

Link zu diesem Dokument: <https://www.fwes.info/extreme-risiken-2024-1.pdf>

## Extreme Risiken

### Zusammenfassung

Die Menschheit wird derzeit von gravierenden Risiken bedroht, insbesondere bezogen auf Künstliche Intelligenz (KI) und Atomwaffen. Führende KI-Wissenschaftler haben erneut vor extremen Risiken durch KI gewarnt. Auch Angriffe auf gegnerische Nuklearstreitkräfte könnten ein Atomkriegsrisiko erheblich erhöhen. Es ist äußerst fraglich, ob die Verbesserung militärischer Stärke und nuklearer Abschreckung geeignet ist, diese Risiken zu reduzieren. Stattdessen könnte dieser Weg in einem Atomkrieg enden. Andere Maßnahmen könnten dringend erforderlich sein, um diese Risiken zu reduzieren. Dazu kann jeder Einzelne etwas beitragen. Nachfolgend werden die Risiken und mögliche Maßnahmen beschrieben.

### Extreme Risiken durch KI

In einem Artikel bei Science warnen am 20.5.24 führende KI-Wissenschaftlern erneut vor extremen KI-Risiken und dass es zu einem Kontrollverlust kommen könnte. Links zu dem Artikel und der PDF-Datei:

- <https://www.science.org/doi/10.1126/science.adn0117>
- [https://www.science.org/doi/suppl/10.1126/science.adn0117/suppl\\_file/science.adn0117\\_sm.pdf](https://www.science.org/doi/suppl/10.1126/science.adn0117/suppl_file/science.adn0117_sm.pdf)

Bereits 2023 gab es Warnungen von führenden KI-Wissenschaftlern und Chefs großer KI-Unternehmen, dass KI zum Aussterben der Menschheit führen könnte.<sup>1</sup> Solche Warnungen sind auch verbunden mit der Erwartung, dass KI-Systeme schon bald ein menschliches Intelligenzniveau erreichen könnten und dann vielleicht sogar eine Superintelligenz möglich sei. Allerdings können extreme Risiken durch KI auch bereits lange vorher entstehen. Auf solche Aspekte wird hier hingewiesen: <https://fwes.info/GenKI-Internet-2024-1.pdf>

Bezogen auf die Risiken durch KI kann es auch gefährliche Wechselwirkungen zwischen solchen KI-Systemen und Atomwaffen sowie anderen Massenvernichtungswaffen geben, die sich einer menschlichen Kontrolle möglicherweise vollständig entziehen. In der neuen Warnung bei Science wird auch bemängelt, dass die Warnungen der führenden KI-Wissenschaftler von Politik und Gesellschaft viel zu wenig ernst genommen werden.

---

<sup>1</sup> <https://www.safe.ai/statement-on-ai-risk>

## **Atomkriegsrisiko**

Am 25.5.24 gab es Meldungen, dass ein russisches Atomraketen-Frühwarnsystem durch einen Drohnen-Angriff getroffen wurde. Solche Vorfälle in Krisen- oder Kriegszeiten können das Atomkriegsrisiko erheblich erhöhen.

Hinweise auf diesen Vorfall:

- <https://www.telepolis.de/features/Eskalation-im-Ukraine-Krieg-Russisches-Atomraketen-Fruehwarnsystem-getroffen-9732435.html>
- <https://www.dailyalert.de/ukrainischer-angriff-auf-russisches-fruehwarnradar-droht-einen-nuklearen-weltkrieg-auszuloesen/>
- <https://www.kurier.de/inhalt.krieg-in-der-ukraine-berichte-treffer-auf-russisches-atomraketen-fruehwarnsystem.79bbacd5-8bc1-4483-a1cb-699ea67d9fd6.html>.

Hinweise auf diesen Vorfall gab es in unseren „Leitmedien“ kaum. Drohungen mit dem Einsatz von Atomwaffen sind genauso unverantwortlich, wie Angriffe auf Komponenten der Nuklearstreitkräfte einer Atommacht. Solche Ereignisse können leicht zu Missverständnissen und Fehlinterpretationen und damit zu einem Atomkrieg aus Versehen führen.

## **Reicht militärische Stärke zur Risikoreduzierung?**

Die westliche Politik setzt auf militärische Stärke und eine nukleare Abschreckung. Diese Haltung wird von den meisten Medien unterstützt oder sogar gefordert. Als Vorbild gilt vielen der Nato-Doppelbeschluss zu Beginn der 1980er Jahre und die damit verbundene nukleare Aufrüstung. Dies habe schließlich zum INF-Vertrag, nuklearer Abrüstung und Entspannung geführt. Deshalb könnte es auch jetzt der richtige Weg sein, auf militärische Stärke, einschließlich Atomwaffen zu setzen. Angenommen diese Strategie funktioniert auch jetzt, dann wäre nichts weiter zu tun und alle Aktivitäten der Friedensbewegung wären irrelevant.

Wie sicher ist es aber, dass eine solche Strategie auch dieses Mal funktioniert? Was könnte dagegensprechen?

## **Risiken der Strategie militärischer Stärke**

Der Nato-Doppelbeschluss hätte auch in den 1980er Jahren schief gehen können. Es gab auch in den 1980er Jahren gefährliche Situationen, die zu einem Atomkrieg aus Versehen hätten führen können. Dass eine solche Strategie der militärischen Stärke gegen eine Atommacht einmal gut ging, bedeutet nicht, dass eine solche Strategie immer funktioniert und immer zu einem positiven Ergebnis führt.

Es ist auch fraglich, wie der weitere geschichtliche Verlauf gewesen wäre, wenn nicht Gorbatschow in der Sowjetunion an die Macht gekommen wäre, sondern jemand anders. Es ist nicht sicher, dass es dann zu einem ähnlichen Ergebnis gekommen wäre.

In den 1980er Jahren gab es keine direkte Konfrontation zwischen den Atommächten. In den jetzigen Ukraine-Krieg droht die Nato immer stärker hereingezogen zu werden. Dies könnte

viele Ebenen betreffen, wie z.B. Waffenlieferungen, Ausbildung, direkte Unterstützung, Cyberangriffe, Weltraumaktivitäten. Damit steigt das Risiko von Zusammenstößen und Konflikten zwischen Atommächten. Folglich sind die nuklearen Risiken größer als in den 1980er Jahren.

In den 1980er Jahren waren die Atommächte vorsichtig bei Aktionen, die zu nuklearen Missverständnissen hätten führen können. Dieses Risikobewusstsein scheint jetzt geringer zu sein. Selbst Komponenten der Nuklearstreitkräfte eines Gegners werden angegriffen und zerstört, wie Ende Mai 2024 (siehe oben).

In den 1980er Jahren spielten Cyberbedrohungen und autonomen Waffensysteme noch keine Rolle. Heute kann es diesbezüglich gefährliche Wechselwirkungen mit Nuklearstreitkräften geben, die kaum kontrollierbar sind.

Es ist also äußerst zweifelhaft, ob auch dieses Mal eine Strategie militärischer Stärke ausreichend ist. Wenn die Strategie schief geht und es zu einem Atomkrieg kommt, ist das Überleben der gesamten Menschheit bedroht.

### **Pyrrhussieg bei Erfolg der Strategie militärischer Stärke**

Falls die Strategie militärische Stärke weiterverfolgt wird und diese sich tatsächlich als erfolgreich herausstellen sollte, dann wird dies vermutlich lediglich ein Pyrrhussieg sein. Denn der Weg dorthin wird mit einem enormen Rüstungswettlauf verbunden sein. Dieser Rüstungswettlauf wird alle militärischen Dimensionen, auch das Weltall und den Cyberraum betreffen und stark von Informatik und KI geprägt sein. Dabei werden Waffensysteme entstehen, die für Menschen nicht mehr kontrollierbar sein werden, aber verheerende Auswirkungen haben werden. Auch ein Wettlauf um eine Superintelligenz wird ungebremsst weiter gehen. Alle großen Industrienationen werden sich hieran beteiligen. Die Folgen werden völlig unkalkulierbar sein.

### **Grenzen der Abschreckung**

Selbst unter der Annahme, dass die nukleare Abschreckung bisher erfolgreich war und Kriege verhindert hat, gilt nicht, dass dies auch in Zukunft so bleibt. Nukleare Abschreckung ist lediglich eine Androhung von Strafe, dieses Konzept kann versagen. Die neuen technischen Entwicklungen vor allem im Bereich der Informatik und speziell der Künstlichen Intelligenz führen zu neuen Bedrohungsarten, gegen die eine nukleare Abschreckung nicht schützen kann. Stattdessen erhöht sich das nukleare Eskalationsrisiko.

Solche neuen Bedrohungsarten können sein:

- Gravierende Cyberangriffe, z.B. auf kritische Infrastruktur,
- Autonome Waffensysteme,
- Einsatz von Biowaffen, eventuell mit Hilfe von KI erzeugt,
- Desinformationen, Deep Fakes, Störung des Informationsaustausches im Internet,
- Superintelligenz

Eine nukleare Abschreckung mag schützen gegen einen Angriff durch einen anderen Staat mit konventionellen oder nuklearen Waffen. Aber eine nukleare Abschreckung

- schützt **nicht** vor einem Atomkrieg aus Versehen als Folge von Fehlern in einem Frühwarnsystem und Missverständnissen in Krisensituationen,
- schützt **nicht** vor schwerwiegenden Cyberangriffen, die auch von privaten Akteuren durchgeführt werden können,
- schützt **nicht** vor massenhafter Anwendung von autonomen Waffen, z.B. durch Terroristen,
- schützt **nicht** vor Biowaffen, eventuell mit Hilfe von KI erzeugt und von Terroristen eingesetzt,
- schützt **nicht** vor Risiken durch Systeme der generativen KI<sup>2</sup>,
- schützt **nicht** vor den Folgen einer Superintelligenz.

Gravierende Folgen der aktuellen technischen Entwicklungen können schon relativ bald eintreten, auch eine AGI oder sogar eine Superintelligenz.<sup>3</sup>

Gegen diese Risiken können keine Atomwaffen schützen. Davor schützen kann nur ein schnelles Ende der aktuellen Kriege und der Aufbau von Vertrauen und Zusammenarbeit mit allen Nationen, auch mit Russland und China.

### Optionen zur Risikoreduzierung

Aufgrund der erheblichen Risiken, die durch neue technische Entwicklungen auf uns zu kommen können, stellt sich die Frage, wie diese reduziert werden können. Als Möglichkeiten Cyberangriffe abzuschrecken, nennen Generalleutnant a.D. Kersten Lahl und Politikwissenschaftler Johannes Varwick vier Optionen:<sup>4</sup>

1. Androhung von Strafe,
2. Sicherheit durch Resilienz
3. internationale Verflechtung
4. internationale Normensetzung

Lahl und Varwick betonen, dass keine dieser vier Optionen alleine für eine bessere Sicherheit reicht, erforderlich sei eine Kombination von mehreren oder allen Optionen. Diese Optionen sind nicht nur als Schutz und Abschreckung von Cyberangriffen relevant, sondern können auch für viele der anderen Risiken herangezogen werden.

Zu Option 1 können entsprechende Gegenangriffe aber auch Sanktionen zählen. Die nukleare Abschreckung ist ein Kernelement dieser Option und bezieht sich nicht nur auf mögliche nukleare Angriffe. Nach Militärstrategien könnten z.B. auch schwerwiegende Cyberangriffe zu einem Einsatz von Atomwaffen führen. Eine solche Option alleine wird nicht reichen und kann versagen.

Option 2 beruht darauf die Sicherheit durch technische und organisatorische Vorsorge zu bessern. Das kann in vielen Fällen sinnvoll sein, wird alleine aber auch nicht reichen. Nicht gegen alle möglichen Angriffe im Cyberraum, gegen Angriffe durch Waffensysteme mit zunehmender Autonomie und nukleare Waffen ist ein vollkommener technischer Schutz möglich. Bei einem Atomkrieg aus Versehen könnten alleine deshalb mehr Raketen als Vergeltung gestartet werden, weil alle Seiten wissen, dass ein Teil der Raketen abgefangen werden könnte.

---

<sup>2</sup> <https://ki-folgen.de/informationsdominanz/> und <https://fwes.info/GenKI-Internet-2024-1.pdf>

<sup>3</sup> siehe <https://ki-folgen.de/>

<sup>4</sup> Lahl, Kersten/Varwick, Johannes (2022): Sicherheitspolitik verstehen – Handlungsfelder, Kontroversen und Lösungsansätze. Wochenschauverlag, 3. Auflage: 117ff

Option 3 ist eine internationale Verflechtung. Dazu schreiben Lahl und Varwick: „Dieser Ansatz nutzt die Erkenntnis, dass in einer global vernetzten Welt jede gewaltsame Auseinandersetzung unter dem Strich nur Verlierer hervorbringt. Je stärker also die Akteure miteinander wirtschaftlich, technologisch, kulturell und ggf. auch militärisch vernetzt sind, desto geringer ist die Chance, durch Aggression einseitige Vorteile erzielen zu können. Dieser oft zu Unrecht als naiv empfundene Gedanke läuft damit auf eine Art der Selbstabschreckung hinaus.“ Die Grenzen dieses Ansatzes sehen Lahl und Varwick darin, dass oft nationale Interessen verfolgt werden und es daher eventuell an der Bereitschaft der betreffenden Akteure zu Vertrauensbildung und gegenseitiger Verflechtung mangelt.

Option 4 ist eine internationale Normensetzung. Internationale Vereinbarungen könnten ein Mittel sein, um einen unkontrollierten Rüstungswettlauf zu dämpfen. Es gibt keine Garantie, dass solche Regeln eingehalten werden, trotzdem könnten diese eine abschreckende Wirkung haben und die Risiken reduzieren.

Für eine dauerhafte globale Sicherheit könnte Option 3 wichtiger sein als Option 1. Auch wenn die nukleare Abschreckung bisher eine Wirkung erzielt hat, gibt es keine Garantie, dass dies so bleibt. Sie kann mit verheerenden Folgen versagen.

### **Notwendigkeit von Zusammenarbeit und Vertrauen**

Die Optionen 1 und 2 schützen nicht vor Risiken in Zusammenhang mit Systemen der generativer KI. Militärische Stärke und Abschreckung helfen nicht, da diese Risiken nicht in erster Linie von Staaten ausgehen. Schutz bieten können nur eine weltweite Zusammenarbeit und Vereinbarungen zwischen allen Nationen. Als Grundlage dafür ist ein gewisses Maß an Vertrauen erforderlich.

Die Globalisierung der Wirtschaft in den letzten Jahrzehnten hat in die Richtung einer Internationalen Verflechtung gewirkt, also Option 3 unterstützt. Nun ist dieser Prozess gebremst oder gestoppt und viele Beziehungen zwischen dem Westen und Russland sind abgebrochen. Beziehungen zu China werden in Frage gestellt. Seit Beginn des Ukraine-Krieges wird immer wieder behauptet, die Strategie „Wandel durch Handel“ sei gescheitert. „Wandel“ ist in diesem Zusammenhang ein ungünstiger Begriff. Besser wäre der Begriff „Sicherheit“. Statt eine Behauptung aufzustellen „Sicherheit durch Handel“ sei gescheitert, wäre eine Unterscheidung von notwendigen und hinreichenden Bedingungen korrekter. Wenn eine Aussage A hinreichend für eine Aussage B ist, dann gilt „aus A folgt B“. Wenn eine Aussage A notwendig für eine Aussage B ist, dann gilt „aus B folgt A“. Die Folgerung gilt hier also genau in der entgegengesetzten Richtung.

Die Strategie „Sicherheit durch Handel“ war also nicht hinreichend, um den Krieg in der Ukraine zu verhindern. Dafür wären noch weitere Maßnahmen erforderlich gewesen. Aber eine solche Strategie ist notwendig, um das Risiko weiterer Eskalationen und weiterer militärischer Konflikte zu reduzieren. „Sicherheit durch Handel“ ist notwendig, um dauerhaft ein gewisses Maß an globaler Sicherheit und damit eine friedlichere Welt zu erreichen.

Für eine globale Sicherheit und eine friedlichere Welt sind alle vier genannten Optionen relevant. Die Optionen 1 und 2 werden alleine nicht ausreichen. Die Optionen 3 und 4 sind besonders wichtig. Bezüglich Option 4 wären dringend wirksame Vereinbarungen zwischen allen Nationen erforderlich zum Klimawandel, zur Rüstungskontrolle bei Atomwaffen, zu autonomen Waffen und Cyberwaffen sowie zur Regulierung der KI. Als Voraussetzung hierfür müssen Vertrauen, Zusammenarbeit und gute Kommunikationskanäle zwischen allen Nationen, auch heutigen Gegnern, aufgebaut und verbessert werden. Die kommenden Risiken sind nicht alleine auf technischer Ebene lösbar.

Je stärker potenzielle Gegner wirtschaftlich, technologisch und kulturell miteinander vernetzt sind, je besser damit Vertrauen und Zusammenarbeit sind, desto geringer werden Motivation und Bestrebungen

sein, Mittel und Prioritäten auf die Entwicklung z.B. von autonomen Waffen zu legen. Die Entwicklung von gefährlichen Waffensystemen kann damit vermutlich nicht ganz verhindert, aber zumindest abgeschwächt werden. Bei einem guten Verhältnis zwischen allen Nationen auf verschiedenen Ebenen verschieben sich Prioritäten, hin zu sinnvollen KI-Anwendungen, die dem Menschen nutzen. Des Weiteren könnten damit auch die Voraussetzungen geschaffen werden, um wichtige weltweite Vereinbarungen zur Rüstungskontrolle und zur Regulierung der KI zu ermöglichen.

Für ein gutes Vertrauensverhältnis zwischen Nato-Staaten und Russland sind enge Beziehungen auf verschiedenen Ebenen wichtig, dazu gehören wirtschaftliche, wissenschaftliche, sportliche, kulturelle und auch private Beziehungen. Auch Städtepartnerschaften spielen hierbei eine wichtige Rolle. War es wirklich notwendig und sinnvoll, all dies, was seit Jahren und Jahrzehnten aufgebaut wurde, jetzt zu zerstören? Ist es sinnvoll, Künstler und Sportler aus Russland von allem auszuschließen, Städtepartnerschaften zu beenden? Zum Erreichen einer dauerhaften globalen Sicherheit wäre es gut, wenn dieser Prozess wieder umgekehrt werden könnte.

Jede Verbesserung von Beziehungen zwischen verschiedenen Nationen ist ein Beitrag zu vertrauensbildenden Maßnahmen und damit zum Erhalt oder zum Schaffen von Frieden.

### **Was kann ich tun?**

Zur Verbesserung von Beziehungen zwischen verschiedenen Nationen kann jeder Einzelne einen Beitrag leisten. Man kann sich lokal für Partnerschaften mit russischen Regionen oder Städten einsetzen oder für mehr kulturellen Austausch. Auch Gespräche mit Freunden und Bekannten sowie mit Politikern und Medienvertretern können relevant sein. Zu Journalisten von lokalen Zeitungen oder Lokalpolitikern findet man häufig auch Kontaktdaten im Netz, auch Telefonnummern. Man kann diese anrufen, mit ihnen sprechen, diskutieren, Argumente vorbringen und auf aktuelle Risiken hinweisen.

Die bevorstehenden Risiken insbesondere im möglichen Zusammenwirken zwischen KI und Massenvernichtungswaffen können so gravierend sein, dass sich das Engagement jedes Einzelnen lohnt. Wenn es in einem Frühwarnsystem für nukleare Bedrohungen zu einem Fehlalarm kommt, sodass ein nuklearer Angriff gemeldet wird, obwohl kein Angriff vorliegt, kann das Überleben der gesamten Menschheit von Kleinigkeiten abhängen. Für die Einschätzung, ob es sich um einen echten Angriff handelt oder nicht, kann wesentlich sein, wie das Vertrauensverhältnis zu dem Gegner gerade ist und ob man ihm aktuell einen solchen Angriff zutraut. Maßnahmen, die zu einer Verbesserung des Verhältnisses zwischen den Nuklearmächten führen, könnten so einen Atomkrieg aus Versehen verhindern. Dies kann auch für wissenschaftliche, kulturelle und viele andere Aspekte gelten. Das eigene Engagement bezüglich solcher Aspekte könnte so entscheidend für die Vermeidung eines Atomkriegs und die Bewahrung des Friedens sein, ohne dass man dies jemals erfahren würde. Entsprechendes Engagement lohnt sich also.