haben wir die chinesische Initiative für die World Artificial Intelligence Cooperation Organization, und auf UNO-Ebene den Global Dialogue on AI Governance und das Scientific Panel. Da sehe ich eine Konvergenz. Beim Verlangsamen bin ich mir, wie gesagt, nicht sicher. Wo ich aber auch einen wachsenden Konsens sehe, ist bei den extremen Risiken der Frontier-Modelle: der Missbrauch für Bio- und Cyberwaffen, Kontrollverlust, militärische Autonomie. Da haben auch die Konkurrenten eigentlich ein Interesse daran, dass das irgendwie eingebunden werden kann. Es gab ja gerade bei diesen Frontier-Modellen sehr problematische, besorgniserregende Vorfälle.

Manfred Götzke: Jetzt sagen Sie, bei einer internationalen Regulierung der KI könnte die Schweiz eine relevante, eine wichtige Rolle spielen. Inwiefern? Was schwebt Ihnen da vor?

Stauffacher: Es ist ein bisschen auch ein glücklicher Zufall: Die Schweiz organisiert im nächsten Jahr einen AI-Weltgipfel, im Nachgang zu den anderen Gipfeln, die in Bletchley begonnen haben und in Seoul, Paris und Neu-Delhi weitergingen. In diesem Prozess haben wir nicht mehr viel Zeit, aber dennoch haben wir die Chance, diese Frage der Konvergenz und der möglichen Regulierung zu diskutieren. Die Schweiz hat eine gewisse Tradition als Brückenbauerin in den UNO-Prozessen, als Gastland verschiedener UNO-Organisationen. Und die Schweiz könnte vielleicht dazu beitragen, gegen eine mögliche Blockbildung der Grossmächte betreffend KI die Leute und die Mitglieder zusammenzubringen. Aber da können wir noch mehr ins Detail gehen, wenn Sie wollen.

Manfred Götzke: Ja, das machen wir gerne. Wie könnte denn eine Regulierung aussehen, und wie könnte man alle relevanten Player an einen Tisch und zu einem Konsens bringen?

Stauffacher: Ich sehe das auf drei Ebenen. Die erste Ebene ist der Bereich der Frontier-Modelle, über den wir gerade gesprochen haben, mit den grössten Risiken, die eigentlich nicht im Interesse der Grossmächte sind. Da könnte man zum Beispiel verbindliche Sicherheitsregeln für die wenigen Entwickler der leistungsstärksten Modelle vereinbaren: Tests vor der Freigabe, Meldepflichten für Zwischenfälle, Rechenschwellen – und die Regeln folgen den Fähigkeiten des Modells, nicht dem Herkunftsland. Die zweite Ebene wäre der Zugang und die Verbreitung; das ist etwas, das für mich sehr wichtig ist: eine Verpflichtung zur Offenheit, zum Aufbau von Rechenkapazität und Kompetenz für Länder ohne eigene Modelle, damit das Abkommen nicht als Kartell der Grossen wahrgenommen wird. Die dritte Ebene bliebe die nationale Ebene – alles, was darunter bleibt: der EU AI Act, die Schweizer Gesetzgebung, Mindeststandards und so weiter. Was man vielleicht auch vereinbaren könnte, wären Institutionen: eine wissenschaftliche Beobachtungsstelle, so wie wir das in der Klimadebatte haben, eine gemeinsame Test- und Zwischenfall-Meldestelle, vielleicht in der Zukunft auch eine Art IAEA. Realistisch wäre für Genf sicher noch keine detaillierte Verhandlung, aber vielleicht eine politische Erklärung, die dann gefolgt würde von einer Rahmenkonvention mit Protokollen, analog zum Rio-Klimaprozess.

Manfred Götzke: Sie haben es gerade schon angedeutet: Die Schweiz ist oder wäre hier besonders prädestiniert, weil andere internationale Abkommen – Atomwaffen, Chemiewaffen – auch in der Schweiz verhandelt und beschlossen wurden.

Stauffacher: Genau, das ist das Know-how, das die Schweiz hat. Wichtig finde ich auch, dass die Schweiz neutral ist und selber keine Frontier-Labs hat, sondern das offene Open-Source-Modell Apertus gegründet hat. Das ist vielleicht auch ein wichtiges Element, um als – sagen wir – Honest Broker, neutral, einen Konsens zwischen all den Blöcken und den Entwicklungsländern herbeizuführen.

Manfred Götzke: Viele halten KI für genauso gefährlich wie Atom- oder Chemiewaffen, für die es ja Abkommen gibt. Könnten diese Verträge eine Blaupause sein?

Stauffacher: Absolut. Oder auch zum Beispiel die Basler Vorgaben für die Grossbanken. Das ist auch ein Modell, das mir vorschwebt, wo man sich auf die höchsten Risiken – die grössten, systemrelevanten Banken – konzentriert und diese eng reguliert, während man die kleineren Banken etwas freier lässt. In dieser Abstufung, von den grössten Risiken her, wo das Interesse wie bei den Nuklearabkommen für alle Seiten – Abschreckung und so weiter – wichtig ist, kann man einen Konsens finden.

Manfred Götzke: Jetzt haben einige Chefs der grossen KI-Firmen für eine Verlangsamung plädiert und sich eine Regulierung quasi gewünscht. Trump will es nicht. Bei China weiss man es nicht so ganz genau. Wie wahrscheinlich halten Sie es, dass in den nächsten Monaten, vielleicht im nächsten Jahr, ein tragfähiges Abkommen zustande kommen könnte?

Stauffacher: Bei der Regulierung muss ich sagen: Da haben eigentlich auch die USA reguliert – das habe ich in meinem Aufsatz auch geschrieben –, obwohl Trump hin und her geht. Ich glaube, er

meinte eher die Verlangsamung als ein Minimum an Regulierung. Die Chinesen sind eigentlich auch für Regulierung. Es gibt viele Beispiele, wo die USA und die Grossmächte faktisch doch schon regulieren. Aber auf der mittleren Ebene wird es schwierig sein, einen Konsens zu erreichen. Bei den grossen Risiken hingegen, die wir besprochen haben, wo eine Gefahr für alle Teilnehmenden besteht, sollten wir einen Konsens hinkriegen.

Manfred Götzke: Im Juni findet, Sie haben es schon gesagt, in der Schweiz, in Genf, ein internationaler KI-Gipfel statt. Und da, sagen Sie, könnte mit entsprechenden Verhandlungen begonnen werden.

Stauffacher: Jawohl. Wir werden sicher noch kein Abkommen haben für diesen Gipfel, dafür ist die Zeit zu kurz. Aber wir könnten eventuell ein Mandat von der Weltgemeinschaft erhalten, die sagt: Jawohl, das ist ein wichtiges Thema, und wir beauftragen die UNO oder gewisse Länder, die Verhandlungen aufzunehmen – zuerst mit einer Rahmenkonvention, dann mit Protokollen – und einen Verhandlungsprozess einzuleiten. Jetzt schon etwas Ausverhandeltes zu haben im nächsten Juni scheint mir unwahrscheinlich. Aber ein Mandat sollte möglich sein, vor allem wenn sich auch das Gastland, die Schweiz, also meine Regierung, stark dafür einsetzt.

Manfred Götzke: Die Schweiz als Land für eine internationale KI-Regulierung – darüber habe ich mit Daniel Stauffacher gesprochen, Ex-Diplomat und Gründer einer NGO, die sich für eine friedliche digitale Welt einsetzt.

AI Risks and the Possible Role of Switzerland

[Interview with Daniel Stauffacher, ICT4Peace Foundation](#)

Broadcaster: Deutschlandfunk (German public radio)

Programme: Europa heute

Presenter: Manfred Götzke, Deutschlandradio, Berlin studio

Aired: Friday, 18 September 2026, approx. 09:12 CEST

Duration: 10:06 min.

English translation

Translated from the German broadcast transcript.

Manfred Götzke: Ten to twenty percent – that is how high the heads of the major AI companies rate the risk that AI could wipe us out in the coming years. The CEOs of OpenAI and Anthropic, even Elon Musk, are therefore calling for regulation and a slowdown of development. But that is unlikely to achieve much unless all the relevant companies and states go along with it. What would be needed is an international agreement – with Switzerland as a neutral body that brings everyone to the table. That, at any rate, is how Daniel Stauffacher sees it. He is the founder of the ICT4Peace Foundation, an NGO that works for a peaceful digital world. Before the programme I asked Daniel Stauffacher how surprised he was by the fairly unanimous demand from the AI companies that AI development must be slowed down.

Stauffacher: This is a very particular issue, because at the moment there is a real muddle among these large companies. Some want a slowdown, others do not. My view is that a slowdown is ultimately not possible at the moment – for geopolitical reasons, given the competition between the United States and China, and generally because of the investment dynamics. So I treat this with caution.

Manfred Götzke: Do you see an international consensus that AI should or must be regulated, or at least that its development must be slowed, so that AI does not wipe us out in the coming years, by the end of the decade – as one could often read this week?

Stauffacher: One has to distinguish. On regulation, I see a growing consensus, a convergence. You can see it not only in declarations but also in actions: we have the EU AI Act, we have the Council of Europe convention, we have national laws. Then we have the Chinese initiative for the World Artificial Intelligence Cooperation Organization, and at UN level the Global Dialogue on AI Governance and the Scientific Panel. There I see convergence. On slowing down, as I said, I am not sure. But where I also see a growing consensus is on the extreme risks of frontier models: misuse for biological and cyber weapons, loss of control, military autonomy. There, even the competitors actually have an interest in seeing this contained in some way. And it is precisely with these frontier models that there have been very problematic, worrying incidents.

Manfred Götzke: You say that in an international regulation of AI, Switzerland could play a relevant, an important role. In what way? What do you have in mind?

Stauffacher: It is partly a happy coincidence: next year Switzerland is organising a global AI summit, following on from the summits that began in Bletchley and continued in Seoul, Paris and New Delhi. There is not much time left in this process, but we still have the chance to discuss this question of convergence and possible regulation. Switzerland has a certain tradition as a bridge-builder in UN

processes, as host country of various UN organisations. And Switzerland could perhaps help, against a possible bloc formation among the great powers on AI, to bring people and members together. But we can go into more detail if you like.

Manfred Götzke: Yes, gladly. What could regulation look like, and how could all the relevant players be brought to the table and to a consensus?

Stauffacher: I see it on three levels. The first level is the area of frontier models we have just discussed, with the greatest risks, which are really not in the interest of the great powers. There one could, for example, agree binding safety rules for the few developers of the most powerful models: testing before release, mandatory incident reporting, compute thresholds – and the rules follow the model's capabilities, not its country of origin. The second level would be access and diffusion; this is something very important to me: a commitment to openness, to building compute capacity and expertise for countries without models of their own, so that the agreement is not perceived as a cartel of the big players. The third level would remain the national level – everything below that: the EU AI Act, Swiss legislation, minimum standards and so on. What one could perhaps also agree on are institutions: a scientific observatory, as we have in the climate debate, a joint testing and incident-reporting body, perhaps in future something like an IAEA. Realistically, Geneva will certainly not yet see detailed negotiations, but perhaps a political declaration, which would then be followed by a framework convention with protocols, along the lines of the Rio climate process.

Manfred Götzke: You have already hinted at it: Switzerland is, or would be, particularly well placed here, because other international agreements – on nuclear weapons, chemical weapons – were also negotiated and concluded in Switzerland.

Stauffacher: Exactly, that is the know-how Switzerland has. I also think it is important that Switzerland is neutral and has no frontier labs of its own, but has instead created the open-source model Apertus. That may also be an important element for bringing about, as an honest broker, neutrally, a consensus between all the blocs and the developing countries.

Manfred Götzke: Many consider AI to be just as dangerous as nuclear or chemical weapons, for which there are agreements. Could those treaties serve as a blueprint?

Stauffacher: Absolutely. Or, for example, the Basel rules for the large banks. That is also a model I have in mind, where one concentrates on the highest risks – the largest, systemically relevant banks – and regulates them tightly, while leaving the smaller banks somewhat freer. With that kind of gradation, starting from the greatest risks, where – as with the nuclear agreements – all sides have an interest, in deterrence and so on, one can find a consensus.

Manfred Götzke: Some heads of the big AI companies have argued for a slowdown and more or less asked for regulation. Trump does not want it. With China one does not quite know. How likely do you think it is that a viable agreement could come about in the coming months, perhaps next year?

Stauffacher: On regulation I have to say: the USA has in fact also regulated – I wrote that in my essay too – even though Trump goes back and forth. I think he meant the slowdown rather than a minimum of regulation. The Chinese are in fact also in favour of regulation. There are many examples where the USA and the great powers are in practice already regulating. But at the middle level it will be difficult to reach a consensus. On the major risks we discussed, however, where there is a danger for all participants, we should be able to reach a consensus.

Manfred Götzke: In June, as you said, an international AI summit is taking place in Switzerland, in Geneva. And there, you say, the corresponding negotiations could begin.

Stauffacher: Yes. We will certainly not have an agreement in time for this summit; the time is too short. But we could possibly obtain a mandate from the world community, which says: yes, this is an important issue, and we task the UN or certain countries with taking up negotiations – first a

framework convention, then protocols – and with launching a negotiation process. To already have something fully negotiated by next June seems unlikely to me. But a mandate should be possible, especially if the host country, Switzerland – my government – pushes hard for it.

Manfred Götzke: Switzerland as the country for international AI regulation – I discussed that with Daniel Stauffacher, former diplomat and founder of an NGO that works for a peaceful digital world.