



Übersetzung aus dem Englischen: „G7 Academies’ Statement 2019: “Artificial intelligence and society”.
Kein offizielles G7-Dokument.

Künstliche Intelligenz und Gesellschaft

Zusammenfassung und Empfehlungen

Künstliche Intelligenz (KI) ist eine der Technologien, die unsere Gesellschaft und viele Aspekte unseres täglichen Lebens verändern. KI wird bereits in vielen Bereichen gewinnbringend eingesetzt und hat das Potenzial, zu beträchtlichem wirtschaftlichem Wohlstand beizutragen. Sie wirft aber auch Fragen zu Themen wie Beschäftigung, Datenschutz, Privatsphäre, Verletzung ethischer Werte und Zuverlässigkeit der Ergebnisse auf.

Deshalb sollten sich Politik und Wissenschaft für Folgendes einsetzen:

- Ein verantwortungsvoller Umgang mit KI ist notwendig, damit alle gesellschaftlichen Gruppen vom Nutzen der KI profitieren können. Dazu ist eine sehr genaue Beobachtung der Auswirkungen von KI auf die Beschäftigungssituation erforderlich, die sich durch verschiedene politische, wirtschaftliche und kulturelle Faktoren sowie Fortschritte in der KI-Technologie verändern wird.
- KI-Systeme und -Daten müssen verlässlich sein. Dies sollte durch Maßnahmen zur Sicherstellung von Datenqualität, Vermeidung von Bias und Datenrückverfolgbarkeit ermöglicht werden. Zwar trägt ein besserer Zugang zu den Daten hierzu bei, doch sollten personenbezogene Daten nicht für unbefugte Dritte verfügbar gemacht werden.
- KI-Systeme und -Daten müssen sicher und geschützt sein. Dies ist vor allem bei Anwendungen wichtig, die eine Gefährdung für Menschen darstellen könnten und beweisbar korrekte Systeme erfordern.
- Weitere Forschung ist erforderlich, um erklärbare KI-Systeme zu entwickeln. Wenn KI-Systeme wichtige Entscheidungen mit Auswirkungen auf Menschen vorschlagen, sollten die Betroffenen hinreichend informiert werden und die Möglichkeit erhalten, diese Entscheidungen zu hinterfragen (z.B. eine Behandlung abzulehnen oder eine Entscheidung anzufechten).
- Um den gesellschaftlichen Nutzen von KI zu maximieren, sind Erkenntnisse aus vielen Bereichen erforderlich. Die interdisziplinäre Forschung sollte verschiedene Felder umfassen, wie Naturwissenschaften, Biowissenschaften und Medizin, Ingenieurwissenschaften, Robotik, Geisteswissenschaften, Wirtschafts- und Sozialwissenschaften, Ethik, Informatik und natürlich die KI selbst.
- Die Bürgerinnen und Bürger müssen auf KI vorbereitet sein. Um die künstliche Intelligenz zu entmystifizieren, müsste eine umfangreiche Aufklärung stattfinden, Informationen müssten zugänglich gemacht und ein fundierter Dialog mit den Bürgerinnen und Bürgern geführt werden.
- Gleichzeitig sollte eine öffentliche politische Debatte über die destruktive/militärische Nutzung von KI angestoßen werden. Internationale Verpflichtungen zur Begrenzung der Risiken autonomer Waffen sollten von den zuständigen Gremien der Vereinten Nationen in Betracht gezogen werden.

- Der Talentaustausch und die Zusammenarbeit zwischen der öffentlichen Forschung und dem Privatsektor müssen gefördert werden. Dies würde zu einer sichereren und schnelleren Bereitstellung von Anwendungen in Bereichen mit hohem Nutzenpotenzial für die Menschen beitragen. Eine solche Zusammenarbeit ist für die großangelegte Erfassung von Daten wichtig, die für die Entwicklung von KI-Systemen entscheidend ist.

Einleitung

Künstliche Intelligenz (KI) bezeichnet Methoden und Technologien, mit denen Computer oder andere Geräte dazu gebracht werden sollen, intelligent zu funktionieren. Im Wesentlichen ist KI eine Sammlung von Algorithmen, die auf der Grundlage von (meist großen) Datenbeständen arbeiten. Maschinelles Lernen (ML) ist eine Unterform der KI, bei der Algorithmen nützliche Informationen aus komplexen Daten extrahieren. In den letzten Jahren haben sich ML-Anwendungen in vielen Bereichen von Wissenschaft und Technik stärker als erwartet durchgesetzt. Es besteht ein breiter Konsens darüber, dass die KI-Forschung immer weiter fortschreitet und in Zukunft wahrscheinlich immer größere Auswirkungen auf die Gesellschaft haben wird.

Die Entwicklung anspruchsvoller algorithmischer Systeme, kombiniert mit der Verfügbarkeit von Daten und Rechenleistung, hat in verschiedenen Spezialgebieten wie Spracherkennung, Bildklassifizierung, Fehlererkennung, autonomes Fahren, Entscheidungsunterstützungssysteme, Robotik, maschinelle Übersetzung, Fortbewegung auf Beinen und Fragebeantwortungssysteme beachtliche Erfolge erbracht. Einige dieser Anwendungen bieten höchst wertvolle Unterstützung für Menschen mit Behinderungen. Über Gehirn-Maschine-Schnittstellen können Gelähmte mit Hilfe eines Computers mit ihrer Umgebung interagieren.

Auch in den Natur- und Sozialwissenschaften ermöglichen ML-Algorithmen Fortschritte und bieten neue Werkzeuge für die Handhabung und Modellierung komplexer Daten und Prozesse mit hohem potenziellem Nutzen. Da ein Großteil der Errungenschaften unserer Zivilisation durch menschliche Intelligenz geschaffen wurde, lässt sich nur erahnen, was alles erreicht werden könnte, wenn diese Intelligenz durch die von KI ermöglichten Instrumente um ein Vielfaches verstärkt würde.

Doch es gibt auch eine Reihe von Fragen und Bedenken zu möglichen Gefahren, die weitere Überlegungen erfordern.

Der Fortschritt in der KI-Forschung bringt es mit sich, dass nicht nur danach gefragt werden muss, wie KI leistungsfähiger gemacht werden kann, sondern auch danach, wie ihr gesellschaftlicher Nutzen unter Wahrung ethischer Werte maximiert werden kann. Der Einsatz und die technische Entwicklung von KI sollte daher ethischen Grundsätzen folgen. Da KI auf statistischen Datenanalysen und maschinellem Lernen basiert, gibt es zunehmend Bedenken, KI-Systeme könnten zu Bias führen.

Vor diesem Hintergrund gehen wir zunächst auf die Probleme ein, die sich durch die transformative Wirkung von KI auf die Wirtschaft ergeben. Anschließend wird es um die allgemeinen Eigenschaften gehen, die KI-Systeme aufweisen sollten, um zufriedenstellend und ethisch einwandfrei mit Menschen interagieren zu können. Schließlich werden spezifischere Fragen zum Einsatz von KI-Systemen im Gesundheitswesen und zu möglichen KI-Anwendungen für autonome Waffensysteme untersucht, und das Potenzial von KI in Robotersystemen betrachtet. Aus dieser Analyse ergibt sich eine Reihe von Empfehlungen, die in der Zusammenfassung aufgeführt sind.

1. Management und Optimierung der Auswirkungen von KI auf unsere Gesellschaften

Zwischen Wirtschaftswissenschaftlern und Informatikern gibt es einen allgemeinen Konsens darüber, dass die Forschung danach streben muss, die wirtschaftlichen Vorteile der künstlichen Intelligenz zu maximieren und gleichzeitig negative Auswirkungen gering zu halten. Wir müssen heute darüber nachdenken, welche Folgen KI im Hinblick auf zunehmende Ungleichheit, Arbeitslosigkeit und unethische Verhaltensweisen haben kann. Auf diese offenen Fragen wird im Weiteren ausführlicher eingegangen.

1.1 Prognosen für den Arbeitsmarkt

Der Einsatz von künstlicher Intelligenz könnte sektorübergreifend signifikante wirtschaftliche Vorteile mit sich bringen. KI-Technologien bergen ein hohes Potenzial für Produktivitätssteigerungen und neuartige Erzeugnisse und Dienstleistungen. Daran schließen sich Fragen zu den Auswirkungen an, die KI auf die Beschäftigung und das Arbeitsleben haben wird.

Man kann davon ausgehen, dass der Einsatz von KI zu einer Disruption in der Arbeitswelt führen wird – bestimmte Arbeitsplätze werden wegfallen, andere neu entstehen, wieder andere werden sich verändern. Studien mit Prognosen zu den Folgen von KI für die Beschäftigung weisen hohe Unsicherheiten auf, was die Änderungsraten und den Anteil der wahrscheinlich von der Automatisierung betroffenen Tätigkeiten oder Arbeitsplätze angeht.

Längerfristig tragen diese Technologien zu mehr Produktivität und Wohlstand für die Bevölkerung bei. Allerdings kann es sein, dass die Entfaltung dieser Vorteile einige Zeit in Anspruch nimmt; bis dahin kann die Entwicklung für bestimmte Bevölkerungsgruppen Nachteile mit sich bringen. Es könnte also für manche Menschen oder Orte durchaus zu negativen Übergangseffekten und kurzfristig zu einem Anstieg sozialer Ungleichheit kommen. Die wirtschaftlichen und gesellschaftlichen Auswirkungen einer solchen Disparität müssen ganz offensichtlich eingehender erforscht werden, vor allem auch im Hinblick auf die Gefährdung von Arbeitsplätzen durch die Automatisierung. Vermutlich ist es einfacher, die Auswirkungen von KI-Systemen auf verschiedene Gruppen von Arbeitnehmern zu analysieren, z.B. auf solche mit geringer Qualifizierung und auf gut ausgebildete Fachkräfte, als vorherzusagen, welche Arbeitsplätze möglicherweise im Rahmen verschiedener Strategien zukünftig geschaffen werden. Für die zukünftige Entwicklung von KI-Technologien sind mehrere Szenarien vorstellbar. Viele Faktoren – u.a. politische, wirtschaftliche und kulturelle Aspekte sowie die KI-technischen Möglichkeiten – werden die Auswirkungen von KI auf die Arbeitswelt mitgestalten. Anhand der besten verfügbaren interdisziplinären Forschungsergebnisse können Strategien erarbeitet werden, die den technologischen Wandel für die gesamte Gesellschaft zu einem Gewinn machen.

1.2 Strategien für das Management der KI-Entwicklung und ihre Integration in die Gesellschaft

KI wird zahlreiche Bereiche der Gesellschaft drastisch verändern, indem sie menschliche Arbeit verbessert oder ersetzt. Die Herausforderung besteht darin, diese Veränderungen vorherzusehen und Strategien zu entwickeln, die die negativen Auswirkungen begrenzen und eine bessere Integration von KI ermöglichen. Bildung ist sowohl für die Akzeptanz von KI als auch für die Bekämpfung von Ungleichheit von zentraler Bedeutung.

Ein Grundverständnis der Nutzung von Daten und KI-Technologien ist erforderlich, und zwar nicht nur bei den Herstellern und professionellen Anwendern von KI, sondern bei allen Bürgerinnen und Bürgern jeden Alters. Eine Vermittlung der zentralen Konzepte im Schulunterricht kann dazu beitragen. Durch Einführung eines breit angelegten und ausgewogenen Lehrplans für Naturwissenschaften, Mathematik, Informatik, Kunst und Geisteswissenschaften könnten junge Menschen mit wichtigen Kompetenzen ausgestattet werden, sodass eine solidere Grundlage für lebenslanges Lernen geschaffen würde.

Darüber hinaus besteht eine große Nachfrage nach hochqualifizierten Arbeitskräften. In zahlreichen Branchen und Berufen sind Kompetenzen gefragt, damit KI in gewinnbringender Weise genutzt werden kann. Mit neuen Initiativen könnte ein Pool fachlich versierter Benutzer von KI-Systemen geschaffen werden. Auch müssen neuartige Ausbildungswege und Infrastrukturen unterstützt werden, um fortgeschrittene KI-Kapazitäten aufzubauen, sodass neue Einsatzgebiete erschlossen und viele neue Arbeitsplätze geschaffen werden können.

Diese Anliegen wurden bereits in der Erklärung von Ottawa – „Realizing our digital future and shaping its impact on knowledge, industry, and the workforce“ – anlässlich des letzten Treffens der G7-Wissenschaftsakademien formuliert. Die Regierungen werden aufgerufen, inklusive Strategien umzusetzen, um allen Bürgerinnen und Bürgern zu ermöglichen, gleichberechtigt von den Vorteilen der KI zu profitieren. Dies setzt voraus, dass Qualität, Sicherheit und Belastbarkeit von Informationen garantiert sind, ebenso wie Transparenz, Offenheit und Interoperabilität der KI-Systeme.

In Bereichen, in denen geltende Regelungen durch die Möglichkeiten der KI überholt wurden, können neue Governance-Ansätze notwendig sein, die ethische Fragestellungen berücksichtigen, die sich aus der Interaktion von Menschen mit intelligenten Maschinen ergeben. Bei der Erforschung der Risiken, die die KI für bestehende ethische Normen birgt, bzw. der Frage, welche neuen ethischen Herausforderungen sich durch KI stellen können, sollte die Rolle der Geistes- und Sozialwissenschaften im Allgemeinen und in der Partnerschaft mit Entwicklern und Nutzern unbedingt mehr Gewicht erhalten.

2. Merkmale von KI-Systemen, die gefördert werden sollten

2.1 Daten

Ob wir die Synergieeffekte zwischen KI und Big Data in vollem Umfang nutzen können, wird zum Teil davon abhängen, wie gut wir in der Lage sind, Daten zu übernehmen, kritisch zu beurteilen und zu verwalten. Viele der heutigen KI-Systeme erfordern Zugang zu riesigen Datenmengen. Um diese Technologie umfassend nutzen zu können, brauchen wir möglicherweise neue Rahmenbedingungen für die Verfügbarmachung von Daten. Dies gilt insbesondere für offene Daten und für persönliche Daten von öffentlichem Interesse, bei denen möglicherweise neue Standards erforderlich sind, um sicherzustellen, dass Daten effektiv verwendet werden können. Dazu gehört, die Bedeutung der Daten deutlich zu machen und den Kontext ihrer Herkunft sowie Informationen zu ihrem Ursprung und ihrer Verarbeitung offenzulegen. All diese Aufgaben können mithilfe von KI bewältigt werden, der somit eine wichtige Rolle zukommen kann, wenn es darum geht, die hohen Erwartungen an offene Daten zu erfüllen und die Interoperabilität verschiedener Datentypen (z.B. soziale, wirtschaftliche, organisatorische und technische Daten) sicherzustellen.

Gleichzeitig sind beim Zugang zu hochwertigen Datenbeständen die Privatsphäre und die Vertraulichkeit personenbezogener Daten zu schützen, und Bedenken wegen unfairer Bias und individueller Rechte muss Rechnung getragen werden. Es sollte alles dafür getan werden, den Zugang von Dritten wie Banken, Versicherungen und potenziellen Arbeitgebern zu vertraulichen Daten durch Vorschriften zu regulieren. Datenbestände müssen gegen böswillige Angriffe geschützt werden. Richtlinien für die Erfassung und den Austausch von sowie den Zugang zu Daten sollten nicht nur für große Konzerne, sondern auch für Open-Source-Initiativen gelten.

2.2 Leistungsfähigkeit und Erklärbarkeit

Einige der erfolgreichsten und populärsten Entwicklungen künstlicher Intelligenz, insbesondere Deep Learning, haben bisher den Nachteil, dass sie nicht erklärbar sind; unterschiedliche KI-Methoden unterstützen dabei unterschiedliche Arten von Erklärbarkeit. Dies kann das Vertrauen der Nutzer in diese Tools beeinträchtigen. In bestimmten Bereichen gilt Nachvollziehbarkeit als unabdingbar: In medizinischen Anwendungen etwa wäre eine Diagnose ohne Erläuterungen kaum akzeptabel. Die Konflikte zwischen Leistungsfähigkeit und Erklärbarkeit sollten klar benannt werden, und es müssen erklärbarere Modelle entwickelt werden. Die Grenzen der implementierten Algorithmen müssen beschrieben werden, damit Nutzer verstehen, warum KI-Systeme bestimmte Entscheidungen treffen.

Eine bessere Erklärbarkeit der künstlichen Intelligenz kann dazu beitragen, dass KI-Systeme keinen Bias produzieren. Disparate Impact („ungleiche Auswirkungen“) ist der inzwischen vorherrschende rechtliche und theoretische Begriff, um eine unbeabsichtigte Diskriminierung durch die Anwendung von Algorithmen zu kennzeichnen, bei der sich ein persönliches Attribut (Ethnizität, soziale Herkunft, Geschlecht oder Alter) unmittelbar auf die Entscheidungsfindung des Algorithmus auswirkt. Wenn KI-Systeme für Entscheidungen eingesetzt werden, die tief greifende Folgen für das tägliche Leben von Menschen haben, sollten sie keine unerwünschten ungleichen Auswirkungen generieren.

2.3 Überprüfung und Validierung von Systemen, die sich im Internet weiterentwickeln

Systeme, die sich im Internet weiterentwickeln, verändern sich im Lauf der Zeit anhand der kontinuierlich erfassten Daten. In jüngster Zeit ist klar geworden, dass ein KI-System ungewollt von seinem ursprünglichen Kurs abweichen kann, beispielsweise in Bezug auf Merkmale wie Geschlecht und Ethnizität. Die Entscheidungen von Systemen, die sich im Internet weiterentwickeln, müssen also überwacht werden, um unerwünschte Entwicklungen zu erkennen.

3. Typische Anwendungsgebiete und Folgen für die Gesellschaft

3.1 Gesundheitswesen

Im Bereich Gesundheit und Pflege bieten KI-Systeme einen signifikanten potenziellen Nutzen. In diesem Sektor können strukturelle Probleme zu Diagnosefehlern, zum Versagen von Fachkompetenz und zu einem mangelhaften Informationsaustausch zwischen Forschung, Technik und Kliniken führen. Mit künstlicher Intelligenz können große Mengen an Publikationen aus der Forschung ausgewertet, unwahrscheinliche und schwache Korrelationen in riesigen Datenbeständen erkannt, Bilder und andere von medizinischen Systemen produzierte Daten analysiert und neue Technologien entwickelt werden. Weil die Verbesserung klinischer Entscheidungsunterstützungssysteme so wichtig ist, kann KI hier Wesentliches leisten, da sie Ärzten Instrumente und Geräte bietet, die sie bei der Diagnosestellung und Therapieentscheidung unterstützen und ergänzen. Ziel ist es, die Interpretation von Beobachtungen und Messungen zu verbessern, exaktere Diagnosen zu stellen und eine genauere, effektivere und besser zugängliche medizinische Versorgung zu ermöglichen. Dies erfordert ein sorgfältiges Systemdesign, bei dem berücksichtigt wird, wie die KI mit menschlichen Nutzern zusammenarbeiten kann, welche Art von Interpretierbarkeit in verschiedenen Kontexten möglicherweise erforderlich ist und wie solche Systeme überprüft und validiert werden können. Es ist wichtig, dass Ärzte und Patienten Vertrauen in solche Systeme haben und dass diese Systeme für verschiedene Nutzergruppen gut funktionieren.

Eine sorgfältige Datenverwaltung ist ebenfalls erforderlich. Um die Fortschritte durch künstliche Intelligenz zu beschleunigen, ist eine grenzüberschreitende Zusammenarbeit für alle Menschen weltweit von Interesse.

3.2 Autonome Waffen

KI eröffnet neue Möglichkeiten für militärische Anwendungen, insbesondere im Hinblick auf Waffensysteme mit erheblicher Autonomie bei den kritischen Funktionen der Auswahl und des Angriffs von Zielen. Solche Waffen könnten zu einem neuen Wettrüsten führen, die Hemmschwelle für Kriege herabsetzen oder zu einem Werkzeug für Diktatoren oder Terroristen werden. Einige Organisationen fordern ein Verbot autonomer Waffen, ähnlich den Abkommen im Bereich der chemischen oder biologischen Waffen. Ein solches Verbot würde eine exakte Definition von Waffen und Autonomie voraussetzen. Solange es kein Verbot von tödlichen autonomen Waffen (Lethal Autonomous Weapons Systems, LAWS) gibt, sollte die Konformität jedes Waffensystems mit dem humanitären Völkerrecht gewährleistet werden. Autonome Waffen sollten so in bestehende Befehls- und Kontrollstrukturen integriert werden, dass Verantwortung und rechtliche Haftung weiterhin den entsprechenden menschlichen Akteuren zugeordnet bleiben. Ganz offensichtlich bedarf es hier größerer Transparenz und einer öffentlichen Debatte der in diesem Bereich aufgeworfenen Fragen.

3.3 Robotik

Roboter sind Maschinen mit integrierter künstlicher Intelligenz, die Dinge wahrnehmen und sich bewegen können. Der physische Kontakt zwischen Maschine und Umwelt, also auch zwischen Maschine und Mensch, ist eine Herausforderung. Roboter müssen sicher und zuverlässig sein. Bis vor Kurzem wurden Roboter hauptsächlich in der Produktion genutzt und waren auf spezifische Einsatzgebiete beschränkt, in denen sie nicht mit Menschen zusammenkamen. Heute, nach der zweiten Welle der Robotikentwicklung, können sich Roboter zunehmend mit Menschen in derselben Umgebung aufhalten und mit ihnen interagieren. Während KI-Anwendungen auf Technologien konzentriert sind, bei denen Daten verarbeitet werden, um Wissen für eine Entscheidungsfindung abzuleiten, besteht das oberste Ziel der Robotik darin, technische Systeme zu schaffen, die mit der physischen Welt interagieren können.

Die Robotik nutzt nicht nur Algorithmen des maschinellen Lernens, sondern ist darüber hinaus mit grundlegenden Einschränkungen im Hinblick auf die physische Sicherheit konfrontiert. Die Entwicklung von Robotern erfordert Softwarezertifizierung und formelle Überprüfung, um die Fehlertoleranz, Zuverlässigkeit und Überlebensfähigkeit zu maximieren.

Trotz des beachtlichen Fortschritts der letzten Zeit wird das Tempo des technologischen Wandels oft überschätzt.

Schließlich werden die landläufigen Vorstellungen von künstlicher Intelligenz im Allgemeinen und von Robotik im Besonderen wohl stärker von Fantasy-Narrativen als von wissenschaftlichen Erkenntnissen beeinflusst. Es ist von zentraler Bedeutung, die Forschungsgebiete von Robotik und KI zu entmystifizieren und das Wissen darüber stärker zu verbreiten – durch öffentliche Aufklärung und durch Diskussionen und Debatten mit allen Bürgerinnen und Bürgern.