

## **Expert:innen drängen auf eine komplexe Systemperspektive zur Bewertung von KI-Risiken**

**Wenn es darum geht, die Potenziale und Risiken von KI zu verstehen, müssen die sozialen Zusammenhänge und deren komplexe Wechselwirkungen berücksichtigt werden, betonen Wissenschaftler:innen in einem *Perspective*. Auch das öffentliche Engagement muss gefördert werden.**

[Wien, 13.11.2024] Künstliche Intelligenz (KI) dringt in immer mehr Aspekte unseres Lebens vor, was bei Expert:innen zunehmend Sorgen über die Gefahren hervorruft. Manche Risiken sind unmittelbar, andere zeigen sich erst nach Monaten oder sogar Jahren.

Ein kohärenter Ansatz zum Verständnis dieser Risiken fehlt aber noch weitgehend – darauf weisen Wissenschaftler:innen im Journal *Philosophical Transactions A* der Royal Society hin. Sie fordern die Risiken aus der Perspektive komplexer Systeme zu betrachten, um die Gefahren besser einschätzen und damit auch eindämmen zu können – insbesondere angesichts langfristiger Unsicherheiten und komplexer Wechselwirkungen zwischen KI und Gesellschaft.

„Um die Risiken von KI zu verstehen, muss man das komplexe Zusammenspiel zwischen Technologie und Gesellschaft verstehen. Es geht darum, sich zurechtzufinden in miteinander verflochtenen, sich gemeinsam entwickelnden Systemen, die unsere Entscheidungen und unser Verhalten beeinflussen“, sagt [Fariba Karimi](#), Mitautorin des *Perspectives*. Karimi leitet das Forschungsteam für [Algorithmische Fairness am Complexity Science Hub \(CSH\)](#) und ist Professorin für Social Data Science an der [Technischen Universität \(TU\) Graz](#).

„Wir sollten nicht nur darüber diskutieren, welche Technologien wie eingesetzt werden sollen, sondern auch, wie der soziale Kontext angepasst werden kann, um die positiven Möglichkeiten zu nutzen. Zudem sollten die Möglichkeiten und Risiken von KI in Debatten, beispielsweise über die Wirtschaftspolitik, einbezogen werden“, erklärt CSH-Wissenschaftler [Dániel Kondor](#), Erstautor der Studie.

### **UMFASSENDE UND LANGFRISTIGE RISIKEN**

Derzeitige Konzepte zur Risikobewertung konzentrieren sich oft auf unmittelbare und spezifische Gefahren wie Vorurteile und Sicherheitsfragen, so die Autor:innen. „Diese Ansätze übersehen oft umfassendere und langfristige systemische Risiken, die sich aus dem weit verbreiteten Einsatz von KI-Technologien und ihrer Interaktion mit dem sozialen Kontext, in dem sie eingesetzt werden, ergeben könnten“, sagt Kondor.

„In unserem Artikel haben wir versucht, die kurzfristige mit der langfristigen Betrachtung der Auswirkungen dieser Technologien auf die Gesellschaft in Einklang zu bringen. Es geht darum, sowohl die unmittelbaren als auch die systemischen Folgen von KI zu erfassen“, fügt er hinzu.

### **ALGORITHMEN IN DER SCHULE**

Um die potenziellen Risiken von KI-Technologien zu veranschaulichen, bedienen sich die Wissenschaftler:innen einer Fallstudie, die zeigt, wie in Großbritannien, während der Covid-19-Pandemie, ein Vorhersagealgorithmus für Schulprüfungen eingesetzt wurde. Diese neue Lösung sollte objektiver und damit gerechter sein [als die Lehrer:innen danach zu fragen, die Leistungen ihrer Schüler:innen vorherzusagen], da sie auf einer statistischen Analyse der Leistungen der Schüler:innen in den vergangenen Jahren beruhte, heißt es in der Studie.

Als der Algorithmus jedoch in der Praxis umgesetzt wurde, traten mehrere Probleme auf. „Sobald der Benotungsalgorithmus angewandt wurde, traten Ungleichheiten deutlich zutage“, berichtet Valerie

Hafez, eine unabhängige Forscherin und Mitautorin der Studie. „Schüler:innen aus benachteiligten Bevölkerungsgruppen trugen die Hauptlast des vergeblichen Versuchs, der Noteninflation entgegenzuwirken, aber auch insgesamt erhielten 40 % der Schüler:innen schlechtere Noten, als sie es angemessenerweise erwarten hätten können.“

Hafez berichtet, dass viele Antworten im Konsultationsbericht darauf hindeuten, dass das von den Lehrkräften als signifikant empfundene Risiko – die langfristige Auswirkung einer schlechteren als der verdienten Benotung – sich von dem Risiko unterscheidet, das die Entwickler:innen des Algorithmus wahrnahmen. Letztere waren besorgt über die Noteninflation, den daraus resultierenden Druck auf das Hochschulsystem und das mangelnde Vertrauen in die tatsächlichen Fähigkeiten der Schüler:innen.

## **DIE SKALIERUNG UND DER UMFANG**

Dieser Fall zeigt mehrere wichtige Probleme auf, die sich beim Einsatz groß angelegter algorithmischer Lösungen ergeben, betonen die Wissenschaftler:innen. „Eine Sache, auf die man unserer Meinung nach achten sollte, ist die Skalierung – und auch der Umfang. Denn Algorithmen lassen sich leicht von einem Anwendungsbereich auf nächsten übertragen, auch wenn diese Kontexte sehr unterschiedlich sein können. Der ursprüngliche Entstehungskontext verschwindet nicht einfach, er überlagert vielmehr die anderen Kontexte“, erklärt Hafez.

„Langfristige Risiken sind nicht die lineare Kombination von kurzfristigen Risiken. Sie können im Laufe der Zeit exponentiell ansteigen. Mit Computermodellen und Simulationen können wir jedoch praktische Erkenntnisse gewinnen, um diese dynamischen Risiken besser zu bewerten“, fügt Karimi hinzu.

## **COMPUTERGESTÜTZTE MODELLE – UND DIE ÖFFENTLICHKEIT**

Dies ist eine der von den Wissenschaftler:innen vorgeschlagenen Richtungen für das Verständnis und die Bewertung der mit KI-Technologien verbundenen Risiken, sowohl kurz- als auch langfristig. „Computermodelle – wie jene, die die Auswirkungen der KI auf die Repräsentation von Minderheiten in sozialen Netzwerken bewerten – können zeigen, wie Vorurteile („Biases“) in KI-Systemen zu Rückkopplungsschleifen führen, die gesellschaftliche Ungleichheiten verstärken“, erklärt Kondor. Solche Modelle können zur Simulation potenzieller Risiken verwendet werden und bieten Einblicke, die mit herkömmlichen Bewertungsmethoden nur schwer zu gewinnen sind.

Darüber hinaus betonen die Autor:innen der Studie, wie wichtig es ist, Laien und Expert:innen aus verschiedenen Bereichen in den Prozess der Risikobewertung einzubeziehen. Kompetenzgruppen – kleine, heterogene Teams, die unterschiedliche Perspektiven vereinen – können ein wichtiges Instrument sein, um die demokratische Beteiligung zu fördern und sicherzustellen, dass die Risikobewertungen unter Mitwirkung der am stärksten von den KI-Technologien Betroffenen entstehen.

„Ein allgemeinere Herausforderung ist die Förderung der sozialen Resilienz, die dazu beitragen kann, Debatten und Entscheidungsprozesse rund um KI zu verbessern und mögliche Fallstricke zu vermeiden. Diese soziale Widerstandsfähigkeit kann von vielen Fragen abhängen, die mit KI wenig oder nur indirekt zu tun haben“, überlegt Kondor. Die Förderung partizipativer Formen der Entscheidungsfindung kann eine wichtige Komponente zur Erhöhung der Resilienz sein.

„Ich denke, wenn man anfängt, KI-Systeme als soziotechnisch zu betrachten, kann man die Menschen, die von den KI-Systemen betroffen sind, nicht von den 'technischen' Aspekten trennen. Wenn man sie vom KI-System trennt, nimmt man ihnen die Möglichkeit, die ihnen auferlegten Klassifizierungsinfrastrukturen mitzugestalten, und verweigert den Betroffenen die Macht, an der Gestaltung von Welten teilzuhaben, die auf ihre Bedürfnisse abgestimmt sind“, sagt Hafez, die als AI Policy Officer im österreichischen Bundeskanzleramt tätig ist. \*

\* Valerie Hafez ist Referentin im österreichischen Bundeskanzleramt, hat diese Forschung jedoch unabhängig durchgeführt. Die in dem Perspective geäußerten Ansichten spiegeln nicht notwendigerweise die Ansichten oder Positionen des Bundeskanzleramtes wider.

## Service

Die Studie „[Complex systems perspective in assessing risks in A.I.](#)“ von D. Kondor, V. Hafez, S. Shankar, R. Wazir und F. Karimi wurde in **Philosophical Transactions A** veröffentlicht und ist online verfügbar (doi: [10.1098/rsta.2024.0109](https://doi.org/10.1098/rsta.2024.0109)).

### KONTAKT

Anja Böck  
Head of Communications  
Complexity Science Hub  
[boeck@csh.ac.at](mailto:boeck@csh.ac.at)  
+43 1 59991 601  
Link

### WISSENSCHAFTLICHER KONTAKT

Fariba Karimi  
Group Lead [Algorithmic Fairness](#)  
Complexity Science Hub  
Professorin für Social Data Science, TU Graz  
[karimi@csh.ac.at](mailto:karimi@csh.ac.at)  
Link



### ÜBER DEN COMPLEXITY SCIENCE HUB

Der Complexity Science Hub (CSH) ist Europas wissenschaftliches Zentrum zur Erforschung komplexer Systeme. Wir übersetzen Daten aus einer Reihe von Disziplinen – Wirtschaft, Medizin, Ökologie, Sozialwissenschaften – in anwendbare Lösungen für eine bessere Welt. Gegründet im Jahr 2016, forschen heute über 70 Wissenschaftler:innen am CSH, getragen von der wachsenden Notwendigkeit für ein fundiertes Verständnis der Zusammenhänge, die unserer Gesellschaft zugrunde liegen – vom Gesundheitswesen bis zu Lieferketten. Mit unseren interdisziplinären Methoden entwickeln wir die Kompetenzen, um Antworten auf heutige und zukünftige Herausforderungen zu finden.

Mitglieder des CSH sind AIT Austrian Institute of Technology, BOKU University, Central European University (CEU), Medizinische Universität Wien, TU Wien, TU Graz, Universität für Weiterbildung Krems, Vetmeduni Wien, Wirtschaftsuniversität Wien (WU) und Wirtschaftskammer Österreich (WKÖ).

**[csh.ac.at](https://csh.ac.at)**