

sustain

Nachhaltige KI in der Praxis



Künstliche Intelligenz:

So lässt sie sich nachhaltiger gestalten

Überblick:

Nachhaltigkeitskriterien

Wie sich die Nachhaltigkeit von KI mithilfe von 13 Kriterien bewerten lässt

Grüne Rechenzentren:

Wärme statt Überhitzung

Wassergekühlte Server lassen die für KI notwendige Energie nicht verpuffen

Standpunkt:

Regulierung

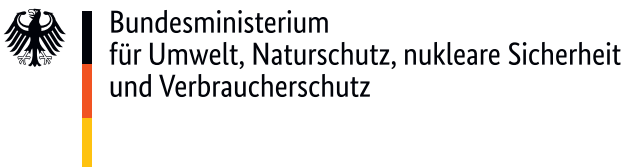
Wir brauchen politische Lösungen für die Risiken von KI



Projektpartner:



Gefördert vom Bundesministerium für Umwelt, Naturschutz,
nukleare Sicherheit und Verbraucherschutz (BMUV) aufgrund
eines Beschlusses des Deutschen Bundestages



Liebe Leser*innen,

eine Zahl taucht immer wieder in der Diskussion zur Nachhaltigkeit von KI auf: Die Emissionen, die in der Forschungs- und Entwicklungsphase großer Sprachmodelle anfallen, entsprechen den Emissionen von fünf Autos während ihrer gesamten Laufzeit. Diese Zahl stammt aus einer wegweisenden Analyse von Emma Strubell, Ananya Ganesh und Andrew McCallum. Wenig hilfreich sind jedoch die vielen ungenauen oder falschen Aussagen, die sich seitdem auf diese Berechnung beziehen – oft mit zweifelhaften Aussagen über die CO₂-Kosten von KI-Systemen.

Die Diskussion über die Nachhaltigkeit von KI verdient mehr Genauigkeit, mehr Differenziertheit, mehr Analyse und mehr Fakten. KI hat ökologische, soziale und wirtschaftliche Kosten zur Folge. Wissenschaft, Industrie, Zivilgesellschaft und Politik müssen dringend auf der Grundlage von Fakten über diese Kosten sprechen und handeln. Mit dem ersten SustAIn-Magazin wollen wir unseren Beitrag zu solch einem Austausch leisten.

Das Magazin ist das Ergebnis des Forschungsprojekts SustAIn, in dem wir einen Rahmen zur Bewertung der Nachhaltigkeit von KI-Systemen entwickelt haben. Hier stellen wir Ihnen vor, wie eine nachhaltigere KI in der Praxis bereits umgesetzt wird. Wir brauchen Good-Practice-Beispiele, methodische Innovationen und differenzierte Perspektiven darauf, wo Forschung und politisches Handeln notwendig sind, um nachhaltigere KI zu fördern.

Ich wünsche Ihnen eine anregende Lektüre!

Dr. Anne Mollen

Projektmanagerin „SustAIn:

Der Nachhaltigkeitsindex für Künstliche Intelligenz“

AlgorithmWatch



Überblick:

SustAI:n: Der Nachhaltigkeitsindex für Künstliche Intelligenz

Was ist nachhaltige KI und wie ist es um sie bestellt? 6

Die Nachhaltigkeit von KI bewertbar machen

13 Kriterien, die zu berücksichtigen sind, damit KI nachhaltiger wird 7

Glossar:

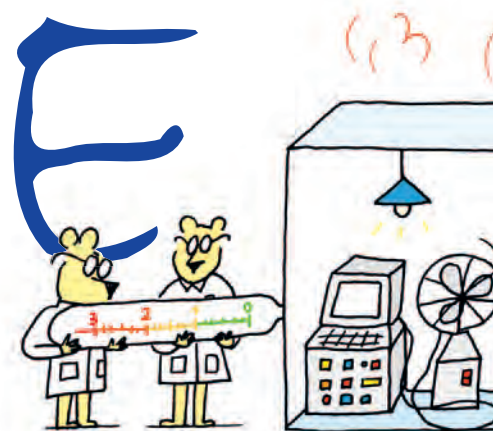
Nachhaltige KI 101

Zentrale Begriffe in der Diskussion um die Nachhaltigkeit von KI 10

Zahlen & Fakten:

AI Index Report 2022

Daten zu KI, die aus Nachhaltigkeitsperspektive relevant sind 14



Methoden für nachhaltige KI:

KI als Gemeinschaftsaufgabe denken

Im Gespräch mit der Soziologin und KI-Ethikerin Alex Hanna, Forschungsleiterin am DAIR-Institut 18

Ressourcenverbrauch:

Monströse Kekse: Cookies als Instrument einer übergriffigen Technologie

Im Gespräch mit der Künstlerin Joana Moll: AdTech und ihr Projekt Carbolytics 22

KI in der Praxis:

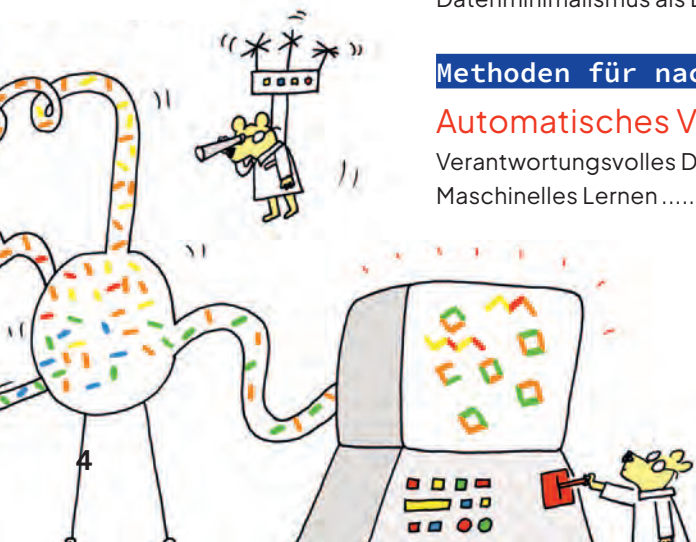
Weniger bringt mehr: Warum Künstlicher Intelligenz eine Datendiät guttun kann

Im Gespräch mit der Computerlinguistin Michaela Regneri: Datenminimalismus als Effizienzstrategie 26

Methoden für nachhaltige KI:

Automatisches Vergessen

Verantwortungsvolles Datenmanagement und Maschinelles Lernen 30





Methoden für nachhaltige KI:

KI sollte nur machen, was wir wollen

Im Gespräch mit der Informatikerin Julia Stoyanovich,
Gründerin des Center for Responsible AI 32

Comic:

Data, Responsibly (Vol. 1)

Welchen Einfluss haben KI und Algorithmen auf uns?
Ein Comic klärt auf..... 36

KI in der Praxis:

Grüne Datenbanken für Algorithmen: der Green Consumption Assistant

Durch Maschinelles Lernen im Online-Handel
nachhaltigen Konsum fördern 38



Ressourcenverbrauch:

Gefangen in einer untragbaren Infrastruktur

Im Gespräch mit der KI-Ethikerin Aimee van Wynsberghe:
KI und Ausbeutung 40

Standpunkt:

Ein gefährliches Ausblenden von Umweltrisiken – Die KI-Verordnung der EU

Die EU scheint beim Ressourcenverbrauch und den
Umweltfolgen von KI blind zu sein 44

Standpunkt:

Eine ökologisch nachhaltige KI braucht Regulierung

Wie lassen sich KI-Lösungen möglichst energieeffizient
entwickeln, steuern und nutzen? 46

KI in der Praxis:

KI aus dem Baukasten

Im Gespräch mit Gregor Blichmann vom Unternehmen elevait:
Größere Energieeffizienz durch modulare KI-Produkte 48

KI in der Praxis:

Wärme statt Überhitzung: Wie Rechenzentren zu nachhaltigen Heizkörpern werden

Im Gespräch mit Ronny Reinhardt von Cloud&Heat:
Nachhaltigere Cloud-Infrastruktur 52

Ausblick:

SustAIIn geht weiter.

In der nächsten Ausgabe: Nachhaltiges Coden 57



Projektpartner:

AlgorithmWatch

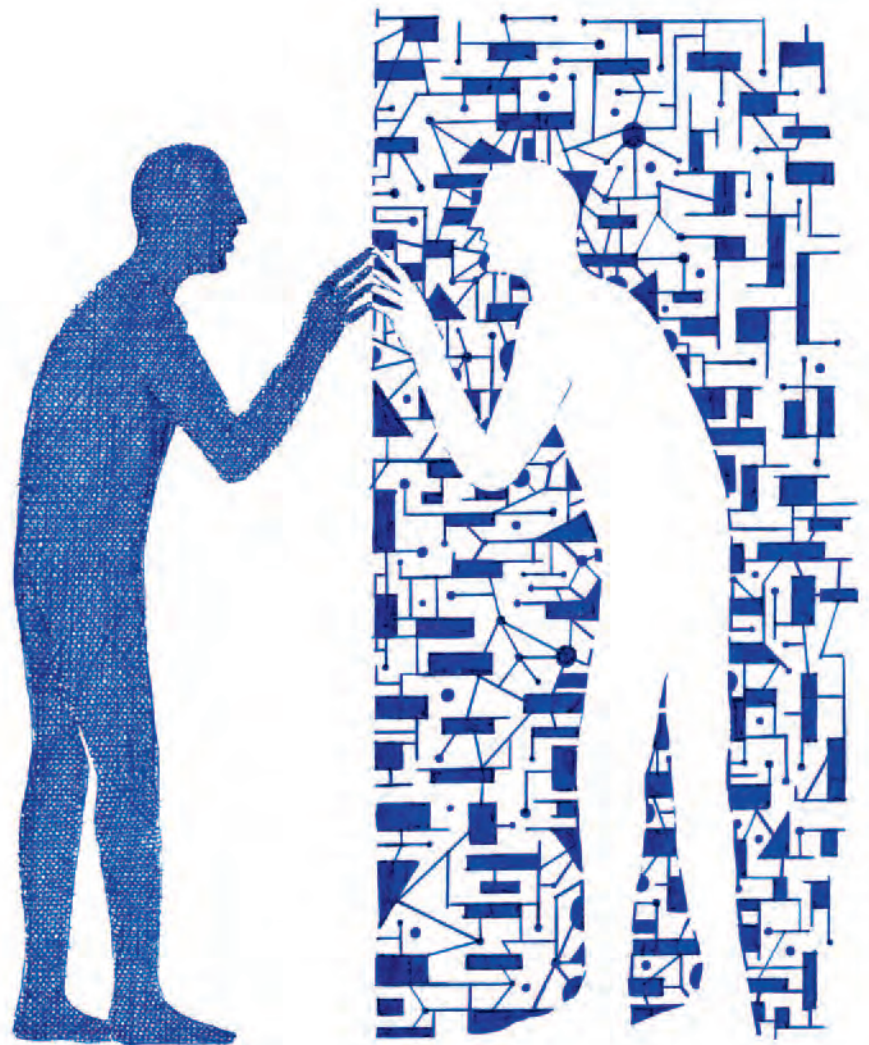
AlgorithmWatch ist eine gemeinnützige Forschungs- und Advocacy-Organisation mit dem Ziel, Systeme automatisierter Entscheidungsfindung (ADM) und deren Auswirkungen auf die Gesellschaft zu beobachten und zu analysieren. Eine Säule unserer Arbeit ist die Auseinandersetzung mit ADM-Systemen und Nachhaltigkeit.

Institut für ökologische Wirtschaftsforschung

Das IÖW ist ein führendes wissenschaftliches Institut auf dem Gebiet der praxisorientierten Nachhaltigkeitsforschung. Wir erarbeiten Strategien und Handlungsansätze für ein zukunftsfähiges Wirtschaften – für eine Ökonomie, die ein gutes Leben ermöglicht und die natürlichen Grundlagen erhält. Seit über 30 Jahren beschäftigen wir uns mit Zukunftsfragen und finden immer wieder neue, oft auch ungewöhnliche Antworten.

Distributed Artificial Intelligence Labor

Das DAI-Labor an der TU Berlin sieht sich als Mittler zwischen universitärer Forschung und industrieller Verwertung. Mit unserem interdisziplinären Team erzeugen wir Innovationen und überführen universitäre Forschung in Anwendungen des täglichen Lebens. Dies geschieht in enger Kooperation mit anderen wissenschaftlichen und industriellen Institutionen.



SustAI:n:

Der Nachhaltigkeitsindex für Künstliche Intelligenz

Was ist nachhaltige KI und wie ist es um sie bestellt? Dieser Frage gehen wir beim Projekt *SustAI:n: Der Nachhaltigkeitsindex für Künstliche Intelligenz* nach. Ein interdisziplinäres Team von AlgorithmWatch, dem Institut für ökologische Wirtschaftsforschung und dem DAI-Labor an der TU Berlin analysiert, wie sich nachhaltige KI definieren und umsetzen lässt. Das Projekt wird vom *Bundesministerium für Umwelt, Naturschutz nukleare Sicherheit und Verbraucherschutz (BMUV)* im Rahmen der Initiative „KI-Leuchttürme für Umwelt, Klima, Natur und Ressourcen“ gefördert.

Nachhaltigkeit von KI bewertbar machen

Nachhaltige KI respektiert die planetaren Grenzen, verstärkt keine problematischen ökonomischen Dynamiken und gefährdet nicht den gesellschaftlichen Zusammenhalt. Im Projekt SustAIIn haben wir auf dieser Grundlage 13 Kriterien definiert, die Organisationen berücksichtigen sollten, um KI nachhaltiger zu produzieren und einzusetzen.



Soziale Nachhaltigkeit

Sozial nachhaltige Entwicklung und Anwendung von Künstlicher Intelligenz stellt den Menschen, die Gesellschaft und gerechte Lebensverhältnisse in den Mittelpunkt.

Damit Menschen ein würdiges Leben führen können, müssen grundlegende Bedürfnisse wie die Nahrungsvorsorgung oder ein ausreichender Wohnraum erfüllt sein. Zugleich müssen sie Zugang zu Infrastrukturen wie

Strom, Wasser oder zum Internet haben. Eine sozial nachhaltige Gesellschaft erlaubt es den Menschen, sich frei zu entfalten. Dem Befähigungsansatz des Ökonomie-Nobelpreisträgers Amartya Sen und der Moralphilosophin Martha Nussbaum zufolge muss eine nachhaltige gesellschaftliche Entwicklung den Menschen Ver-

wirklichungschancen bieten. Sie müssen auf ein Fundament von materiellen und kulturellen Handlungsressourcen zurückgreifen können, um ihre Rechte wahrzunehmen.

Auch KI-Systeme müssen die Würde des Menschen wahren. Sie dürfen niemanden ausschließen, benachteiligen oder diskriminieren und die menschliche Autonomie und Handlungs-

Praxisbeispiele für die Nachhaltigkeitskriterien? Achten Sie auf die Infoboxen im gesamten Magazin.

freiheit nicht einschränken. Werte wie Gerechtigkeit, Inklusion oder Freiheit müssen im Design, bei der Entwicklung und in der Anwendung von KI einbezogen werden. Insbesondere sollte auch die Fähigkeit, auf menschliche Art und Weise zu denken, zu argumentieren und zu handeln, nicht durch die Systeme eingeschränkt werden.



Transparenz und Verantwortungsübernahme

Wer KI nutzt oder mit ihr interagiert, sollte vorab darüber informiert worden sein, dass KI eingesetzt wird, und die daraus hervorgehenden Ergebnisse nachvollziehen können. Hierzu müssen zentrale Informationen zu KI-Systemen offengelegt werden und Verantwortlichkeiten für deren Ergebnisse geklärt sein.



Nicht-Diskriminierung und Fairness

Bei der KI-Entwicklung und -Anwendung sollte ein Bewusstsein für Fairness vorhanden sein. Außerdem sollte die KI regelmäßig auf mögliche Diskriminierungen hin überprüft werden.



Technische Verlässlichkeit und menschliche Aufsicht

Schwachstellen in KI-Systemen sollten systematisch über Risikobewertungen identifiziert werden. Zudem sollte eine hohe Datenqualität sichergestellt sein. Systemeingriffe durch Menschen sollten ermöglicht werden.



Selbstbestimmung und Datenschutz

Kleine Datensätze, Verschlüsselung oder ein Widerspruchsrecht bei der Verwendung personenbezogener Daten stärken neben weiteren Maßnahmen die informationelle Selbstbestimmung und den Datenschutz.



Inklusives und partizipatives Design

Endnutzer*innen, Betroffene und weitere Stakeholder*innen sollten am Designprozess der KI beteiligt werden. Die Teams in der KI-Planung und -Entwicklung sollten divers und interdisziplinär aufgestellt sein.



Kulturelle Sensibilität

Der Anwendungskontext einer KI muss in deren Entwicklung berücksichtigt werden. KI-Systeme sollten daher anpassbar und neu trainierbar sein, wenn sie in unterschiedlichen kulturellen Kontexten eingesetzt werden. Es sollte insbesondere auf lokale Wissensbestände und Datensätze zurückgegriffen werden.



Ökologische Nachhaltigkeit

KI wird häufig als wichtiges Werkzeug zur Bewältigung der Klimakrise gesehen. Die Einsatzmöglichkeiten von KI-Systemen sind breit gefächert: Sie sollen den Ressourcenverbrauch effizienter gestalten, eine effizientere Verkehrs- und Stadtplanung herbeiführen, ein nachhaltigeres Energiesystem kreieren oder auch die Erforschung neuer Materialien erleichtern. Für den Einsatz der Systeme sind allerdings seinerseits Rohstoffe notwendig. Zu den Materialien für die Chips und Platinen gehören seltene Erden, die aufwendig abgebaut werden müssen. Die fertige Hardware muss an den Einsatzort transportiert werden. In den Rechenzentren wird Energie aufgewendet, um KI-Systeme zu entwickeln und einzusetzen, gleichzeitig muss die IT-Infrastruktur gekühlt werden, um die Bauteile vor Schäden zu schützen. Die Hardware muss dennoch regelmäßig ausgetauscht werden, was wiederum zu neuem Materialbedarf und zu einer großen Menge von Elektroschrott führt. Eine unsachgemäße Entsorgung kann schlimmstenfalls schädliche Chemikalien freisetzen.

Generell zielt ökologische Nachhaltigkeit darauf ab, die Natur zu erhalten, um künftigen Generationen einen lebenswerten Planeten zu hinterlassen. Die unter anderem von Johan Rockström entwickelten „Belastbarkeitsgrenzen des Planeten“ legen Schwellen fest, deren Überschreiten unumkehrbare Umweltschäden nach sich ziehen würde. Auf viele dieser Belastungsgrenzen wirken KI-Systeme sich direkt oder indirekt aus. KI-Systeme sind das Gegenteil von ökologisch nachhaltig, wenn sie mehr Ressourcen verbrauchen, als durch ihren Einsatz geschont oder sogar reproduziert werden. Neben dem Materialverbrauch für die Hardware sind ihr immenser Energieverbrauch und die damit verbundenen Emissionen ein Hindernis auf dem Weg zur ökologischen Nachhaltigkeit.



Energieverbrauch

Während der KI-Entwicklung sollte die Energieeffizienz erfasst und ggf. durch geeignete Methoden wie beispielsweise die Modellkomprimierung optimiert werden.



CO₂- und Treibhausgasemissionen

Die CO₂-Effizienz kann beispielsweise durch einen nachhaltigen Energiemix, die geeignete Wahl von Trainingszeitpunkt und -standort sowie durch eine Kompensation von anfallenden CO₂-Emissionen erhöht werden.



Nachhaltigkeitspotenziale in der Anwendung

KI-Systeme können nachhaltig wirken, wenn sie in ihrer Entscheidungsfindung Nachhaltigkeit berücksichtigen – sie beispielsweise nachhaltige Produkte im Handel fördern oder sie den allgemeinen Ressourcenverbrauch minimieren.



Indirekter Ressourcenverbrauch

Um die Hardware zu produzieren, die für KI-Systeme notwendig ist, werden Ressourcen und Energie benötigt. Hier gilt es, auf eine ressourcenschonende und umweltfreundliche Produktion und Entsorgung zu achten.





Ökonomische Nachhaltigkeit

Ökonomische Nachhaltigkeit erweitert den Horizont wirtschaftlicher Aktivitäten: Statt Ökonomie nur auf die Befriedigung der Bedürfnisse der heute lebenden Menschen auszurichten, wird aus einer ökonomisch nachhaltigen Perspektive auch die Bedürfniserfüllung der zukünftigen Menschheit eingeplant. Dieser Bewusstseinswandel drängt sich vor dem Hintergrund der „Grand Challenges“ auf: der Klimakrise, dem fortschreitenden Biodiversitäts- und Artenverlust sowie der fortschreitenden Ressourcenknappheit. Im Zuge einer sozial-ökologischen Transformation stellen sich grundsätzliche Gerechtigkeitsfragen, da Produktions- und Konsumzyklen bestimmen, ob die Verteilung natürlicher Ressourcen mit einem menschenwürdigen und selbstbestimmten Leben in Einklang gebracht werden kann. Ökonomische Nachhaltigkeit bettet die Ökonomie zwischen sozialen und ökologischen Leitplanken ein. Vor diesem Hintergrund sind auch die Auswirkungen von KI-Systemen einzuordnen. Gerade Systeme, die weitreichende Auswirkungen auf die gesellschaftliche Wohlstandsverteilung (etwa bei der Vergabe von Sozialleistungen, Krediten oder Wohnraum) sowie auf ökonomische Strukturen und Dynamiken haben, müssen besonders verantwortungsvoll eingesetzt werden.



Marktvietfald und Ausschöpfung des Innovationspotenzials

Um eine Konzentration in KI-Märkten zu verhindern, müssen faire Zugangsmöglichkeiten zur KI-Entwicklung geschaffen werden, zum Beispiel durch offene Datenpools, einen offenen Quellcode oder auch Schnittstellen (APIs).



Verteilungswirkung in Zielmärkten

Der Zugang zu KI-Anwendungen ist nicht allen Wirtschaftsakteuren gegeben. Es kommt mitunter zu Wettbewerbsverzerrungen oder Marktkonzentration. Hier könnte Inklusivität geschaffen werden, indem Modelle auch mit kleinen Datensätzen arbeiten oder kleine und mittlere Unternehmen durch Förderangebote in die Lage versetzt werden, KI einzusetzen.



Arbeitsbedingungen und Arbeitsplätze

Entlang der gesamten Wertschöpfungskette der KI-Entwicklung sollten faire Arbeitsbedingungen sichergestellt sein. Wenn KI am Arbeitsplatz eingesetzt wird, sollten vorab die Auswirkungen auf die Mitarbeitenden abgeschätzt und ggf. kompensiert werden.

Diese Kriterien haben wir in über 40 Indikatoren ausdifferenziert und operationalisiert. Nachzulesen sind sie im Diskussionspapier:

Rohde, F., Wagner, J., Reinhard, P., Petschow, U., Meyer, A., Voß, M., & Mollen, A. Nachhaltigkeitskriterien für künstliche Intelligenz. Schriftenreihe des IÖW, 220, 21.



Download: https://www.ioew.de/fileadmin/user_upload/BILDER_und_Downloaddateien/Publikationen/2021/IOEW_SR_220_Nachhaltigkeitskriterien_fuer_Kuenstliche_Intelligenz.pdf



ANDREAS MEYER

... ist wissenschaftlicher Mitarbeiter am Distributed Artificial Intelligence Labor der TU Berlin und forscht dort unter anderem an Anwendungen von Machine-Learning-Verfahren zur Lastprognose und Nachhaltigkeit von KI-Systemen.



ANNE MOLLEN

... ist Senior Policy und Advocacy Managerin bei AlgorithmWatch. Dort arbeitet sie unter anderem zur Nachhaltigkeit von Systemen des automatisierten Entscheidens (ADM-Systeme). Weitere Schwerpunkte sind ADM-Systeme im Arbeitsleben und im öffentlichen Sektor.



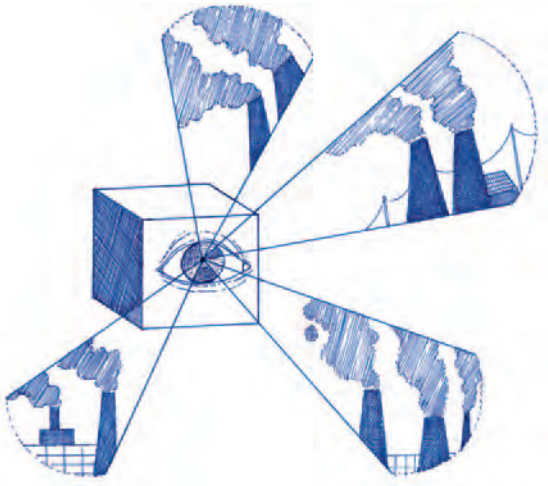
FRIEDERIKE ROHDE

... ist Nachhaltigkeitsforscherin und Techniksoziologin am Institut für ökologische Wirtschaftsforschung (IÖW) und promoviert an der TU Berlin. Sie beschäftigt sich mit sozio-technischen Zukünften im Kontext des digitalen Wandels, sozialen Innovationen und algorithmischen Entscheidungssystemen.



JOSEPHIN WAGNER

... ist wissenschaftliche Mitarbeiterin am Institut für ökologische Wirtschaftsforschung. Im Forschungsfeld Umweltökonomie und Umweltpolitik arbeitet sie schwerpunktmäßig zu den Themen „Digitalisierung und sozialer Wandel“ sowie „ökonomische und institutionelle Analyse von Umweltpolitiken“.



Nachhaltige KI 101

Immer wieder tauchen bestimmte Begriffe in der Diskussion um die Nachhaltigkeit von KI auf: *Foundation Models* sind besonders energieintensiv, *Model Cards* können mehr Transparenz schaffen, *Schnittstellen* versprechen Offenheit... Wir erklären zentrale Begriffe.

CO₂-Tools zum Messen des Verbrauchs von KI-Anwendungen

KI-Systeme verursachen über ihren Lebenszyklus hinweg direkte und indirekte CO₂-Emissionen. Treibhausgase entstehen zum einen durch die benötigte Hardware: während der Rohstoffgewinnung, der Herstellung, dem Transport und bei der Entsorgung. Die Emissionen lassen sich aufgrund fehlender Informationen und komplexer Lieferketten zumeist jedoch nicht genau beziffern. Außerdem ist es schwierig, den direkten Bezug zu einzelnen KI-Modellen herzustellen. Der Großteil der durch KI-Systeme verursachten CO₂-Emissionen entsteht durch den Stromverbrauch der Hardware, während die Systeme entwickelt und eingesetzt werden. Das Ausmaß der Emissionen hängt hauptsächlich von zwei Faktoren ab: dem Stromverbrauch der Hardware und der CO₂-Intensität des zugrundeliegenden Strommixes. Um den Emissionswert zu ermitteln, lassen sich manche Tools in den Quellcode der Anwendungen integrieren (CodeCarbon, experiment-impact-tracker, Carbontracker). Es ist auch möglich, unter Angabe der Systemspezifikationen Hochrechnungen der zu erwartenden Emissionen zu erstellen (zum Beispiel mit dem Machine Learning Emissions Calculator oder mit Carbontracker).

Datenpools

Datenpools sind Datensätze, die selbst durch KI generiert werden oder aus internen oder externen Quellen stammen. Man unterscheidet zwischen offenen und geschlossenen Datenpools. Bei offenen Datenpools teilen sich mehrere Unternehmen den Datenzugriff, nachdem sie sich hinsichtlich der Nutzung und Anpassung des Pools auf bestimmte Rahmenbedingungen geeinigt haben. Bei geschlossenen Datenpools handelt es sich um Datensätze, auf die nur ein Unternehmen Zugriff hat. Solche Datenpools haben für Unternehmen den Vorteil, dass sie mit bestimmten Daten verbundene Wettbewerbsvorteile nicht mit Konkurrenten teilen müssen. Aus einer ökonomischen Nachhaltigkeitsperspektive sind geschlossene Datenpools heikel. Sie führen zu Lock-in-Effekten, das heißt zu einer Bindung an das Produkt oder dessen Anbieter, die es den Kund*innen zum Beispiel wegen hoher Wechselkosten sehr schwer macht, zu anderen Produkten oder Anbietern zu wechseln. Eine zunehmende Marktkonzentration bis hin zur Monopolisierung ist die Folge. Darunter leidet die Innovationskraft des Marktes. Die Marktvietalt wird stark eingeschränkt, da einzelne große Unternehmen unverhältnismäßig von KI profitieren. Gleichzeitig bieten geschlossene Datenpools und die damit verbundenen Wettbewerbsvorteile Unternehmen Anreize, bei der Datenbeschaffung auf wettbewerbsverzerrende Praktiken zu setzen.

Offene Datenpools

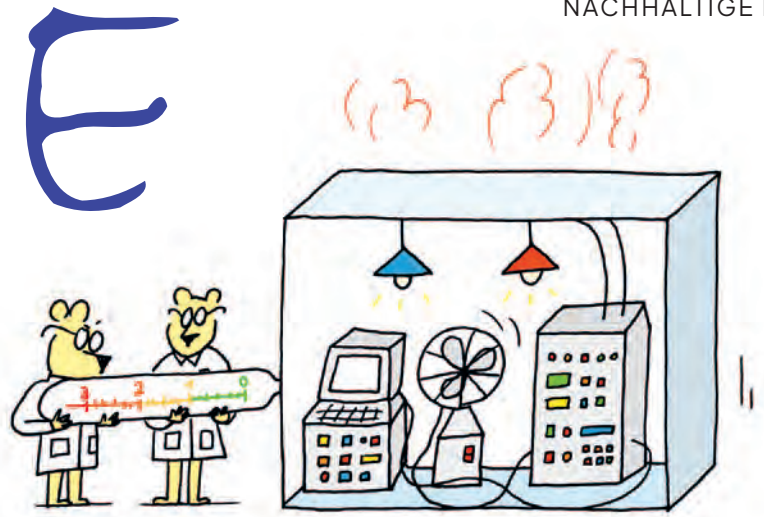


Geschlossene Datenpools



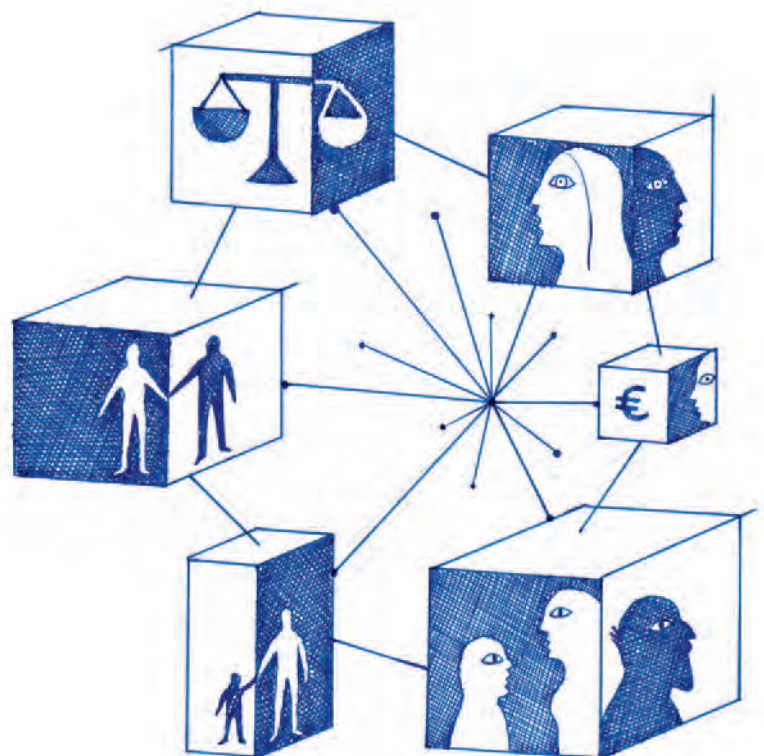
Effizienzmetriken für Rechenzentren

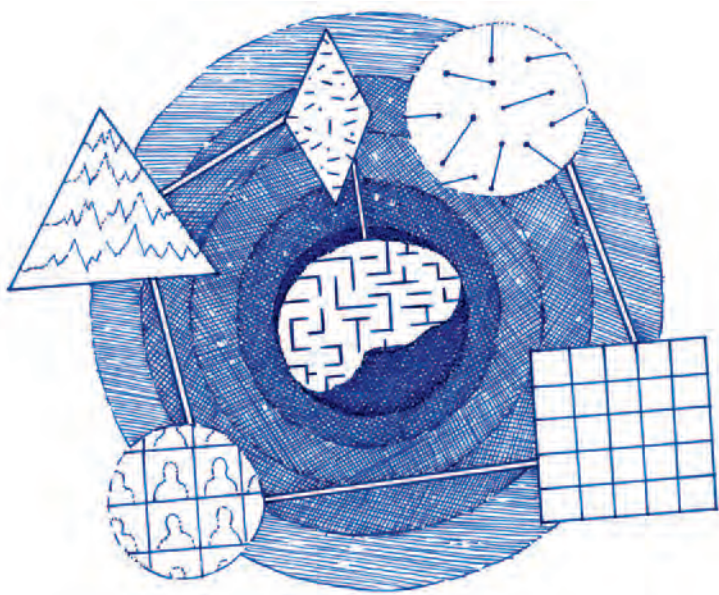
Effizienzmetriken sind dazu da, die Effizienz des Rechenzentrumsbetriebs messbar zu machen und vorhandene Schwachstellen in unterschiedlichen Bereichen zu identifizieren. Die jeweiligen Metriken bilden verschiedene Ressourcenverbräuche ab. Die Power Usage Effectiveness (PUE) ist eine der am häufigsten eingesetzten Effizienzmetriken. Durch sie wird die Effizienz des Energieeinsatzes bemessen und erfasst, welches Verhältnis zwischen dem Gesamtenergieverbrauch und dem der IT-Infrastruktur besteht. Bei einem PUE von 1 würde die gesamte aufgewendete Energie in die Infrastruktur fließen. Ein Wert von 2 würde bedeuten, dass Kühlung, Beleuchtung und Räumlichkeiten genau so viel Strom benötigen wie die IT-Infrastruktur. Alles, was nicht direkt dem Betrieb des Computings dient, gilt als Nicht-Infrastruktur. Im Jahr 2020 lag der durchschnittliche PUE von deutschen Rechenzentren bei 1,63. Analog lässt sich die Einsatzeffizienz anderer Ressourcen bemessen. Durch die *Water Usage Effectiveness* (WUE) und die *Carbon Usage Effectiveness* (CUE) werden beispielsweise bei einem gegebenem Energieverbrauch der Wasserverbrauch und der Kohlenstoffdioxidausstoß erfasst. Es existieren viele solcher Effizienzmetriken, die zum Beispiel auch die Auslastung der technischen Infrastruktur oder die Weiterverwendung der anfallenden Abwärme widerspiegeln.



Fairness-Metriken

Algorithmen können Menschen unter anderem aufgrund von Alter, Geschlecht oder Hautfarbe diskriminieren, wenn zum Beispiel die Daten, mit denen die Modelle trainiert wurden, einen Bias (also eine Verzerrung) enthalten und gesellschaftliche Vorurteile reproduzieren. Um solche diskriminierenden Effekte zu vermeiden, wird die Fairness von Modellen gemessen und getestet. Die verschiedenen statistischen Ansätze dazu werden als Fairness-Metriken bezeichnet. Mit diesen Messgrößen kann beispielsweise die Wahrscheinlichkeit von vorteilhaften Entscheidungen des Algorithmus für Gruppen mit verschiedenen demographischen Merkmalen wie Alter oder Einkommen gemessen werden, oder auch geprüft werden, ob die Genauigkeit des Modells für verschiedene Subgruppen gleich ist (zum Beispiel ob die Ergebnisse von Kreditwürdigkeitsprüfungen erhebliche Abweichungen zwischen männlichen und weiblichen Personen zeigen). In der Online-Bibliothek *Tensorflow* (<https://github.com/tensorflow/fairness-indicators>) werden verschiedene Messverfahren für Fairness aufgeführt. Der Messbarkeit von Fairness sind allerdings Grenzen gesetzt, da viele Einflussgrößen nur schwer quantifizierbar sind. Aus diesem Grund ist es wichtig, Fairness klar zu definieren und schützenswerte Gruppen anhand von geschützten Attributen wie Ethnizität, Herkunft oder Hautfarbe zu benennen.





Foundation Models

Der Begriff Foundation Models (zu Deutsch: Basismodelle) wurde durch eine Forschungsgruppe der Universität Stanford geprägt, die damit auf einen voranschreitenden Paradigmenwechsel im Bereich der KI reagierte. Gerade KI-Modelle zur Sprachverarbeitung werden seit einiger Zeit immer größer und leistungsstärker. Beispiele aus den letzten Jahren sind Open-AIs GPT-3 und Googles BERT. In einem häufig sehr aufwendigen Prozess lernen KI-Modelle eine vordefinierte Zuordnung durch Verarbeitung vorher annotierter Daten. Dieses sogenannte überwachtes Lernen wurde zum Beispiel zum Erkennen von Objekten in Bildern eingesetzt. Basismodelle wie GPT-3 und BERT versuchen beim selbstüberwachten Lernen zunächst, allgemeine Muster in den Daten zu erkennen und zu erlernen. Da Daten hierfür nicht manuell annotiert werden müssen, können die Modelle auf deutlich größere Datensätze zurückgreifen und komplexe Zusammenhänge erkennen. Diese Modelle lassen sich deutlich flexibler und vielfältiger als Basis für verschiedene Anwendungen einsetzen. Diese Anwendungen übernehmen allerdings dabei deren Eigenschaften und mögliche Verzerrungen. Mit der steigenden Komplexität von KI-Systemen steigen auch die Kosten für deren Entwicklung. Infolgedessen sinkt die Anzahl der Akteure, die in der Lage sind, so große KI-Modelle zu entwickeln, was zu einer zunehmenden Zentralisierung der KI-Entwicklung führt.



Modelcards

Mit sogenannten Modelcards werden Machine-Learning-Modelle dokumentiert. Die Modelcards sollen darüber Aufschluss geben, in welchem Kontext algorithmische Entscheidungssysteme genutzt werden. Es handelt sich um kurze Dokumente, in denen die Leistungsmerkmale des jeweiligen Algorithmus strukturiert aufgeführt sind. Sie enthalten auch Informationen zum Kontext, in dem das Training des Modells stattfindet: Informationen über verschiedene kulturelle, demografische oder phänotypische Gruppen (z. B. Ethnizität, geografischer Standort, Geschlecht, Fitzpatrick-Hauttyp) und intersektionale Gruppen (z. B. Alter und Ethnizität oder Geschlecht und Fitzpatrick-Hauttyp). Eine vollständige Dokumentation sollte außerdem den Typ und die Details des Machine-Learning-Modells enthalten, ebenso wie Informationen zum geplanten Anwendungsfall und zu möglichen Einflussfaktoren. Darüber hinaus sollten in den Modelcards Test- und Trainingsdaten sowie ethische Aspekte oder Bedenken festgehalten werden. Diese Dokumentation soll dafür sorgen, dass das Maschinelle Lernen der jeweiligen Modelle transparenter und nachvollziehbarer wird. Zu diesem Zweck hat etwa Google eine Modelcard für einen Algorithmus veröffentlicht, der auf Fotos und in Videos Gesichter erkennt. Ein anderes Beispiel ist die Modelcard für das häufig genutzte Sprachverarbeitungsmodell BERT, die auf der Entwicklungsplattform *Hugging Face* veröffentlicht wurde (<https://huggingface.co/distilbert-base-uncased>).



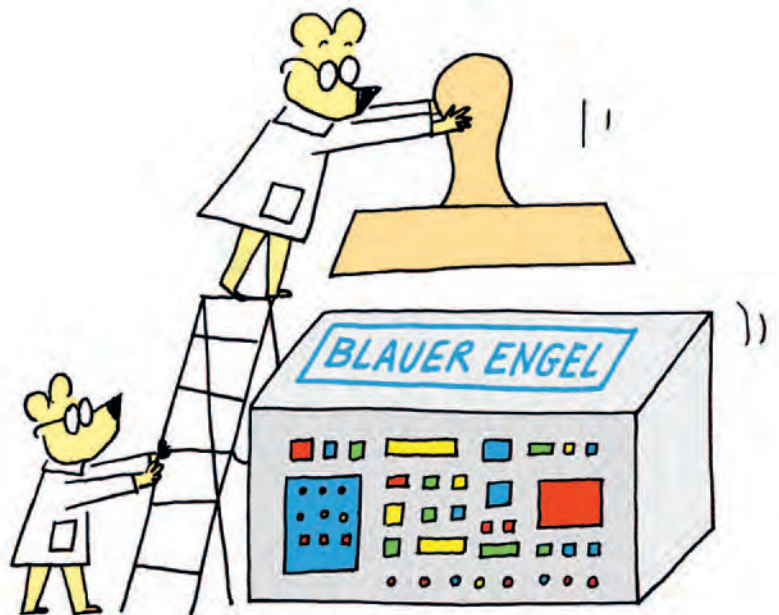


Zertifizierung von Hardware und Rechenzentren

Für das Betreiben von Rechenzentren wird immer mehr Energie benötigt. In Frankfurt sind Rechenzentren mittlerweile für 20 Prozent des gesamten Stromverbrauchs verantwortlich, Tendenz steigend. Um bei der Auswahl von Rechenzentren und von Hardware nachhaltigere Alternativen zu finden, können Zertifizierungen helfen. Die Zertifikate werden für einzelne Hardwarekomponenten vergeben (zum Beispiel *Der Blaue Engel* oder *Energy Star* für Server und Datenspeicherprodukte), aber auch für ganze Rechenzentren und Co-Location-Rechenzentren, in denen Unternehmen Flächen anmieten, um dort ihre eigene IT-Hardware zu betreiben. Neben dem Blauen Engel existiert für Rechenzentren beispielsweise auch das CEEDA-Zertifikat. Der Blaue Engel zertifiziert eine effiziente Stromversorgung und Klimatisierung sowie die Verwendung von energieeffizienten, langlebigen und umweltfreundlichen Hardwarekomponenten. Durch ein konsequentes Monitoring des Rechenzentrumsbetriebs und einen jährlichen Bericht wird fortlaufend die Einhaltung der Vorgaben sichergestellt.

Schnittstellen

KI-Systeme oder -Plattformen können Schnittstellen (APIs = Application Programming Interfaces) für Drittanbieter und andere KI-entwickelnde Unternehmen bereitstellen, um KI-Systeme durch externe Services zu erweitern. Schnittstellen erleichtern den Datentransfer und die Kommunikation zwischen zwei Programmen. Entwickler*innen können sie nutzen, um Daten zu suchen, zu sammeln und zu teilen, oder um Funktionen in eigene Softwareprodukte einzubauen und zu adaptieren. Grundsätzlich wird zwischen *Private APIs*, *Partner APIs* und *Public APIs* unterschieden. Auf *Private APIs* können nur interne Entwickler*innen und Mitarbeiter*innen innerhalb eines Unternehmens zugreifen. Eingeschränkte APIs, sogenannte *Partner APIs*, können nur von bestimmten Unternehmen unter bestimmten vertraglichen Bedingungen verwendet werden. *Public APIs*, also öffentlich zugängliche Schnittstellen, sind frei verfügbar und können von jedem Unternehmen ohne Einschränkungen verwendet werden. Im Gegensatz zu Open-Source-Software erlauben *Public APIs* keine Einblicke in den Quellcode und keine freie Anpassung. Von offenen APIs profitieren insbesondere kleine und mittelständische Unternehmen sowie NGOs. Sie besitzen oftmals nicht die Ressourcen und das Know-how, um eine eigene wettbewerbsfähige KI-Software zu entwickeln. Mit offenen APIs können sie Prozesse effizienter und ressourcenschonender gestalten und flexibel KI-Funktionen in ihre eigene Software integrieren.



Nachhaltige KI: Zahlen & Fakten

Für den jährlich erscheinenden AI Index Report der Stanford University werden Daten zu KI erhoben, extrahiert und visualisiert. Der Bericht soll zu einem umfassenderen und nuancierteren Verständnis rund um das Thema KI führen. Hier sind einige Daten aus der diesjährigen Ausgabe, die aus der Nachhaltigkeitsperspektive relevant sind.

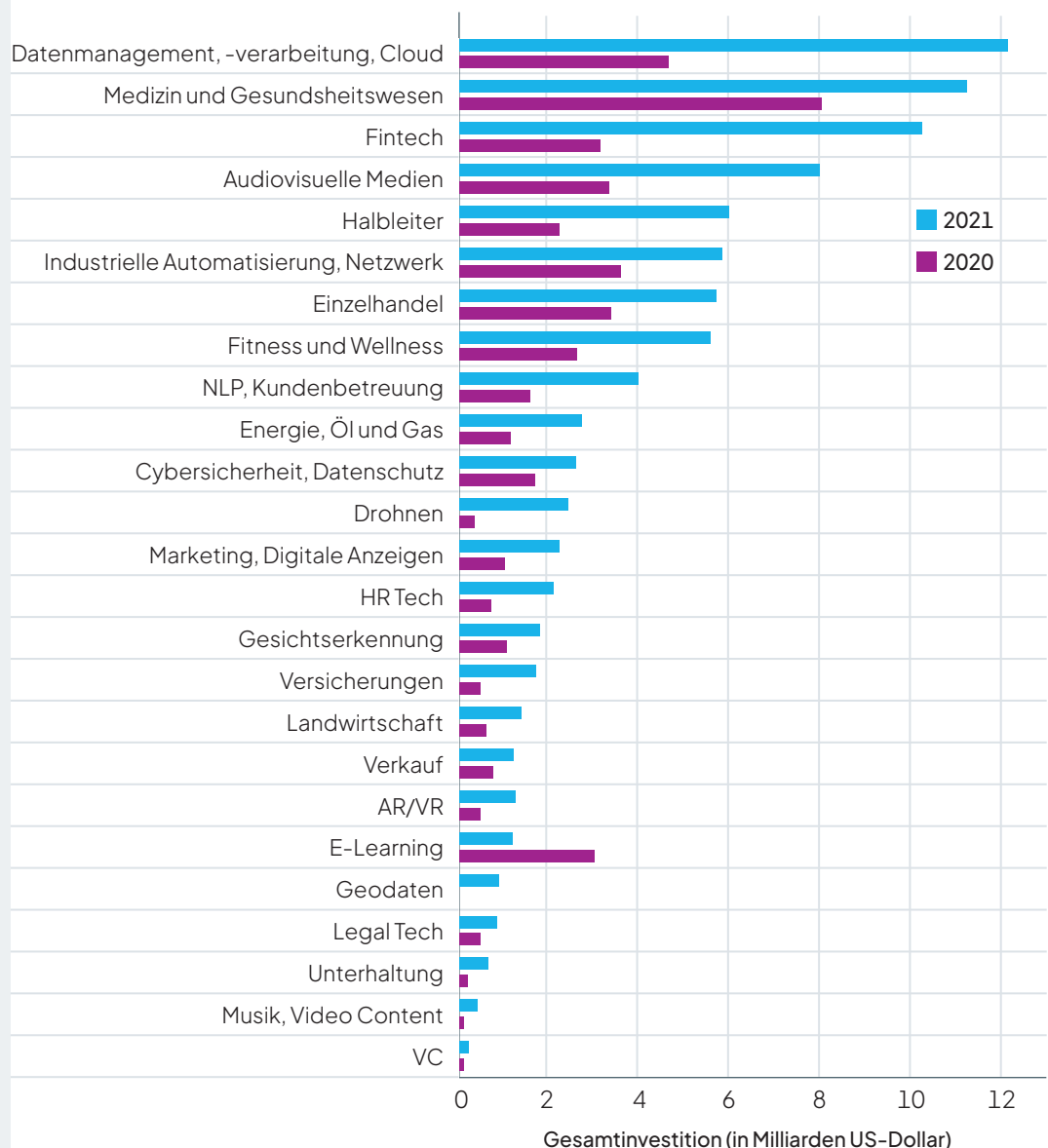
Wachstum:

Private Investitionen in KI

Die KI-Branche wächst rasant, vor allem die Investitionen in Datenmanagement, -verarbeitung und Clouds. Sie nahmen 2021 im Vergleich zum Vorjahr um mehr als das Zweieinhalbfache zu und lagen bei etwa 4,69 Milliarden US-Dollar. Zwei der vier größten privaten Investitionen flossen 2021 in Datenmanagement-Unternehmen.

PRIVATE INVESTITIONEN IN KI NACH SCHWERPUNKTBEREICHEN, 2020 VS. 2021

(Quelle: NetBase Quid, 2021 | Grafik: 2022 AI Index Report)



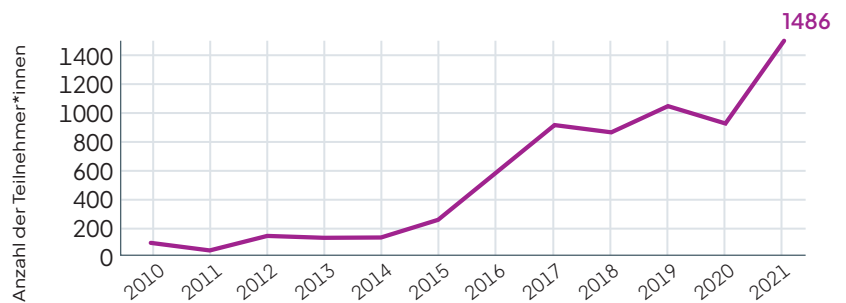
Diversität:**Women In Machine Learning**

Diversität ist ein Schlüsselement im Kampf gegen Diskriminierung. Die 2006 gegründete Organisation Women in Machine Learning (WiML) arbeitet daran, Frauen im Bereich des Maschinellen Lernens größeren Einfluss zu ermöglichen.

Die Daten zeigen, wie viele Teilnehmer*innen die WiML-Workshops im Laufe der Jahre bei der NeurIPS-Konferenz hatten, die eine der wichtigsten auf dem Gebiet von KI und ML ist.

TEILNAHME AN NEURIPS WOMEN IN MACHINE LEARNING WORKSHOPS, 2010–2021

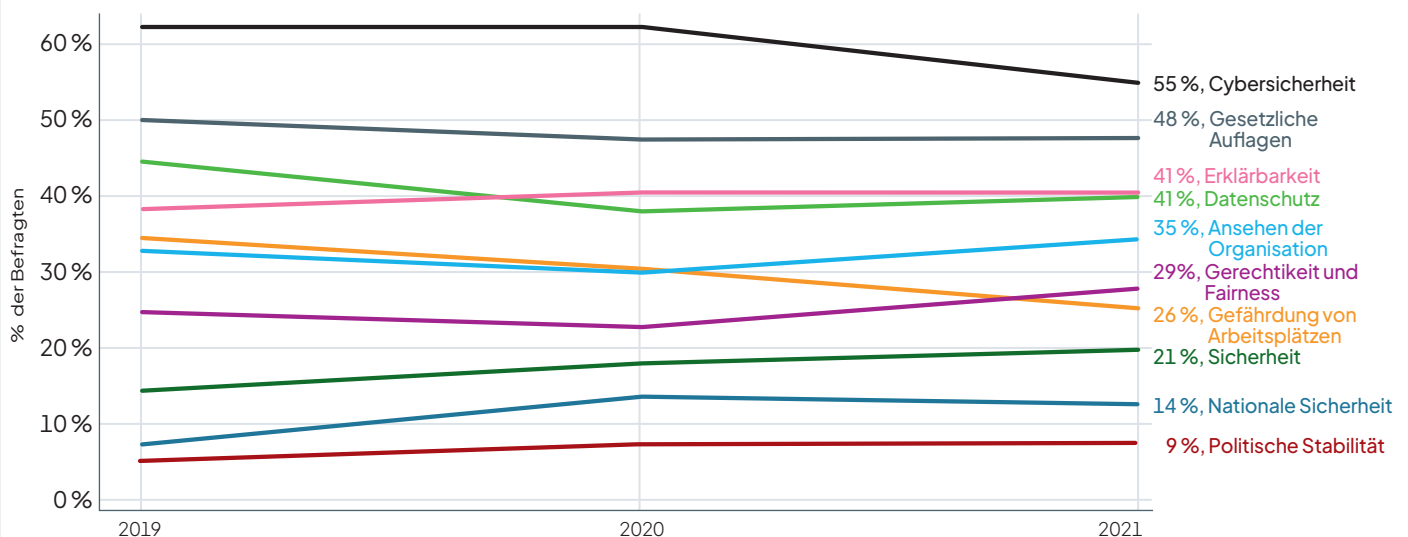
(Quelle: Women in Machine Learning, 2021 | Grafik: 2022 AI Index Report)

**Risiken:****Diskrepanz zwischen Problembewertung und Problembehebung**

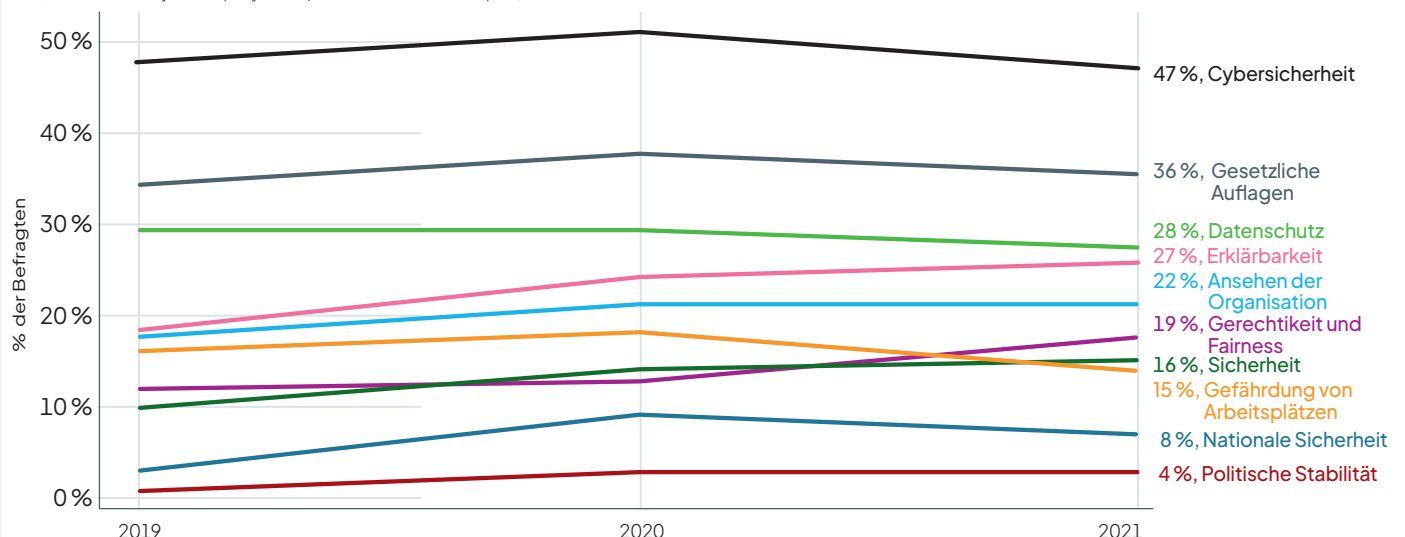
Es reicht nicht, Risiken von KI zu erkennen, sie müssen auch aktiv angegangen werden. Bislang klafft hier eine Lücke. Zum Beispiel sehen 29 Prozent der befragten Organisationen hinsichtlich der Gleichbehandlung und Fairness ein Risiko darin, KI-System einzusetzen, aber nur 19 Prozent ergreifen bei der Implementierung Maßnahmen, um die Risiken einzudämmen. Auch bei anderen Risiken ist diese Diskrepanz vorhanden.

RISIKEN BEI DER EINFÜHRUNG VON KI, DIE VON ORGANISATIONEN ALS RELEVANT WAHRGENOMMEN WERDEN, 2019–2021

(Quelle: McKinsey & Company, 2021 | Grafik: 2022 AI Index Report)

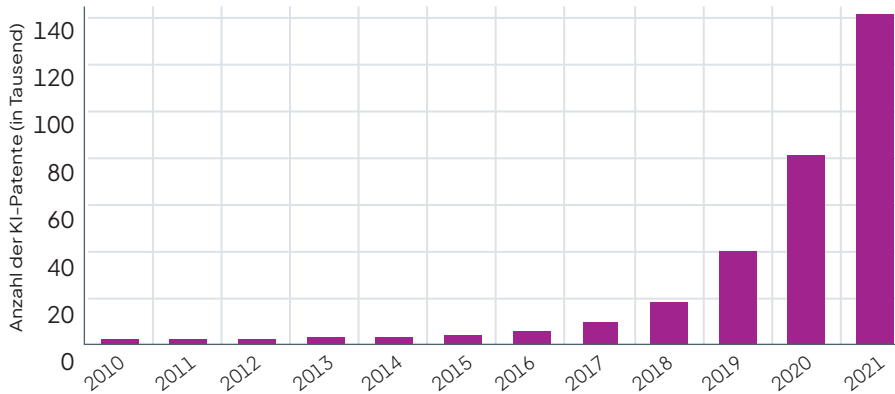
**RISIKEN BEI DER EINFÜHRUNG VON KI, GEGEN DIE ORGANISATIONEN MASSNAHMEN ERGREIFEN, 2019–2021**

(Quelle: McKinsey & Company, 2021 | Grafik: 2022 AI Index Report)



ANZAHL DER KI-PATENTANMELDUNGEN, 2010–2021

(Quelle: Center for Security and Emerging Technology, 2021 | Grafik: 2022 AI Index Report)



Wachstum:

KI-Patente

KI ist auf dem Vormarsch: Die Anzahl der 2021 eingereichten Patente ist im Vergleich zu 2015 um mehr als das 30-fache gestiegen, was einer durchschnittlichen jährlichen Wachstumsrate von 76,9 Prozent entspricht.

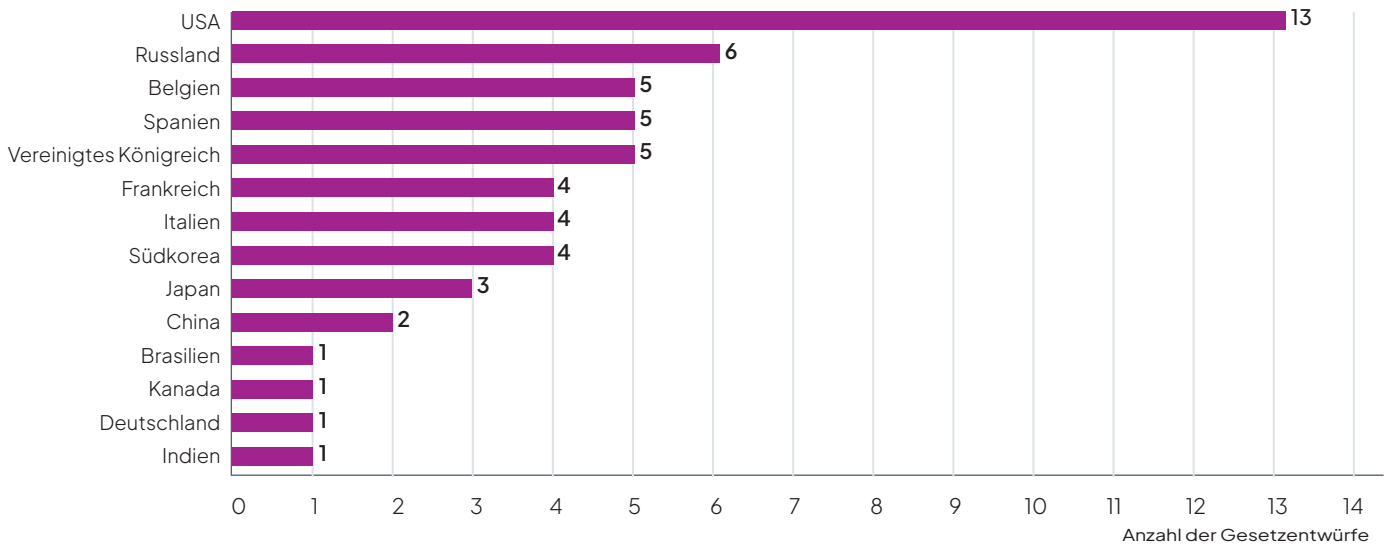
Governance:

Globale KI-Gesetzgebung

Regierungen und Rechtsorgane weltweit arbeiten an Gesetzen, um die Entwicklung von KI zu regulieren. In 25 Ländern wurden 55 Gesetze verabschiedet, deren Wortlaut den Begriff Künstliche Intelligenz enthielt.

ANZAHL DER VERABSCHIEDETEN KI-GESETZESENTWÜRFE IN AUSGEWÄHLTEN LÄNDERN, 2016–2021 (GESAMT)

(Quelle: AI Index 2021 | Grafik: 2022 AI Index Report)



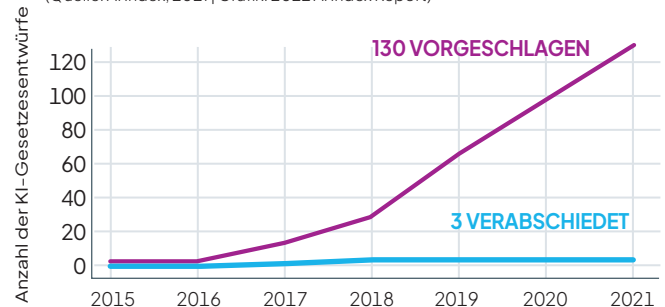
Governance:

KI-Gesetzgebung auf Bundesebene in den USA

In den US-amerikanischen Bundesstaaten ist ein starker Anstieg von KI-Gesetzesentwürfen zu verzeichnen. Für 2015 ist nur ein Bundesgesetzentwurf dokumentiert, während es 130 im Jahr 2021 waren. Die Zahl der verabschiedeten KI-Gesetze hält nicht mit der wachsenden Menge an KI-Gesetzesentwürfen Schritt. 2021 zeigte sich dies besonders deutlich, da nur zwei Prozent der Entwürfe auch als Gesetz verabschiedet wurden.

ANZAHL DER KI-GESETZESENTWÜRFE IN DEN USA, 2015–21 (VORGESCHLAGEN VS. VERABSCHIEDET)

(Quelle: AI Index, 2021 | Grafik: 2022 AI Index Report)

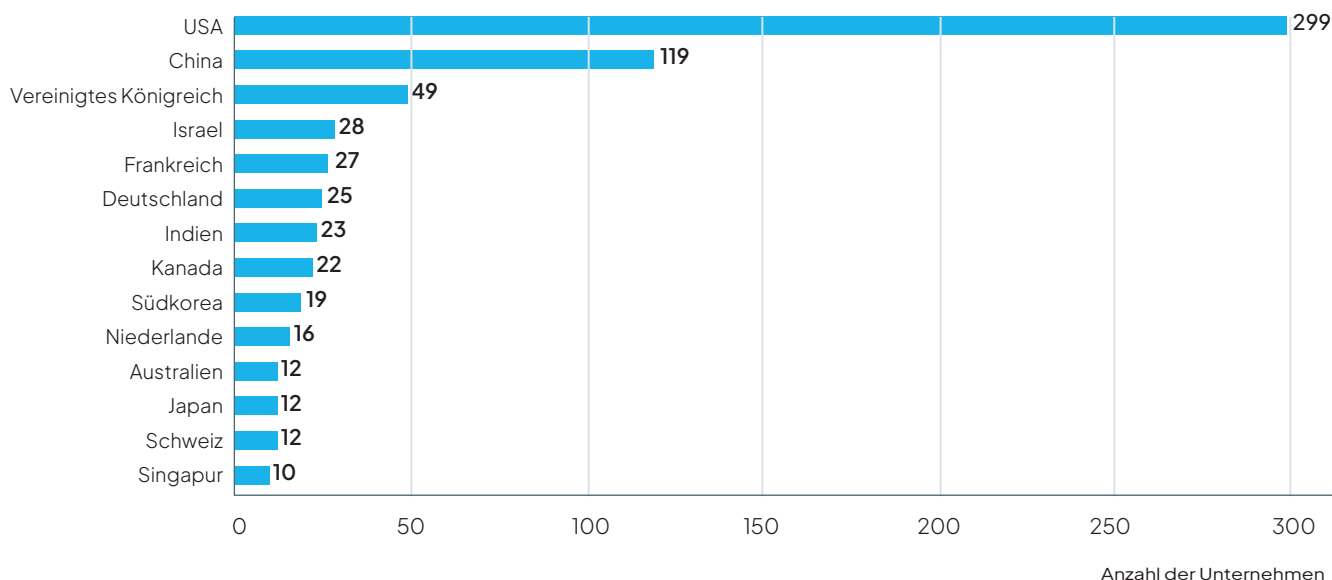


Wachstum:**Anzahl der neugegründeten KI-Unternehmen**

Die Kluft zwischen den zwei führenden Ländern in der KI-Entwicklung und dem Rest der Welt ist enorm: Die USA dominieren den Markt, gefolgt von China. Andere asiatische und europäische Länder liegen weit dahinter.

ANZAHL DER NEUGEGRÜNDETEN KI-UNTERNEHMEN NACH LÄNDERN, 2021

(Quelle: NetBase Quid, 2021 | Grafik: 2022 AI Index Report)

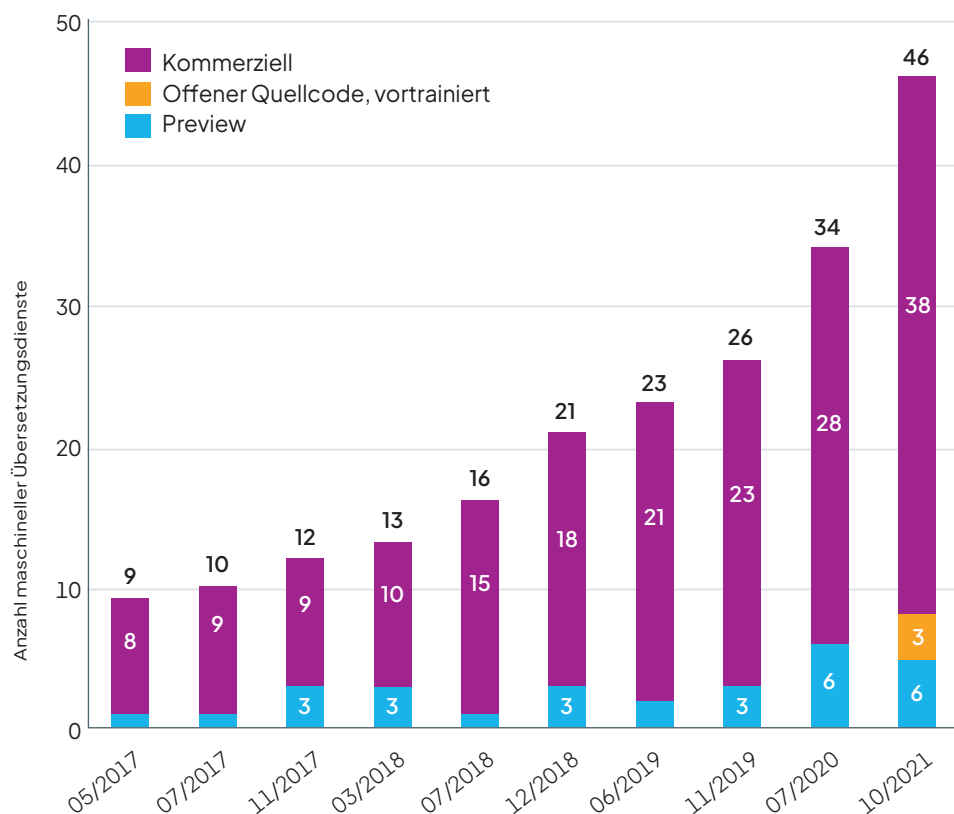
**Vortrainiert:**

**Kommerziell
verfügbare
maschinelle
Übersetzungs-
systeme**

Das wachsende Interesse an maschinellen Übersetzungen spiegelt sich in der steigenden Zahl an kommerziell verfügbaren maschinellen Übersetzungsdiensten wie Google Translate wider. Seit 2017 hat sich deren Zahl beinahe verfünffacht – während im gleichen Zeitraum nur wenige vortrainierte Modelle eingeführt wurden: die Open-Source-Dienste M2M-100, mBART und OPUS im Jahr 2021.

ANZAHL MASCHINELLER ÜBERSETZUNGSDIENSTE

(Quelle: Intento, 2021 | Grafik: 2022 AI Index Report)





KI als Gemeinschafts- aufgabe denken

Die Soziologin Alex Hanna untersucht, wie der Einsatz von Daten in neuen Technologien dazu beiträgt, dass sich vorhandene Ungleichheiten im Hinblick auf Gender, Ethnizität und Klasse verschärfen. Wir haben mit ihr darüber gesprochen, warum sie ihren Job im Ethik-Team von Google aufgegeben hat, um sich ihrer früheren Vorgesetzten Timnit Gebru anzuschließen – die zuvor bei Google entlassen worden war.

Sie sind seit Kurzem Forschungsleiterin am DAIR-Institut. Das erklärte Ziel des Instituts ist, dem alles bestimmenden Einfluss der großen Tech-Konzerne bei der Forschung, Entwicklung und Implementierung von KI etwas entgegenzusetzen. Inwiefern ist dieser Einfluss problematisch?

Die Finanz- und Fördermittel konzentrieren sich dort. Google, Microsoft oder auch Facebook finanzieren ausschließlich Forschung, die bestehenden wissenschaftlichen Paradigmen folgt und auf die Optimierung ihrer Geschäftsmodelle ausgerichtet ist. Das ist entweder direkt oder indirekt der Fall, also entweder durch die Art der wissenschaftlichen Artikel, die sie veröffentlichen, oder durch die Förderprojekte, die sie an Universitäten, gemeinnützige Organisationen oder auch unter dem Label *AI for Good* vergeben. Die Förderung gibt vor, an welchen Problemen die Forscher*innen arbeiten. Die Konzerne fördern keine Projekte, die nicht im Einklang mit ihren Interessen stehen, oder nur in sehr begrenztem Ausmaß.

Im Moment ist also KI ein Teil des Problems?

Sie ist ein Teil des Problems, wenn sie dazu benutzt wird, Macht zu gewinnen und Macht zu konsolidieren. Ihr gegenwärtiger Einsatz verschärft die existierende Ungleichheit. Die großen Tech-Konzerne nutzen sie vor allem für Empfehlungssysteme, Zielgruppenwerbung und um zu vermeiden, dass Kund*innen

abwandern, kurz: zur Optimierung ihrer Geschäftsmodelle. Im öffentlichen Sektor wird KI eingesetzt, um den Arbeitsaufwand für die Bewilligung von Sozialleistungen zu minimieren oder Betrug aufzudecken. Gleichzeitig entwickelt sie sich aber zu einem Überwachungsinstrument. Im wirtschaftlichen Kontext verschlimmert sie die Bedingungen, unter denen Menschen arbeiten müssen, entweder indem sie ein neues Prekariat von Arbeiter*innen schafft, die Daten für KI generieren, oder indem sie als Taktgeberin bei der Plattformarbeit fungiert.

Was werden Sie bei DAIR anders machen?

Menschen, deren Leben durch KI und automatisierte Entscheidungssysteme betroffen sind, müssen einen größeren Einfluss darauf ausüben können, wie diese Systeme eingesetzt werden. Wir fangen damit an, indem wir die betroffenen Gemeinschaften in unsere Forschungsaktivitäten einbeziehen.

Wie setzen Sie das um? Wie können wir eine Diskussion über einen bewussteren Einsatz von KI anstoßen?

Wenn wir KI neu denken, müssen wir darüber nachdenken, was die Betroffenen wirklich brauchen, vor allem wenn es um Gruppen geht, die durch rassistische oder sexistische Vorur-

Gefährliches Nachplappern: Was richten Riesensprachmodelle an?

Im 2021 veröffentlichten Artikel *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* („Die Gefahr stochastischer Papageien: Können Sprachmodelle zu groß sein?“) erörtern Timnit Gebru, Emily Bender und weitere Autor*innen die Risiken, die von sehr großen Sprachmodellen ausgehen. Diese gerade sehr beliebten KI-Modelle werden mit enormen Mengen an Textdaten trainiert. Sie sind unter den richtigen Voraussetzungen erstaunlich gut darin, Sinngehalt aus Sprache abzuleiten und ihrerseits Texte zu produzieren, die Sinn zu ergeben scheinen.

Schädlicher als Fliegen

Die in *Stochastic Parrots* ausgeführte Analyse baute auf vorangegangener Forschungsarbeit auf, vor allem ein 2019 erschienenes Paper von Emma Strubell, Ananya Ganesh und Andrew McCallum über die Kohlendioxid-Emissionen und finanziellen Kosten von sehr großen Sprachmodellen: *Energy and Policy Considerations for Deep Learning in Natural Language Processing* („Deep Learning und natürliche Sprachverarbeitung: gesellschaftliche und energiepolitische Erwägungen“). Um große KI-Modelle trainieren zu können,

ist eine gewaltige Rechenleistung nötig, womit wiederum ein immenser Strombedarf verbunden ist. Ihr Energieverbrauch und ihre CO₂-Bilanz sind seit 2017 ausgeufert, da das in die Modelle eingespeiste Datenvolumen immer größer geworden ist. Ein Trainingsdurchlauf für eine Version von Googles Sprachmodell BERT, das sehr wichtig für die Funktion der Suchmaschine ist, setzt gut 700 Kilo CO₂ frei, also ungefähr so viel wie ein Rundflug zwischen New York City und San Francisco. Solche Modelle werden allerdings etliche Male während des Forschungs- und Entwicklungsprozesses trainiert.

teile marginalisiert werden. Manche KI-Instrumente könnten genutzt werden, um diesen Gemeinschaften etwas von ihrer Entscheidungsgewalt zurückzugeben oder sie zumindest bei ihren Entscheidungen zu unterstützen. In diese Richtung gehen einige unserer Projekte. Wir untersuchen zum Beispiel in Südafrika, wie sich räumliche Apartheid mithilfe von KI rückgängig machen lässt. In einem anderen Forschungsprojekt beschäftigen wir uns damit, wie KI und Tools zur Verarbeitung von natürlicher Sprache verwendet werden können, um hetzerische Social Media-Konten von Regierungsvertreter*innen ausfindig zu machen. Wir versuchen also, neue Ansätze für den Einsatz von KI zu finden, die der Machtkonzentration entgegenwirken und den Menschen ein wenig Selbstbestimmung zurückgeben.

Im Dezember 2021,

... am Jahrestag ihrer Entlassung bei Google, veröffentlichte Timnit Gebru eine Pressemitteilung, in der sie die Gründung des *Distributed AI Research Institute (DAIR)* verkündete. Das DAIR soll als unabhängiges Institut bei der Erforschung, Entwicklung und Implementierung von KI den Einflussbereich der großen Tech-Konzerne eindämmen und stärker die Bedürfnisse von betroffenen Gemeinschaften berücksichtigen. Timnit Gebru will mit ihrem Institut die Prozesse und die Prinzipien der aktuellen KI-Forschung hinterfragen. Ihre Arbeit stützt sich auf den Gedanken, dass die mit KI-Technologie einhergehenden Schäden verhindert werden können, wenn der Einsatz von KI die Interessen verschiedener Communitys und eine Vielfalt an Perspektiven stärker einbeziehen würde.

Bei einem aktuellen Projekt verwendet das Institut Satellitenbilder und Bilderkennungssoftware, um die Folgen der Apartheid in Südafrika zu analysieren. Bei *Datasheets for Datasets*, einem weiteren Projekt des DAIR, geht es darum, beim Maschinellen Lernen weitgehend inexistenten Branchenstandards bei der Dokumentation von Datensätzen zu etablieren. Im Bereich des Maschinellen Lernens sollen Forschende und Praktiker*innen dadurch ein größeres Bewusstsein für notwendige Transparenz- und Rechenschaftspflichten entwickeln, Bias in ML-Modellen vermindern und bei der Auswahl passender Datensätze unterstützt werden.

Fortsetzung ...

Die Vervielfältigung sozialer Spaltungen

Timnit Gebru und ihre Co-Autor*innen geben in ihrem Paper zu bedenken, dass nur reiche Unternehmen Zugang zu den Ressourcen haben, die notwendig sind, um solche großen KI-Modelle zu bauen und zu pflegen. Die Folgen der Klimakrise, die auch durch einen exzessiven Energieverbrauch ausgelöst werden, treffen hingegen besonders marginalisierte Gemeinschaften. Die Trainingsdaten für die Modelle werden üblicherweise im Internet gesammelt, so dass das Risiko besteht, dass rassistischer, sexistischer oder ähnlich abwertender Sprachgebrauch darin Eingang findet. Da die Datensätze so groß sind, ist es nahezu unmöglich, diskriminierende Sprachmuster zu finden und herauszufiltern. Daraus folgern die Autor*innen, dass große Modelle

Sprache in einer Weise interpretieren, die überholte gesellschaftliche Normen und Diskriminierungsmuster reproduziert. Die Modelle würden zudem die Sprache und Normen von Ländern und Kulturen ausblenden, in denen das Internet weniger verbreitet ist, weshalb ihr linguistischer Abdruck dort deutlich kleiner ist.

Die Kosten des Profits

Ein weiteres Problem sehen die Autor*innen des *Stochastic Parrots* in einem „fehlgeleiteten Forschungsverständnis“, das sich hinter großen Sprachmodellen verstecken kann. Diese Modelle verstünden Sprache nämlich nicht wirklich. Sie errechnen die Wahrscheinlichkeit, dass bestimmte Wortsequenzen für uns Sinn ergeben und plappern die Wörter auf dieser Grundlage bloß nach. Sinn ist allerdings etwas, das innerhalb einer Sprachgemeinschaft verhandelt wird. Wenn Sinn als ma-

schinell ermittelbares, objektives Ergebnis definiert wird, besteht die Gefahr, dass Sprache durch diese Pseudo-Objektivität von KI manipuliert wird. Die Investitionen der großen Tech-Konzerne in KI sind insofern per se heikel, dass sie auf Profite ausgerichtet sind. Gesellschaftlich wäre es allerdings wünschenswerter, KI-Modelle zu entwickeln, die ein wechselseitiges Verständnis fördern und Energie sparen, indem sie mit kleineren und besser kuratierten Datensätzen zu guten Ergebnissen gelangen. Timnit Gebru und ihre Mitstreiter*innen sind allerdings pessimistisch, dass es auf Seiten der Konzerne zu einem Umdenken und einer Abkehr vom reinen Profitdenken kommt, obwohl der vorgegaukelte Sinngehalt der Sprachmodelle eine weitere Gefahr mit sich bringt: Sie könnten als Instrument missbraucht werden, um Desinformationen zu streuen.

Chronik einer angekündigten Trennung

Die Entwicklungen, die zur Gründung des DAIR-Instituts führten, erzählen nicht bloß die Geschichte eines Rauswurfs. Sie werfen die Frage auf, ob die großen Tech-Konzerne eine toxische Arbeitsumgebung schaffen.

Am 2. Dezember 2020 verkündete Timnit Gebru, die stellvertretende Leiterin des KI-Ethikteams bei Google, per Twitter, dass das Unternehmen sie rausgeworfen habe. Sie war 2018, als sie noch für Microsoft arbeitete, als Co-Autorin der bahnbrechenden Studie *On the gender and racial biases embedded in commercial face recognition systems* („Gender-basierter und rassistischer Bias in kommerziellen Gesichtserkennungssystemen“) bekannt geworden. In der Studie ging es darum, dass Gesichtserkennungstools Frauen und People of Color weniger verlässlich identifizieren und potenziell diskriminieren. Timnit Gebru ist eine der Mitgründer*innen der Gruppe Black in AI, die sich für mehr Diversität in der Tech-Industrie einsetzt. In ihrer kritischen Arbeit fordert sie häufig gängige Praktiken in der KI- und Tech-Industrie heraus.

Der Rauswurf von Timnit Gebru war letztlich die Konsequenz eines Konflikts, der sich an einem Artikel entzündet hatte, an dem sie als Mitautorin beteiligt war. Ihre Vorgesetzten bei Google forderten sie auf, entweder von der Veröffentlichung abzusehen oder den Namen aller Google-Mitarbeiter*innen auf dem Artikel zu entfernen (fünf der sechs Autor*innen). Jeff Dean, der Leiter von Google AI, schrieb in einer internen E-Mail (die er später auf Twitter teilte), dass der Artikel „nicht unseren Anforderungen an eine Publikation entspricht, da es die relevante Forschung nur unzureichend berücksichtigt“. Im Artikel sind allerdings

128 Forschungsreferenzen aufgeführt. Bei anderen KI-Ethiker*innen fachte dieser Umstand Spekulationen darüber an, ob Google Timnit Gebru nicht einfach loswerden wollte, weil ihr Text unangenehme Wahrheiten über ein zentrales Forschungsfeld des Unternehmens enthält. Über 1.400 Google-Mitarbeiter*innen und 1.900 weitere Unterstützer*innen unterzeichneten daher nach der Entlassung einen Protestbrief.

Der Artikel, um den es ging, trägt den Titel *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? („Über die Gefahren stochastischer Papageien: Können Sprachmodelle zu groß sein?“)*. Er sollte eine Bestandsaufnahme darüber sein, wie es um die Forschung zur Verarbeitung von natürlicher Sprache geht. Google hat einen Großteil der Grundlagenforschung auf dem Gebiet großer Sprachmodelle geleistet: 2017 wurde sowohl das *Transformer language model* vorgestellt, das die Basis für das spätere Google-Modell BERT werden sollte, als auch für die von OpenAI entwickelten Modelle GPT-2 und GPT-3. BERT bildet inzwischen das

Fundament für die Google-Suche, der Haupteinnahmequelle des Unternehmens. Führende KI-Ethiker*innen befürchten, dass das Vorgehen von Google einen abschreckenden Effekt auf die zukünftige Forschung zur KI-Ethik haben könnte. Denn viele der Forschenden in diesem Themenfeld sind bei den großen Tech-Konzernen beschäftigt.

Das KI-Ethikteam bei Google hatte inzwischen zwei weitere Abgänge zu verkräften. Die leitende Forscherin Alex Hanna und Dylan Baker aus der Software-Entwicklung haben sich Timnit Gebru und ihrem gemeinnützigen DAIR-Forschungsinstitut angeschlossen. In ihrer Ankündigung auf Medium, dass sie Google verlässt, nutzte Alex Hanna die Gelegenheit, um das „toxische“ Arbeitsumfeld bei Google zu kritisieren. Daneben beklagte sie den Mangel an Diversität in den Forschungseinrichtungen bei Google, vor allem das Fehlen von schwarzen Frauen. Sie kam zu dem Schluss, dass die Tech-Riesen ein „Whiteness“-Problem haben und für People of Color ein schreckliches Arbeitsumfeld bieten.

DR. ALEX HANNA



... ist Forschungsleiterin am *Distributed AI Research Institute (DAIR)*. Sie beschäftigt sich mit KI-Ethik und sozialen Bewegungen. Außerdem ist sie Vizepräsidentin der Vereinigung *Sociologists for Trans Justice* und Senior Fellow am *Center for Applied Transgender Studies*. Sie sitzt bei der *Human Rights Data Analysis Group* im Beirat und im Scholars Council des *UCLA Center for Critical Internet Inquiry*.

Monströse Kekse: Cookies als Instrument einer übergriffigen Technologie

Carbolytics ist ein Projekt an der Schnittstelle von Kunst und Forschung. Die Künstlerin Joana Moll hat dafür mit dem Hochleistungsrechenzentrum in Barcelona (*Barcelona Supercomputing Center*) zusammengearbeitet. *Carbolytics* soll ein größeres Bewusstsein dafür schaffen, wie groß die ökologischen Folgen einer allgegenwärtigen Überwachung durch Werbetechnologie (AdTech) sind.

Wie kam es dazu, dass Sie die CO₂-Bilanz von AdTech untersuchen wollten?

AdTech ist das grundlegende Geschäftsmodell des Internets, aber niemand kann die Kohlenstoffemissionen beziffern, die damit verbunden sind. Die Unternehmen, die dafür verantwortlich sind, müssen dafür zur Verantwortung gezogen werden. Dabei stellt die AdTech-Industrie nur die Spitze des Eisbergs dar. Hinter Cookies und Browsern versteckt sich ein gewaltiges Ökosystem. Da wir nicht sagen können, wie die Daten der Nutzer*innen innerhalb dieses Geschäftsmodells ausgewertet werden, liegt auch der dazu notwendige Energieverbrauch im Dunklen.

Was haben Sie bei dem Projekt herausgefunden?

In den meisten Fällen war es sehr schwierig herauszufinden, welche Organisationen hinter den Cookies stecken. Die am weitesten verbreiteten Cookies wie die von Google Analytics sind außerdem nicht diejenigen, die am schädlichsten für die Umwelt sind. Was allein die Cookies angeht, ist die umweltschädlichste Website Netflix, allerdings setzt Netflix nicht viele Cookies ein.



»Die meisten unserer Trans- und Interaktionen werden quantifiziert und kommerzialisiert. Durch AdTech wird alles zu Geld gemacht, was wir online tun: Alle Mausbewegungen, alle getippten Wörter, alle Klicks sind Umsatzquellen.«



Wie kann eine derart emissionsintensive Technologie wie AdTech so ein blinder Fleck sein?

Der ganze Prozess ist völlig unsichtbar. Alles geschieht so schnell und wir sehen nur unsere Geräte, nicht was durch die Geräte in Gang gesetzt wird. Wenn wir alle diese Vorgänge kennen würden, könnte das ganze System wahrscheinlich nicht funktionieren. Das Problem ist, dass wir im Prinzip keine Ahnung haben, wie alles abläuft. Ich habe in Slowenien eine Ausstellung mit einer Installation gehabt. Die Besucher*innen gingen dort in einen immersiven Raum hinein, wo sie von vier Projektoren angestrahlt wurden. Sie hatten diese ganzen Cookie-Logos auf ihren Gesichtern. Dieses Bild hat sich mir eingeprägt, weil es genau das illustriert, was gerade tatsächlich passiert. Viele konnten es nicht glauben, dass niemand weiß, was wir dagegen tun können, und dass uns als Nutzer*innen die Hände gebunden sind. Das Problem liegt im System. Wir sind einer bizarren Menge an Cookies ausgesetzt und gleichzeitig wissen wir kaum etwas darüber, wie sie funktionieren.



„Besucher*innen gingen bei meiner Installation in einen immersiven Raum hinein, wo sie von vier Projektoren angestrahlt wurden. Sie hatten diese ganzen Cookie-Logos auf ihren Gesichtern. Dieses Bild hat sich mir eingeprägt, weil es genau das illustriert, was gerade tatsächlich passiert.“

netflix.com	
tmall.com	139,9
microsoft.com	118,2
adobe.com	70,0
webmd.com	60,6
cloudflare.com	59,5
google.com	58,8
nypost.com	57,6
windows.com	57,2
newyorker.com	55,7
sharepoint.com	54,3
foxnews.com	48,3

Sie sehen AdTech als Phänomen des kognitiven Kapitalismus. Was bedeutet das?

Kognitiver Kapitalismus ist das ökonomische System, in dem wir uns alle bewegen. In diesem System wird Vermögen nicht nur durch materielle Güter generiert, sondern durch ein Zusammenspiel von unbestimmbaren Handlungen, Erfahrungen, Kommunikationsmustern und Wahrnehmungen. Die meisten unserer Trans- und Interaktionen werden quantifiziert und kommerzialisiert. Durch AdTech wird alles zu Geld gemacht, was wir online tun: Alle Mausbewegungen, alle getippten Wörter, alle Klicks sind Umsatzquellen.

Haben Sie etwas herausgefunden, was Sie nicht erwartet hätten?

Die Cookies, mit denen wir der Verarbeitung unserer Daten zustimmen, waren in unserer Analyse die am drittstärksten vertretene Gruppe von Cookies. Es ist ziemlich pervers, dass Datenschutz eine zusätzliche Problemebene mitbringt. Deshalb meine ich, dass Datenschutz und Nachhaltigkeit immer zusammen gedacht werden müssen. Sie sind zwei Aspekte desselben Problems: der mangelnden Rechenschaftspflicht von Unternehmen, die unsere Umwelt verschmutzen.

Google hat angekündigt, dass es 2030 CO₂-positiv sein will. Halten Sie das für realistisch?

Es ist allein schon problematisch, dass sie so etwas überhaupt legal behaupten dürfen, es ist nämlich schlicht unmöglich. Google ist ja nicht

nur irgendein riesiges Rechenzentrum. Google steckt in all unseren Geräten. Wie wollen sie diesen ganzen Energieverbrauch überhaupt quantifizieren? Wie soll diese Technologie allein klimaneutral werden? Sie saugen ständig Daten von unseren Geräten ab, wie soll das ohne Energieaufwand möglich sein? Als wir die Kohlenstoffemissionen von Cookies berechnen wollten, ist uns aufgefallen, dass es für Daten im Allgemeinen kaum unabhängige Einschätzungen und Gutachten über den damit verbundenen Energieverbrauch gibt. Ein weiteres Problem ist, dass sich die Forschung überhaupt nicht einig darin ist, wie sich der Verbrauch und die Emissionen quantifizieren ließen.

Was nehmen Sie aus dem Projekt mit?

Nachdem es abgeschlossen war, ist mir etwas Interessantes aufgefallen. Ich hätte mit einem deutlich größeren Medieninteresse gerechnet. Immerhin geht es um das wichtigste Geschäftsmodell des Internets. Zwar habe ich mit einigen



Die interaktive, webbasierte Installation CarboLytics demonstriert in Echtzeit das Ausmaß des durchschnittlichen globalen Cookie-Traffics. So bekommen wir ein Gefühl dafür, wie parasitär sich Cookies in unseren Geräten einnisten, um persönliche Daten zu extrahieren, und wie viel Energie dafür notwendig ist. (<https://carboLytics.org/web2x>)

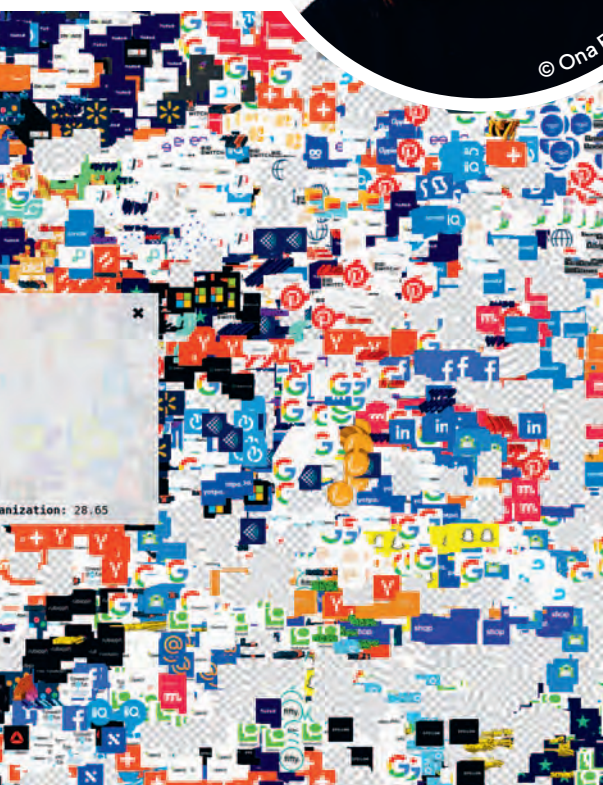
428,6

Die CO₂-emissionsreichsten Websites (Tonnen pro Monat)

großen Zeitungen gesprochen, aber keine von ihnen brachte dazu einen Beitrag. Das war sehr frustrierend für mich. Dann bemerkte ich, dass die *New York Times* und ähnliche Unternehmen Cookies einsetzen, die in unserem Ranking in den Top 20 auftauchen. Inzwischen bin ich mir ziemlich sicher, dass die Medien die Geschichte nicht aufgegriffen haben, weil sie ihr eigenes Geschäftsmodell betrifft. Es ist sehr schwer, die Öffentlichkeit auf etwas aufmerksam zu machen, wenn die Medien nicht mitspielen, weil sie sich damit selbst entblößen würden.



JOANA MOLL
Die Künstlerin
hinter Carbolytics



Über Carbolytics

(Auszug aus der Projektbeschreibung von Fernando Cucchietti, Joana Moll, Marta Esteban, Patricio Reyes, Carlos García Calatrava)

In den letzten zehn Jahren ist das Tracking des Online-Verhaltens von Nutzer*innen zu einem gewaltigen Geschäftsmodell geworden. Beim Online-Tracking werden Daten gesammelt, die Nutzer*innen online hinterlassen, beim Lesen der Nachrichten, bei Online-Einkäufen, bei ihren Social-Media-Interaktionen oder wenn sie Suchanfragen abschicken. Die Unternehmen, die Online-Tracking betreiben, befassen sich bislang nicht mit den Umweltfolgen ihres Trackings. Das massive Sammeln von Daten bildet oft die Grundlage auf der KI-Systeme erst angewandt werden können.

AdTech analysiert, verwaltet und distribuiert Online-Werbung. Es ist ein Geschäftsmodell, auf dem das gesamte Ökosystem der Datenwirtschaft beruht. 2021 erreichten die globalen Werbeausgaben über alle Plattformen hinweg 763,2 Milliarden US-Dollar. 2022 soll dieser Wert noch um zehn Prozent übertroffen werden. 2020 wurden 97,9 Prozent des globalen Umsatzes von Facebook und 80 Prozent des globalen Umsatzes von Google durch Werbeeinnahmen erzielt. Den chinesischen Markt ausgenommen, haben diese Unternehmen zusammen mit Amazon im Jahr 2022 einen Marktanteil von 80 bis 90 Prozent. Trotz der außerordentlichen Bedeutung von AdTech für die globale Wirtschaft sind die Methoden und Prozesse, die diesem Geschäftsmodell zugrunde liegen, weitestgehend unbekannt. Es ist daher sehr schwierig, diesen Bereich zu kontrollieren und zu regulieren.

Das Datensammeln durch AdTech ist häufig die Voraussetzung für KI-Anwendungen wie Empfehlungssysteme. Um die Nachhaltigkeit von KI zu bewerten, ist AdTech deshalb ein wichtiger Einflussfaktor.

Daten werden normalerweise durch Cookies und andere Tracking-Technologien gesammelt, die in Geräte, Websites, Apps und sonstige interaktive und audiovisuelle digitale Inhalte integriert sind. Obwohl Cookies auf den Geräten von Nutzer*innen erstellt und gespeichert werden, bleiben diese und andere Tracking-Technologien für Nutzer*innen intransparent und manchmal sogar unsichtbar. Auch wenn sie fast unbemerkt sind, setzen sie weltweit unzählige algorithmische Prozesse in Gang, die auf der Ausbeutung des Online-Verhaltens der Nutzer*innen beruhen und einen direkten Einfluss darauf haben, wie viel Energie die Endgeräte der Nutzer*innen verbrauchen.

Bei der das Carbolytics-Projekt begleitenden Forschung wurden die CO₂-Emissionen aller Cookies ermittelt, die zu den eine Millionen meistbesuchten Websites gehörten. Bei der Untersuchung konnten diesen Websites insgesamt mehr als 21 Millionen Cookies zugeordnet werden – bei einem einzigen Besuch. Diese Cookies konnten mehr als 1.200 unterschiedlichen Unternehmen zugeordnet werden. Im Durchschnitt entspricht dies 197 Billionen Cookies pro Monat, was wiederum etwa 11.442 Tonnen an CO₂-Emissionen monatlich bedeutet. Diese Zahlen spiegeln nur den Browser-basierten Cookie-Traffic wider und schließen keine anderen verhaltensbasierten WerbETOOLS wie App-Aktivitätsverfolgung oder Profilbildungsalgorithmen ein.

»Eine aufwendige Datenverarbeitung erfordert mehr Personal und teure Systeme und nimmt Zeit in Anspruch.«



Weniger bringt mehr:
Warum Künstlicher
Intelligenz eine
Datendiät guttun kann

Die Computerlinguistin Michaela Regneri untersuchte für den Hamburger Plattformanbieter OTTO KI-Anwendungen für den E-Commerce. Sie entwickelte Empfehlungsalgorithmen weiter, die für die Shopping-Plattform strategisch relevant sind und gleichzeitig in der KI-Entwicklung Ressourcen schonen. Als Verfechterin des Datenminimalismus möchte sie den Ansatz als Effizienzstrategie für Künstliche Intelligenz in der Data Economy etablieren. Dabei gilt es, möglichst wenige Daten so gezielt einzusetzen, dass keine Qualitätseinbußen bei der Leistung des Systems zu erwarten sind.



Was ist Datenminimalismus und welche Vorteile hat datenminimalistisches Arbeiten?

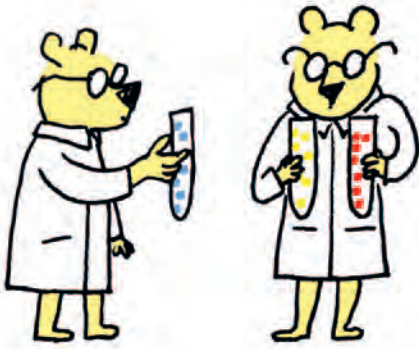
Minimalismus beziehen wir auf die Datenmengen, die wir mit KI verarbeiten. Wir sind von der Frage ausgegangen: Was sind eigentlich die Daten wert, die wir verarbeiten? Die Frage ist ökonomisch und ökologisch interessant, weil wir Effizienz anstreben. Wir wollen eine KI, die das tut, was sie soll. Aber sie soll das so effizient und effektiv wie möglich machen. Je kleiner also die Datenmenge für die gleiche Leistung, desto besser das Ergebnis.

Bei unserem einjährigen Projekt bei OTTO haben wir uns zunächst gefragt, wie wir berechnen können, was von den jeweiligen Datenpunkten oder vom gesamten Datentopf überhaupt wertvoll für die KI ist. Wir wollten also den nutzungsbasierten Datenwert ermitteln. Dazu haben wir praktische Experimente durchgeführt und die einzelnen Algorithmen daraufhin geprüft, ob es das System wirklich weiterbringt, wenn wir sie mit Daten füttern: Welche Daten führen beim Algorithmus zu besseren Ergebnissen und welche lassen ihn schlechter werden? Irgendwann trat eine Sättigung ein und es wurde ineffizient, neue Daten hinzuzufügen.

Wie funktioniert eine Verschlankung von Datenmengen?

Wir haben unterschiedliche Wege ausprobiert. Die einfachste Methode besteht darin, Daten wegzulassen und zu schauen, wie gut der Algorithmus mit der Hälfte der Daten, mit einem Drittel oder mit zwei Drittel der Daten funktioniert. So eine Sensibilitätsanalyse funktioniert im Prinzip wie ein umgekehrter Allergietest. Bei Allergien wird nach schädlichen Einflüssen gesucht, wir haben positive Einflussdaten identifiziert. Wenn ein Datenpunkt ausgelassen wird und das System wird besser, dann war es wahrscheinlich ein schlechter Datenpunkt. Wenn er ausgelassen wird und es passiert gar nichts, dann ist er für das Funktionssystem unnötig. Wenn das System schlechter wird, haben wir einen wertvollen Datenpunkt entdeckt.

Oft ist die Tendenz da, alles reinzustecken, was zur Verfügung steht. Aber gerade in industriellen Umgebungen liegen oft sehr viele Daten vor. Dann ist es manchmal gar nicht nötig, alles zu nutzen. Es lassen sich dann vielleicht noch minimale Verbesserungen von einem halben Prozent herbeiführen. Aber wirklich effizient kann der Algorithmus nur sein, wenn schädliche Daten entfernt werden. Wir hatten zum Beispiel konkret das Problem, dass manchmal Bots auf der Seite abgelassen werden. Die klicken wild im Shop herum oder sie klicken sogar speziell auf bestimmte Artikel, damit die so aussehen, als seien sie beliebt. So etwas schadet dem System natürlich. Manchmal schaden wir der KI sogar selber unabsichtlich, wenn wir zum Beispiel eine Marketingaktion wie den Deal des Tages haben. Das ist ein günstiger Artikel, den viele anklicken.



„Wir wählen diejenigen Daten aus, die wirklich wichtige Informationen enthalten.“

Das Problem ist aber, dass die Menschen sich nicht zwingend für den Artikel selbst interessieren, sondern nur für den Rabatt. Wir haben es also nicht mit ihrem natürlichen Kaufverhalten zu tun, sondern mit ihrer Reaktion auf etwas, was wir provoziert haben. Aus diesem Grund ist es dann schwierig, daraus Rückschlüsse für die KI zu ziehen.

Welche Vorteile bietet Datenminimalismus in wirtschaftlicher Hinsicht?

Datenminimalismus senkt Kosten. Eine aufwendige Datenverarbeitung erfordert mehr Personal und teure Systeme und nimmt darüber hinaus Zeit in Anspruch. Datenschutz und Compliance bringen einen zusätzlichen Aufwand mit sich. Je weniger Daten in der Cloud verwaltet werden, desto weniger Kosten verursacht die Cloud. Das Modell kann schneller trainiert werden und es hat ein größeres Innovationspotenzial, weil die Innovationsschleifen kürzer werden. Die Algorithmen können schneller getestet werden. Letzten Endes führt das Weglassen schädlicher Datenpunkte zu einer besseren Performance und dies bedeutet im wirtschaftlichen Kontext immer einen größeren Gewinn.

Kann Datenminimalismus auch soziale Vorteile haben?

Wo immer Daten gespart werden, werden auch Privacy und Cybersicherheit besser. Jeder sensible Datenpunkt, der unnötig bewegt wird, stellt ein unnötiges Sicherheitsrisiko dar. Außerdem kann Datenminimalismus dabei helfen, diskriminierende Muster in Datensätzen einzudämmen.

Inwiefern?

Normalerweise wird immer argumentiert, dass sehr viele Daten benötigt werden, damit der Datensatz ausgeglichen und nicht diskriminierend ist. Allerdings gibt es ein paar alte Techniken, um Diskriminierung zu vermeiden, die darauf basieren, Datensätze kleiner zu machen. Sie kommen aus der Medizin und der medizinischen Statistik. Die interes-

santeste Gruppe für die Medizin ist die kleinste, nämlich die der Kranken. Um bei medizinischen Anwendungen dieser Gruppe gerecht zu werden, besteht die Möglichkeit, die größere Gruppe der Gesunden zu verkleinern. Dies lässt sich datenminimalistisch verstehen: Wir wählen diejenigen Daten aus, die wirklich wichtige Informationen enthalten, und sorgen dafür, dass eine Gruppe nicht größer ist, als es für das angestrebte Ergebnis nötig ist. Dadurch bekommen relativ gesehen auch Minderheiten oder kleinere Gruppen ein höheres Gewicht.

Ein völlig konsequenter Datenminimalismus würde vollständig erklärbare KI-Systeme zur Folge haben. Zumindest in dem Sinne, dass wir ganz genau wissen würden, was einzelne Datenpunkte oder Datentöpfe im KI-System bewirken. Was wiederum bedeutet, dass wir auch vorhersagen könnten, welche Daten im Algorithmus einen diskriminierenden Einfluss haben werden.



Dimension:
Ökologische Nachhaltigkeit



Kriterium:
Energieverbrauch



Indikator:
Es werden Maßnahmen zur Reduktion der Datenmengen eingesetzt

Der Energieverbrauch eines KI-Systems unterscheidet sich je nach der jeweiligen Phase im Lebenszyklus. Die Entwicklungsphase neuer KI-Modelle kann trotz zunehmend effizienter Hardware extrem energieintensiv sein. Insbesondere das Finden der gewünschten Modellarchitektur erfordert mitunter eine immense Rechenleistung. Der Energieverbrauch in der Trainings- und vor allem der Anwendungsphase ist deutlich geringer. Jedoch werden diese Phasen im Gegensatz zur Entwicklung, die lediglich einen einmaligen Vorgang darstellt, mitunter massenhaft wiederholt. Datenminimalistische Ansätze sind eine Möglichkeit, den Energiebedarf in der Trainings- und Anwendungsphase zu reduzieren, indem die für Training und Anwendung genutzten Datensätze klein gehalten werden.

Wie beurteilen Sie Ihre Erfahrung bei OTTO im Rückblick? Welche Einsichten hatten Sie im Hinblick auf die Entwicklungsmöglichkeiten nachhaltiger KI?

Das Besondere an dem Projekt war für mich, überhaupt angewandte Forschung betreiben zu können. Universitäten stehen nur begrenzt Daten zur Verfügung. Für große Konzerne ist es kein Problem, viele Daten zu verwalten, weil sie die Rechen- und Finanzkapazitäten haben. Bei OTTO konnten wir diese Daten einsetzen, um selber KI-Algorithmen zu bauen. Wir wollten aber nicht unökologisch blind alles irgendwo reinstecken, nur weil es möglich war. Weniger Daten bedeuten nämlich weniger Rechenzeit und weniger CO₂-Emissionen. Einfach dadurch, dass weniger verbraucht wird. Der Energieverbrauch zieht sich durch die ganze Daten-Wertschöpfungskette. Der Stromverbrauch wächst linear mit der Datenmenge, die im Algorithmus verarbeitet wird.

Aber es gibt weder in der Wissenschaft noch in der Industrie zuverlässige und bewährte Methoden, die diese Folgen berücksichtigen und eindämmen. Und da müssen wir unbedingt etwas tun. Wenn das mit KI und Nachhaltigkeit was werden soll, müssen Regierungsstellen genau solche Kooperationen fördern, wie wir sie bei OTTO zwischen Wissenschaft und Industrie hatten.



Dimension:
Soziale Nachhaltigkeit



Kriterium:
Selbstbestimmung und Datenschutz



Indikator:
Ein Privacy-by-Design-Ansatz wird konsequent umgesetzt.

Ein konsequenter Privacy-by-Design-Ansatz berücksichtigt bereits in der Planung und Entwicklung von digitalen Technologien Datenschutz- und Sicherheitsinteressen. Die Datenschutz-Grundverordnung (DSGVO) fordert ein Privacy-by-Design sogar ein, lässt aber bei der Umsetzung Spielräume. Privacy-by-Design bedeutet zum Beispiel, dass Daten verschlüsselt und anonymisiert, sparsam eingesetzt und nicht unnötig zusammengeführt werden. Datenminimalismus ist somit ein integraler Bestandteil von Privacy-by-Design. Bei einer datenminimalistischen KI-Entwicklung können durch ein gutes Datenmanagement und eine hohe Qualität der ausgewählten Daten Diskriminierungsrisiken bei KI-Anwendungen reduziert werden. Gleichzeitig lassen sich dadurch Ressourcen bei der erforderlichen Rechenleistung einsparen.

**DR. MICHAELA
REGNERI**



... begeistert sich für Künstliche Intelligenz, insbesondere für Cognitive Computing mit Sprache, Bildern und allen anderen Arten von Mensch-Maschine-Interaktionen. Nach ihrer Promotion in Computerlinguistik war sie zunächst beim SPIEGEL-Verlag verantwortlich für Suche und Data Mining. Bei OTTO war sie seit 2016 zunächst als Produktmanagerin für Business Intelligence Analytics tätig. Sie interessiert sich besonders für Themen rund um Corporate Digital Responsibility und Organisationsentwicklung: KI und die Zukunft der Arbeit, KI und Mitarbeiter*innenentwicklung oder auch das nachhaltige Design von KI-Systemen.



Automatisches Vergessen

Verantwortungsvolles Datenmanagement und Maschinelles Lernen

Angesichts des immer weiter reichenden Einsatzes von KI-Systemen in unserem Alltag, muss unser Recht auf Vergessenwerden besser durchgesetzt werden – wie ein Fall aus den Niederlanden besonders drastisch darlegt. Dort hatte ein Paar durch eine Totgeburt das gemeinsame Kind verloren. Als zusätzliche traumatisierende Erfahrung wurde den beiden weiterhin dauernd Online-Werbung für Babyprodukte angezeigt. Auf Nachfrage antwortete das verantwortliche E-Commerce-Unternehmen, dass es nicht über die technischen Mittel verfügen würde, die Online-Empfehlungen auszuschalten.

Solche Online-Empfehlungen basieren auf automatisierten Entscheidungssystemen (ADM-Systemen). Das sind KI-Anwendungen, die immer häufiger eingesetzt werden, um automatisierte Entscheidungen zu treffen, die einen Einfluss auf unser Leben haben: zum Beispiel bei Kredit- und Leihgeschäften, medizinischen Diagnosen oder auch bei Einstellungsverfahren. Diese Systeme basieren auf Beispieldaten. Mithilfe verschiedener Techniken des Maschinellen Lernens werden

aus diesen Beispielen Vorhersagemodelle abgeleitet. Mit ihnen gehen bestimmte Risiken einher, die sich unter anderem aus den technischen Herausforderungen beim Verwalten der Daten ergeben, die von ADM-Systemen gespeichert und verarbeitet werden. Das obige Beispiel zeigt: Datenschutz und Selbstbestimmung über die eigenen Daten müssen verteidigt werden.

TECHNISCHER BIAS KANN DURCH DATENVERARBEITUNGSSYSTEME HERVORGERUFEN WERDEN.

Ein großer Teil der aktuellen Diskussion über die Fairness automatisierter Entscheidungen konzentriert sich auf soziale Vorurteile, die häufig über Beispieldaten Eingang in die Systeme finden. Sogenannter technischer Bias in den Daten entsteht durch bestimmte Verarbeitungsschritte. Das Risiko, dass technischer Bias in datengestützten Entscheidungssystemen zu finden ist, ist sehr hoch. Jedoch kann dieses Problem technisch gelöst werden, was im Fall des gesellschaftlichen Bias nicht möglich ist.

Technischer Bias resultiert aus den Laborbedingungen, unter denen Fachleute normalerweise die algorithmischen Komponenten von ADM-Systemen designen. Sie arbeiten mit einem festen und eindeutigen Trainingsdatensatz und probieren verschiedene algorithmische Ansätze aus, um ein Vorhersagemodell zu finden, das mit dem gegebenen Datensatz gut funktioniert. Sobald das vollständige ADM-System allerdings entwickelt wurde und in der Praxis eingesetzt wird, ändert sich der Modus der Datenproduktion für das Vorhersagemodell.

ADM-Systeme verarbeiten im Regelfall Daten aus unterschiedlichen Quellen, die kontinuierlich neue Daten produzieren. Das Vorhersagemodell muss regelmäßig an diese neuen Daten angepasst werden. Daten aus allen Quellen müssen im System zusammengeführt und in einer Form aufbereitet werden, die das Vorhersagemodell versteht. In diesem Verarbeitungsprozess kann technischer Bias durch Programmierfehler oder die falsche Darstellung von Gruppen in den generierten Daten auftreten. Auch scheinbar harmlose Abläufe können dazu führen: wenn etwa demografische Daten bei der Datenvorbereitung auf Grundlage von Postleitzahlen gefiltert werden, da der Wohnort mit sensiblen demografischen Daten wie Alter, Einkommen oder Ethnizität korreliert. In der Folge trifft das Modell möglicherweise weniger verlässliche Voraussagen für Gruppen oder Personen, da sie in den Daten durch Filtervorgänge falsch repräsentiert sind.



Dimension:
Soziale Nachhaltigkeit



Kriterium:
Selbstbestimmung und Datenschutz



Indikator:
Sicherstellung der informationellen Selbstbestimmung

Menschen sollten stets die Autonomie über ihre persönlichen Daten behalten. Zu diesem Zweck eignen sich einfache Visualisierungen sowie Benachrichtigungs-, Zustimmungs- und Widerrufsmechanismen. Nutzer*innen sollten darüber informiert werden, wenn KI-Systeme persönliche Daten verwenden oder sammeln. Sie sollten über den Einsatz ihrer Daten mitbestimmen können und nicht durch verhaltensbeeinflussende Mechanismen wie Nudging oder Dark Patterns in ihrem selbstbestimmten Handeln eingeschränkt werden.

DAS RECHT AUF VERGESSENWERDEN SIEHT EIN BALDIGES LÖSCHEN VOR.

Das Datenmanagement in ADM-Systemen weist eine orthogonale Dimension auf, was das Einhalten von Gesetzen betrifft, die den Schutz der Privatsphäre und Rechte der digitalen Selbstbestimmung garantieren. Ein bekanntes Beispiel ist das „Recht auf Vergessenwerden“ (Artikel 17 der Datenschutz-Grundverordnung). Es verpflichtet Unternehmen und Institutionen, die persönliche Daten verarbeiten, auf Anfrage Daten von Nutzer*innen ohne unnötigen Verzug durch angemessene und effektive Maßnahmen zu löschen. In der Praxis ist dies gegenwärtig ein eher zäher und langwieriger Prozess. Es kann beispielsweise bis zu zwei Monate dauern, bis Daten aus aktiven Systemen in der Google-Cloud gelöscht werden.

Deshalb brauchen wir ADM-Systeme mit Funktionen, die Lernvorgänge rückgängig machen. Das würde ihnen ermöglichen, Interaktionsdaten von Nutzer*innen nach einer entsprechenden Anfrage zu löschen und gleichzeitig die Voraussagen anzupassen, die das System erstellt. Beim Aktualisieren bestehender Vorhersagemodelle stellt dies jedoch hinsichtlich der Effizienz der Algorithmen und Berechnungen eine große Herausforderung dar.

Industrie und Forschung haben gerade erst begonnen, sich diesen Herausforderungen zu stellen. Selbst wenn die technologischen Grundlagen für ein verantwortungsvolles Datenmanagement in ADM-Systemen gegeben wären, würden wir Best-Practice-Lösungen benötigen. Um solche zu finden, müssen wir Zugang zu in der Praxis eingesetzten ADM-Systemen haben. Bislang bleibt uns dieser Zugang verwehrt, da die meisten dieser Systeme proprietär gehandhabt und von Privatunternehmen betrieben werden.

**DR. SEBASTIAN
SCHELTER**



... ist Juniorprofessor an der Universität Amsterdam und forscht dort an der Schnittstelle zwischen Datenmanagement und Maschinellem Lernen (ML). Er beschäftigt sich insbesondere mit datenspezifischen Problemen bei ML-Anwendungen: mit der automatischen Validierung von Datenqualität, mit Prüfverfahren für ML-Pipelines oder auch mit dem Design von ML-Anwendungen, die Daten vergessen. Sebastian Schelter veröffentlicht den Großteil des im Rahmen seiner Forschung entstehenden Quellcodes unter einer Open-Source-Lizenz. Er ist gewähltes Mitglied in der Apache Software Foundation.



KI sollte nur machen, was wir wollen

Julia Stoyanovich ist Informatikerin und forscht hauptsächlich zu verantwortungsvollem Datenmanagement und KI. Sie hat es sich zur Aufgabe gemacht, die Öffentlichkeit darüber aufzuklären, welchen Einfluss KI und Algorithmen auf unser Leben haben, zum Beispiel durch frei zugängliche Kurse in öffentlichen Bibliotheken oder auch durch Comics.

Sie haben das Center for Responsible AI gegründet. Dieses Zentrum für verantwortungsvolle KI soll „eine Zukunft schaffen, in der verantwortungsvolle KI die einzig vorhandene KI ist“. Wie sind wir da aktuell aufgestellt?

Können wir sagen, dass KI heute gleichbedeutend mit verantwortungsvoller KI ist? Leider sind wir noch nicht so weit. Wir sind sogar ziemlich weit davon entfernt. Wir müssen noch die richtige Mischung aus Lösungsansätzen finden, um weiterzukommen. Dazu gehören auch technische Lösungsansätze. Wir brauchen bessere Algorithmen und Programmierer*innen, die sich der allgemeinen Problemlage

»Die Automatisierung ist in diesem Bereich so weit fortgeschritten, dass wir alle bei der Arbeitssuche entweder bereits mit KI-Systemen zu tun hatten oder es in naher Zukunft zu tun haben werden.«



bewusst sind. Die Algorithmen selbst sind aber im Moment nicht das eigentliche Problem.

Was ist das eigentliche Problem?

Wir haben noch kein gemeinsames Verständnis darüber entwickelt, welche Funktion KI in unserer Gesellschaft erfüllen soll. Die Voraussetzung dazu wäre, uns zunächst einmal darüber klar zu werden, was KI überhaupt leisten kann, und was eben auch nicht. Die Allgemeinheit stellt sich KI noch immer als magisches Instrument vor, was den Entwickler*innen dieser Technologie sehr viel Macht verleiht. Aber die Entwickler*innen haben es auf diese Macht gar nicht unbedingt abgesehen, weil sie in der Regel nichtdafür ausgebildet sind, die rechts- und sozialwissenschaftlichen Implikationen der Technologie zu begreifen. Wir sollten Techniker*innen gar nicht erst in die Verlegenheit bringen, durch die Art, wie sie ihre Codes gestalten, gesellschaftlich relevante Grundsatzentscheidungen zu treffen.

Brauchen wir eine umfassendere Diskussion darüber, was KI für uns tun kann?

Wir geben gerade einen für alle zugänglichen Kurs in der öffentlichen Bibliothek in Queens, New York. Er heißt „We Are AI“ („Wir sind KI“). Dafür sind keine bestimmten Voraussetzungen notwendig, weder Programmierkenntnisse noch irgendwelche Abschlüsse. Die Teilnahme an dem Kurs ist kostenlos. Wir schenken den Teilnehmenden sogar Gutscheine. Bei dem Kurs wollen wir den Menschen klarmachen, dass KI genau das

macht, was wir wollen. Sowohl diejenigen, mit denen ich diesen Kurs entwickelt habe, als auch die Teilnehmer*innen sagen, dass das aber noch nicht der Fall ist. Wir haben keine Kontrolle darüber, wie KI eingesetzt wird. Wir sind ihr heute einfach ausgeliefert. Aber indem wir sagen „KI macht genau das, was wir wollen“, setzen wir dieser Realität etwas entgegen. Unser Ziel ist es, Einfluss auf die Gegenwart zu nehmen, indem wir sagen, wie wir uns die Zukunft vorstellen. Allein dadurch bewegen wir uns in die richtige Richtung.

Das Center for Responsible AI setzt sich für Transparenz auf allen Ebenen ein. Was bedeutet Transparenz im Hinblick auf KI-Systeme und wie können wir sie wirkungsvoll regulieren?

Wenn ich über Transparenz- und Interpretationsmechanismen nachdenke, habe ich immer ganz normale Menschen im Sinn. Ein Beispiel: Algorithmen werden in großem Stil bei Bewerbungsverfahren eingesetzt. Irgendwann suchen wir alle nach einem Job. Die Automatisierung ist in diesem Bereich so weit fortgeschritten, dass wir alle bei der Arbeitssuche entweder bereits mit KI-Systemen zu tun hatten oder es mit ihnen in naher Zukunft zu tun haben werden. Einen Zugang zu Jobs zu haben ist eine schiere ökonomische Notwendigkeit für uns alle. Deshalb müssen wir genau an dieser Stelle überlegen, wie wir den Einsatz von KI regulieren wollen. Alle Arbeitssuchenden müssen wissen, in welcher Weise sie von diesen Systemen betroffen sind. Die Unternehmen, die die Entwicklung solcher Systeme für ihre Zwecke in Auftrag geben, sollten ebenso

Nährwertkennzeichnungen als Inspiration für Transparenz bei Ranking-Algorithmen

ADM-Systeme berechnen oft Scores und Rankings, um ihre Ergebnisse darzustellen. Beispielsweise wird oft die Kreditwürdigkeit einer Person über die automatisierte Ermittlung eines Scores bewertet. Auch können Bewerber*innen von einem Ranking-Algorithmus nach vermeintlicher Qualifikation und Eignung für eine Stelle gelistet werden. Solche Scores und Rankings sind dafür bekannt, dass sie mitunter unfair, einfach zu manipulieren und wenig transparent sind. Zudem werden sie oft in Situationen eingesetzt, für die sie ursprünglich nicht entwickelt wurden – was zu ungenauen und problematischen Ergebnissen führen kann. Julia Stoyanovich hat aus diesem Grund zusammen mit weiteren Forscher*innen ein Bewertungssystem entwickelt, das in Anlehnung an Nährwertkennzeichnungen bei Lebensmitteln Informationen zu Ranking-Algorithmen bereitstellt. Die Anwendung *Ranking Facts* stellt über Visualisierungen Transparenz her. Sie zeigt unter anderem, welche Entscheidungskriterien in welcher Gewichtung in ein Ranking einfließen und wie beständig und fair die Berechnungen sind. Die Anwendung soll auch Laien dabei helfen, die Güte und Tauglichkeit eines Rankings bewerten zu können.



Quelle:

<https://dl.acm.org/doi/10.1145/3183713.3193568>



Dimension:

Soziale
Nachhaltigkeit



Kriterium:

Transparenz und Ver-
antwortungsübernahme



Indikator:

Öffentlich zugängli-
che Informationen über
Einsatz des Systems

Wenn Menschen Gegenstand einer automatisierten Entscheidung geworden sind, müssen die Betroffenen darüber in Kenntnis gesetzt werden. Ebenso müssen relevante Informationen über KI-Systeme öffentlich zugänglich sein, damit die Funktionsweise, Entscheidungskriterien und die technische Verlässlichkeit dieser Systeme von unabhängigen Instanzen geprüft werden können. Als Mindeststandard gilt, die relevantesten Informationen zur Zielsetzung eines Systems, den Nutzer*innen und Anwendungsfällen, zu Trainings- und Testdaten, den verwendeten Modellen, Feature-Selektion-Prozessen, Eingaben, Tests, Metriken etc. zu dokumentieren. Solche Informationen lassen sich in öffentlichen Registern hinterlegen.

wissen, was in diesen Systemen eigentlich passiert. Wenn es um den Zugang zur Arbeit geht, müssen wir alle wissen, was die Algorithmen genau tun. Identifizieren sie Kandidat*innen für eine Stelle, die vielleicht gar nicht wissen, dass ihr Profil auf einen Job passt? Gleichen die Algorithmen Lebensläufe mit Stellenausschreibungen ab? Bestimmen sie die Gehaltseinstufung? All diese Möglichkeiten sind mit unterschiedlichen Voraussetzungen und Fehlerquellen verbunden – und in allen Szenarien kommen unterschiedliche Instrumente zum Einsatz. Wenn wir mit bestimmten Interessensvertreter*innen sprechen, müssen wir ihnen aus diesem Grund sehr genau erklären, wo der Hund begraben liegt.

Wie sieht das in der Praxis aus?

Wir haben sowas wie eine „Lebensmittelkennzeichnung“ für Bewerber*innen entwickelt, um die verschiedenen Szenarien abzudecken, in denen Bewerber*innen und die Unternehmen, die einstellen, über ein Einstellungssystem miteinander interagieren. Ein Kennzeichnungslabel würde dann zu einer Stellenanzeige gehören, ich nenne es das „Profillabel“. Darin dokumentieren die Unternehmen, welche speziellen Qualifikationen für die entsprechende Stelle erforderlich sind. Menschen, die einen Job suchen, können dadurch einschätzen, ob sie für die entsprechende Stelle qualifiziert sind, welche ihrer Daten für das Screeningverfahren genutzt werden, welches Screening angewendet wird, welche Opt-Out-Optionen es gibt und in welcher Instanz bestimmte Informationen relevant sind. So ein Label würde es ermöglichen, dass Bewerber*innen der Verarbeitung ihrer Daten informiert zustimmen könnten. Sie würden einschätzen können, ob sie sich auf die Stelle bewerben wollen, oder ein alternatives Screening-Formular beantragen können. Mit der Zu- oder Absage bei einem Job würde ein anderes Label verbunden sein, das „Entscheidungslabel“. Bewerber*innen würden darüber informiert werden, was sie tun können, um in Zukunft größere Chancen bei einer ähnlichen Ausschreibung zu haben. Außerdem ist damit eine Möglichkeit verbunden, die Entscheidung anzufechten, zum Beispiel wenn Daten falsch verarbeitet wurden.

Sehen Sie in dieser Hinsicht noch Spielraum für zusätzliche Regulierungsmaßnahmen?

Wie alle Mitwirkenden im Center for Responsible AI bin ich sehr stolz darauf, dass wir große Befürworter*innen eines Gesetzes waren, das am 21. Dezember 2021 in New York City verabschiedet wurde. Dieses Gesetz hat den ersten Präzedenzfall zur Regulierung algorithmischer Bewerbungsprozesse geschaffen: Das Gesetz unter dem Namen „Local Law 144“ sieht vor, dass Tools für automatisierte Entscheidungen vor ihrem Einsatz auf Bias hin überprüft werden müssen. Außerdem müssen Bewerber*innen vor der Bewerbung darüber informiert werden, wenn Systeme verwendet werden, um ihre Daten zu screenen, und welche Daten aus ihrer Bewerbung dies betrifft. Das Gesetz wird im Januar 2023 in Kraft treten, so dass KI-Anbieter und Unternehmen, die KI einsetzen, ein Jahr lang Zeit haben, um sich auf die neue Gesetzeslage einzurichten. Das ist ein wichtiger erster Schritt.

**PROF. JULIA
STOYANOVICH**



...ist Professorin für Computer Science & Engineering sowie Data Science und außerdem Direktorin des Center for Responsible AI an der New York University. Ihre Forschung konzentriert sich auf die Operationalisierung von Fairness, Vielfalt, Transparenz und Datenschutz in allen Phasen des Datenmanagements. Sie war Mitglied der New York City Automated Decision Systems Task Force und trug zu einer Gesetzgebung bei, die den Einsatz von KI-Systemen in Einstellungsprozessen regulieren wird. Stoyanovich unterrichtet verantwortungsvolle KI und ist Mitautorin eines preisgekrönten Comics zu diesem Thema. Sie ist Empfängerin eines NSF CAREER-Preises und Mitglied der Association for Computing Machinery (ACM).



Dimension:
Soziale Nachhaltigkeit



Kriterium:
Nicht-Diskriminierung
und Fairness



Indikator:
Erfassung, Bewusstsein
& Sensibilisierung
für Fairness und Bias

Um Nicht-Diskriminierung und Fairness im Kontext von KI zu etablieren, muss mehr für das Thema sensibilisiert werden. Der erste Schritt besteht darin, Fairness fallspezifisch zu definieren und diese Definition im Planungs- und Entwicklungsprozess breit zu kommunizieren. Diskriminierungspotenziale können bereits in der Entwicklungsphase von KI-Systemen durch Impact-Assessments erkannt werden. Zur Messung von Fairness und Bias gibt es bewährte Methoden wie Equalized Odds oder Equal Opportunities. Damit kann Bias in Trainingsdaten, Inputdaten, Modellen, Methoden und Designs identifiziert werden, um in der Entwicklung nachzubessern. Fairness-Tests müssen geschützte Attribute wie Ethnizität, Hautfarbe, Herkunft, Religion, Geschlecht etc. berücksichtigen, um Diskriminierung auf Grundlage dieser Attribute zu verhindern. Das gilt auch für sogenannte Proxy-Variablen (Stellvertreter-Variablen), die mit den geschützten Attributen korrelieren.

Auf den folgenden Seiten: „Mirror, Mirror“ von Falaah Arif Khan und Julia Stoyanovich. Data, Responsibly 1 (2020).

Julia Stoyanovich und Falaah Arif Khan (die sowohl als Zeichnerin als auch als Datenwissenschaftlerin an der Arbeit beteiligt ist) haben bislang zwei Comic-Reihen veröffentlicht. Eine heißt „We Are AI“. Sie erschien zuerst auf Englisch und wurde kürzlich ins Spanische übersetzt. Weitere Übersetzungen werden folgen. Es ist die Begleitpublikation zu einem öffentlichen Erwachsenenbildungskurs gleichen Namens. Alle, die ein Interesse an Künstlicher Intelligenz und ihren gesellschaftlichen Folgen haben, können daran teilnehmen, ohne bestimmte Technikenkenntnisse mitbringen zu müssen. Comics sind allein dadurch ein außergewöhnliches Bildungsmedium, da sie humoristisch sind. Es erlaubt Stoyanovich und Arif Khan, KI-Technologie und ethische Überlegungen dazu allgemein zugänglicher zu machen. Die zweite Reihe heißt „Data, Responsibly“ und richtet sich an Studierende der Informatik und Interessierte. Sie ist etwas technischer ausgerichtet, da sie zur Pflichtlektüre eines Seminars über verantwortungsvolle Datenwissenschaft gehört, das Julia Stoyanovich an der New York University für Studierende anbietet.



Download:
https://dataresponsibly.github.io/comics/vol1/mirror_en.pdf

DIGITAL ACCESSIBILITY

DID YOU KNOW?

15% OF THE ENTIRE POPULATION EXPERIENCE SOME FORM OF DISABILITY- VISUAL, AUDITORY, MOTOR OR COGNITIVE. (3)

"THE POWER OF THE WEB IS IN ITS UNIVERSALITY. ACCESS BY EVERYONE REGARDLESS OF DISABILITY IS AN ESSENTIAL ASPECT"

-TIM BERNERS-LEE



SO, WHAT IS DIGITAL ACCESSIBILITY?

THIS VOLUME IS ABOUT ML AND DATA, SO YOU'RE PROBABLY IMAGINING ROBOTIC ARMS TRAINED ON HUNDREDS OF THOUSANDS OF RUNS OF SIMULATED MOVEMENT AND CUSTOMIZED TO THE WEARER'S MEASUREMENTS AND MOTION OF ACTION.



OR HOW ABOUT A FULLY AUTOMATED, HYPER SENSITIVE ROBOTIC ARMOUR THAT SELF-LEARNS AND AUTO-NAVIGATES FOR THE PHYSICALLY DISABLED?



OR GROUND-BREAKING, HYPER-INTELLIGENT GOGGLES FOR THE BLIND, THAT COLLECT THE DISTORTED IMAGE FROM THE WEARER'S RETINAS AND RECONSTRUCT IT TO A SHARP, 10800000 PIXEL IMAGE FOR SUPERHUMAN VISION?




MAYBE, IF ELON MUSK DECIDED TO GET INTO THE ACCESSIBILITY GAME...



The Anti-Elon 
@antiElon

Accessibility rocks!

 2.3K  9.2K  126K

IN OUR REALITY, DIGITAL ACCESSIBILITY IS FOCUSED ON MAKING SURE WEB PLATFORMS ARE EASILY NAVIGABLE AND USABLE BY PEOPLE WITH ANY KIND OF DISABILITY

IT IS THIS VERY WORK THAT MAKES SURE THAT THE IMAGE YOU JUST POSTED ON INSTAGRAM HAS CAPTIONS



SO THAT THE BLIND USERS OF THE PLATFORM CAN ALSO PARTAKE IN YOUR TRIUMPH OVER THAT SOURDOUGH RECIPE.

OR WHEN YOU DROP A NEW TUTORIAL VIDEO FOR ALL ONE SQUILLION OF YOUR SUBSCRIBERS TO ENJOY,

HOW TO BUILD AGI



IT IS THIS WORK THAT CONVERTS YOUR VOCAL PEARLS OF WISDOM INTO TEXT FOR YOUR DEAF FOLLOWERS.



ACCESSIBILITY NEEDS TO BE A FUNDAMENTAL DESIGN PRINCIPLE FOR BUILDING WEBSITES AND SOFTWARE,

BUT IN OUR QUEST FOR OPTOPIA, IT IS USUALLY OVERLOOKED.

WITHOUT **A11IES** (4), THE DEMOGRAPHIC THAT WAS HOLDING ON TO THE ACCESSIBILITY ROPE IS NOW CUT OFF.

LET'S GET RID OF THE **MAGPIE MENTALITY?**

FOR YOUR NEXT FUN DATA SCIENCE PROJECT, INSTEAD OF SOME COMMUNITY-OVERFITTED IMAGE RECOGNITION CHALLENGE, MAYBE CHOOSE AN **OPEN PROBLEM IN DIGITAL ACCESSIBILITY**, SUCH AS AUTOMATIC VIDEO CAPTIONING. THEN HOPEFULLY ONE DAY THERE WILL BE **"NO MORE CRAPTIONS"** (5)

PART 2: GHOSTS IN THE SHELL

(WHO ARE WE BUILDING MODELS FOR?)

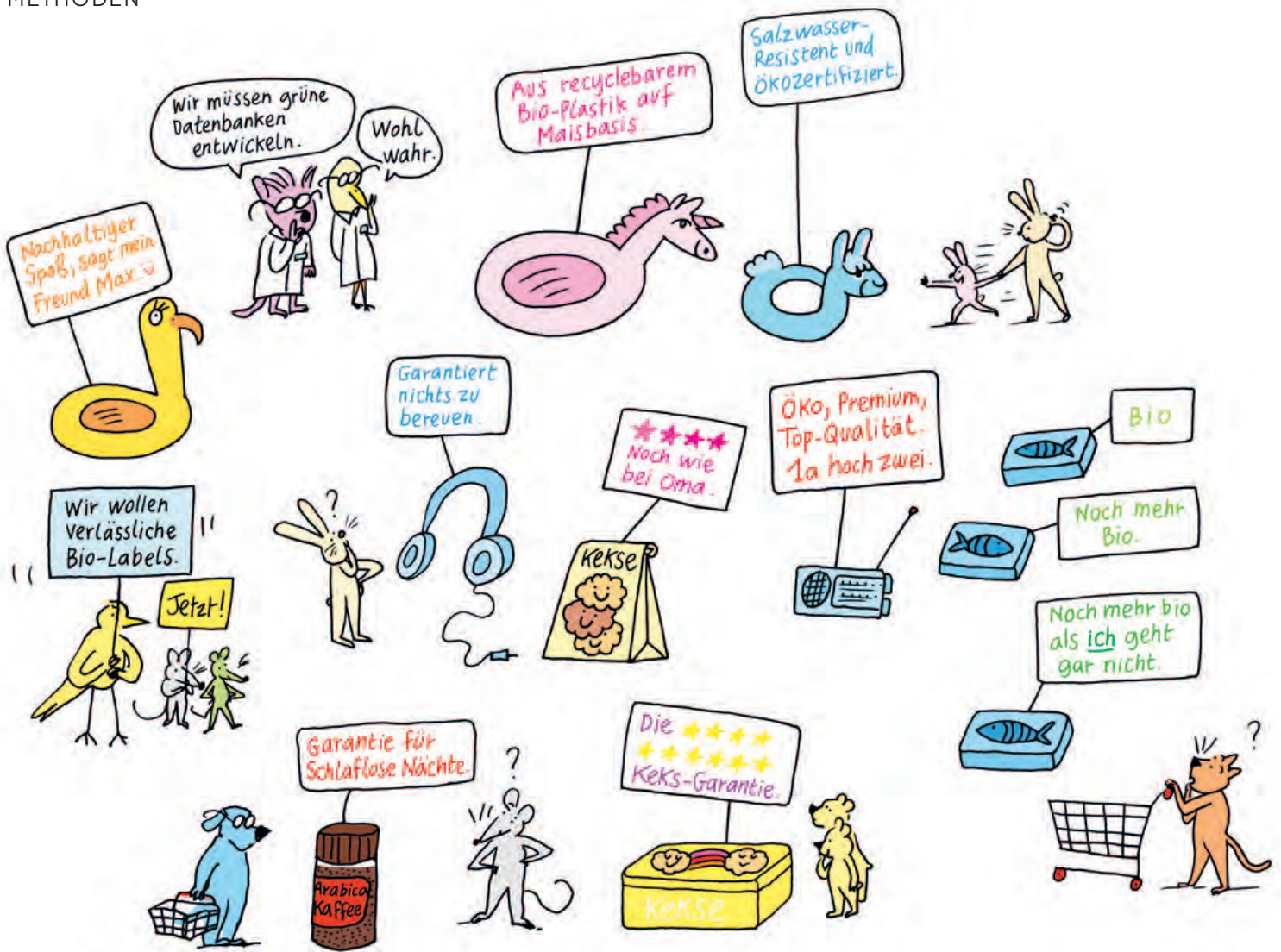
WE HAVEN'T YET FIGURED OUT HOW TO MAKE EXISTING DIGITAL PLATFORMS ACCESSIBLE TO EVERYONE, YET WE'RE ALREADY JUMPING TO FORGE A NEW "INTELLIGENT" CLASS OF WEB APPLICATIONS.

WE'RE SO CAUGHT UP IN THE **"HOW"** (USING ML/AI/DL/DS !!!) THAT WE FORGET TO ASK, **"FOR WHOM"?**

WHEN PLATFORMS ARE NOT DESIGNED FOR EVERYONE, THEY GIVE OFF THE STENCH OF **"ENCODED INHOSPITALITY"** (6).

SEEMINGLY INNOCUOUS THINGS SUCH AS **POP-UPS** AND **EXPIRING FORMS** ON WEBSITES COMPLETELY HIJACK THE ONLINE EXPERIENCE OF USERS WITH DISABILITIES WHO RELY ON SCREEN READERS.





Grüne Datenbanken für Algorithmen

Der Green Consumption Assistant

Ein vermeintlicher Mangel an nachhaltigen Produkten kann beim Online-Kauf längst nicht mehr als Entschuldigung dafür dienen, sich für nicht nachhaltige Optionen zu entscheiden. Das Projekt *Green Consumption Assistant* nutzt die bestehenden Infrastrukturen für Maschinelles Lernen im Online-Handel, um Verbraucher*innen beim Finden nachhaltiger Produkte zu unterstützen.

Das Projektteam um Tilman Santarius und Maike Gosse von der TU Berlin, Felix Biessmann von der Berliner Hochschule für Technik sowie Mitarbeiter*innen der grünen Suchmaschine Ecosia hatte zwei Dinge beobachtet. Erstens geben Konsument*innen zwar an, dass sie beim Online-Kauf nachhaltigere Entscheidungen treffen wollen, aber sie tun es nicht. Zweitens könnten im Online-Handel bereits genutzte Systeme des Maschinellen Lernens eingesetzt werden, um nachhaltige Konsumententscheidungen viel einfacher zu machen. Es fehlen jedoch wesentliche und um-



Dimension:
Ökologische
Nachhaltigkeit



Kriterium:
Nachhaltigkeits-
potenziale in der
Anwendung



Indikator:
Förderung nachhaltiger
Produkte

fassende Daten zu nachhaltigen Produkten für diese Systeme.

Die Lösung: grüne Datenbanken, damit Algorithmen bei ihren automatisierten Entscheidungen Nachhaltigkeitsaspekte berücksichtigen können. Automatisierte Empfehlungssysteme könnten beim Online-Shopping nachhaltige Produkte prominenter anzeigen als nicht nachhaltige Produkte. Wenn solche Datenbanken frei zugänglich wären, könnten Verbraucher*innen sich über Nachhaltigkeitszertifizierungen und -kennzeichnungen informieren und nachhaltigere Kaufentscheidungen treffen.

Das Projektteam des *Green Consumption Assistant* will genau diesen Mangel beheben. Es hat die *GreenDB*-Datenbank entwickelt, die Informationen zur Nachhaltigkeit von Konsumgütern enthält. Die Datenbank wird wöchentlich aktualisiert und enthält aktuell über 220.000 Produkte der größten Online-Händler aus mehreren europäischen Ländern. Im Gegensatz zu früheren Ansätzen umfasst die *GreenDB* Produkte, für die eine große Nachfrage besteht. Eine Analyse der Suchprotokolle von Ecosia-Nutzer*innen diente als Auswahlbasis für die 26 Produktkategorien, derzeit hauptsächlich Mode und Elektronik.

Die Datenbank zeigt Informationen darüber an, wie verlässlich die vorhandenen Nachhaltigkeitsinformationen über ein bestimmtes Produkt sind: ob es mit einem sehr glaubwürdigen, von externen Stellen verifizierten Nachhaltigkeitslabel ausgezeichnet ist oder mit einem nicht-verifizierten Label. Die Datenbank wird auf der Ecosia-Suchseite in der Registerkarte „Shopping“ eingesetzt, um umweltfreundliche Produkte hervorzuheben. Die *GreenDB* ist für Forschungszwecke öffentlich zugänglich und kann für andere KI-Anwendungen wie Such- und Empfehlungssysteme verwendet werden, um transparen-

ter zu machen, wie vertrauenswürdig Nachhaltigkeitsinformationen sind.

Die *GreenDB* kann nicht nur nachhaltigen Konsum fördern. Sie gibt auch eine Übersicht darüber, wie verfügbar Nachhaltigkeitsinformationen im Online-Handel sind. Nur 14 Prozent der Datenbankprodukte sind mit glaubwürdigen Nachhaltigkeitslabels zertifiziert: Das zeigt, wie schwierig es für Verbraucher*innen ist, nachhaltige Produkte zu erkennen. Die vielen nicht zertifizierten Labels verhindern eine Vergleichbarkeit und sorgen für Verwirrung und Unsicherheit. Mehr Klarheit und Informationen sind dringend erforderlich: Wir brauchen insbesondere politische Initiativen, die die vielen unzureichenden Labels auf ein mögliches Greenwashing hin untersuchen.

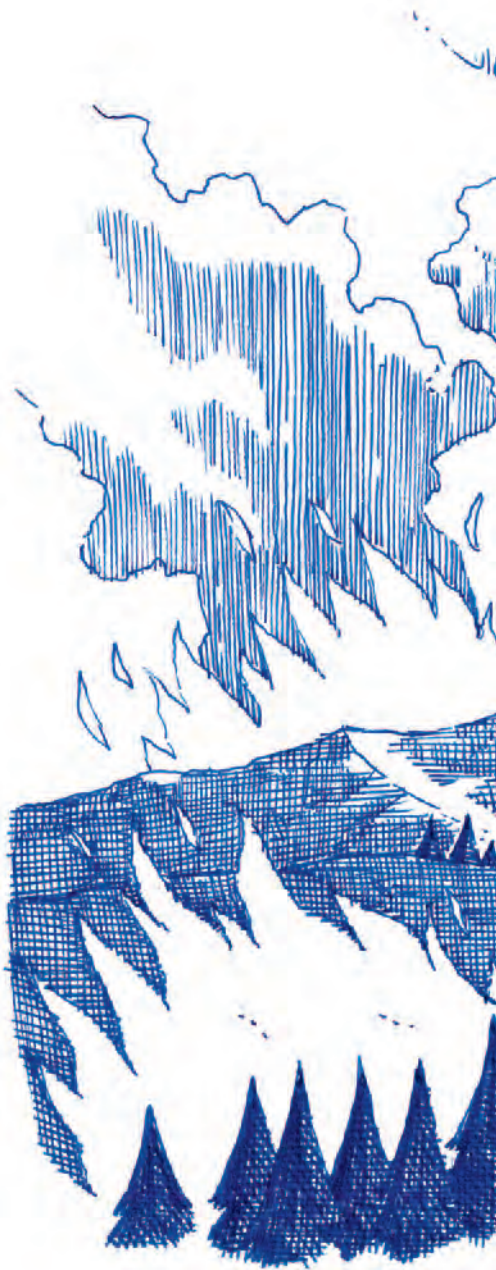
Green Consumption Assistant

Der *Green Consumption Assistant* (GCA) unterstützt Konsument*innen dabei, nachhaltigere Kaufentscheidungen im Internet zu treffen. Der GCA zeigt dazu auf der Suchmaschine Ecosia grüne Produktalternativen an und informiert über nachhaltigere Alternativen, beispielsweise in Form von Hinweisen auf Repair-, Verleih- oder Sharing-Optionen. Basis für die Empfehlungen des GCA ist eine mithilfe von Maschinellen Lernverfahren entwickelte Produktdatenbank (*GreenDB*) mit ökologischen und sozialen Nachhaltigkeitsinformationen.

Der GCA ist ein Kooperationsprojekt zwischen der Technischen Universität Berlin, der Berliner Hochschule für Technik und der grünen Suchmaschine Ecosia. Es wird als Leuchtturmprojekt für Künstliche Intelligenz im Einsatz für ökologische Herausforderungen vom Bundesumweltministerium gefördert.

Gefangen in einer untragbaren Infrastruktur

KI-Systeme bestehen nicht nur aus Daten, Knoten in Netzwerken oder Reencode – wie beliebte Visualisierungen es uns glauben machen wollen. Hinter ihnen verbirgt sich die Ausbeutung natürlicher und sozialer Ressourcen. Die KI-Ethikerin Aimee van Wynsberghe hält Regulierung für den einzigen Weg, die großen Tech-Unternehmen für die versteckten Nachhaltigkeitskosten von KI zur Verantwortung zu ziehen.



In Ihrem Artikel über nachhaltige KI unterscheiden Sie zwischen KI für Nachhaltigkeit und der Nachhaltigkeit von KI. Warum ist es wichtig, über die Nachhaltigkeit von KI zu sprechen?

Ich wollte die Menschen daran erinnern, dass eine physische Infrastruktur notwendig ist, um KI-Technologie herzustellen und einzusetzen. Und diese physische Infrastruktur ist zum gegenwärtigen Zeitpunkt alles andere als nachhaltig, da sie Kohlendioxid-Emissionen zur Folge hat. Wenn Algorithmen trainiert und eingesetzt werden, benötigen wir außerdem für die dafür notwendige

Infrastruktur Mineralien. Diese Mineralien stecken in Batterien und Mikroprozessoren. Die Bedingungen, unter denen Menschen arbeiten müssen, um diese Mineralien zu gewinnen, sind furchtbar. Deswegen werden sie auch „Blutmineralien“ genannt. Auch die für die Infrastruktur notwendige Wasser- und Landnutzung ist problematisch. Und was machen wir mit all dem Elektroschrott, den Servern? Wir laden sie in asiatischen Ländern ab und die Menschen dort leiden unter den Umweltfolgen. Was ich also mit der Unterscheidung zum Ausdruck bringen wollte: Es reicht nicht zu sagen, dass wir die Technologie einsetzen wollen, um nachhaltiger zu leben, wir müssen überprüfen, ob sie selbst nachhaltig ist.



Der Diskurs über KI für Nachhaltigkeit ist sehr lebendig. Warum ist das beim Diskurs über die Nachhaltigkeit von KI nicht so?

Erst wenn wir die versteckten Kosten aufgedeckt haben, können wir sagen, wie teuer uns KI zu stehen kommt. Das würde die Blase zum Platzen bringen, die Unternehmen wie Google, Amazon, Facebook und Apple geschaffen haben. Es hat aber strategische Gründe, die wirklichen Kosten zu verbergen, denn Europa hat allein 2020 3,2 Milliarden Euro in KI investiert. Die großen Konzerne wollen sich der Verpflichtung entziehen, diese Kosten zu dokumentieren und ihre Beschaffungsketten offenlegen zu müssen. Der Diskurs über die Nachhaltigkeit von

Desirable Digitalisation

Beim Projekt *Desirable Digitalisation: Rethinking AI for Just and Sustainable Futures* wird KI-Entwicklung auf der Grundlage ethischer Prinzipien untersucht. Unter der Leitung von Prof. Aimee van Wynsberghe arbeiten Wissenschaftler*innen mit der KI-Industrie zusammen, um nachhaltige und gerechte Prinzipien für das Design und die Bildung von KI zu entwickeln. Das Projekt startete im April 2022 und hat eine Laufzeit von fünf Jahren.



Dimension:
Ökologische
Nachhaltigkeit



Kriterium:
Indirekter
Ressourcenverbrauch



Indikator:
Kennzahlen des
Verwertungsbetriebs

KI geht nicht voran, weil a) die großen Konzerne nicht am damit verbundenen Aufwand interessiert sind und b) wir es uns nicht eingestehen wollen, dass die verborgenen ökologischen Kosten sehr viel größer sind als wir glauben mögen. Der Diskurs lahm aber bereits durch die schiere Komplexität des Problems. Darüber hinaus lässt sich unsere Gesellschaft nur ungern ihren Enthusiasmus für KI-Technologie trüben. Ökologische Folgekosten passen da einfach nicht ins Bild.

Glauben Sie, dass die großen Tech-Unternehmen dieses Problem in absehbarer Zeit angehen werden?

Nein, uns wird wahrscheinlich nichts anderes als Regulierung übrigbleiben. Wenn wir es den Unternehmen überlassen wollen, dürfen wir nicht vergessen, dass es ihnen in erster Linie um Gewinne geht. Sie sind ihren Aktionär*innen gegenüber in der Bringschuld und richten KI-Ethikräte bloß strategisch ein, um sich einen verantwortungsbewussten Anstrich zu geben. Wenn wir die großen Konzerne dazu bringen wollen, die ökologischen Folgen von KI zu ermitteln, brauchen wir Regulierung. Erst dann können wir beurteilen, wie ernst die Lage wirklich ist.

Wie ernst ist denn die Lage?

Wir müssen sofort handeln. Der jüngste IPCC-Bericht hat für die Menschheit Alarmstufe Rot ausgerufen. Wir müssen jede Technologie, die wir nutzen, daraufhin über-

Um Computerhardware zu produzieren, werden sogenannte Konfliktrohstoffe oder Konfliktmineralien benötigt – seltene Erden oder Edelmetalle, deren Abbau mit Menschenrechtsverletzungen, verheerenden Arbeitsbedingungen und Umweltverschmutzungen einhergeht. Wenn bei der Entsorgung der Hardware die darin enthaltenen Wertstoffe zurückgewonnen werden, müssen sie nicht für neue Hardware erneut abgebaut werden. Zertifizierte Recycling-Fachbetriebe können die Hardware in die enthaltenen Wertstoffe auftrennen und so wiederverwertbar machen. Alternativ kann gebrauchte Hardware von *Original Equipment Manufacturers* (OEMs) oder Aufarbeitungsunternehmen gesammelt und wiederverwendet werden. Sie entnehmen einzelne Hardwarekomponenten und setzen sie in gebrauchte oder neue Produkte wieder ein. Der gewichtsmäßige Anteil von recycelten oder wiederverwendeten Materialien ist eine wichtige Kennzahl, um die ökologische Nachhaltigkeit von Hardware zu beurteilen.

prüfen, welche Folgen sie für die Umwelt hat. Auf der ganzen Welt wurden bereits Milliarden in KI investiert, in jeder erdenklichen Branche. Das betrifft nicht nur die Algorithmen, sondern auch die Infrastruktur, auf der sie beruhen. Es gibt keinen Weg mehr zurück. KI ist eine so um sich greifende Technologie, dass sie global allgegenwärtig geworden ist. Wenn wir nicht sofort handeln, werden die Emissionen nicht mehr zu stoppen sein, da die Infrastruktur selbst nicht nachhaltig ist. Die Folgen werden für die ganze Weltgemeinschaft zu spüren sein.

Es ist oft zu hören, dass es für Branchenakteure nicht so einfach wäre, den Nachweis für die Nachhaltigkeit von KI zu erbringen. Glauben Sie, dass es für die Industrie machbar wäre, nachhaltige KI zu designen und herzustellen?

Ich höre in Diskussionen rund um KI und Ethik oft, dass ethische Bedenken bloß die KI-

Innovation aufhalten würden, weil damit unweigerlich diese ärgerlichen Kontrollmechanismen einhergingen. Dabei setzen wir uns bloß dafür ein, soziale und ökologische Standards für Innovationen festzulegen. Warum sollte das Innovationspotenziale eindämmen? Vor nicht allzu langer Zeit schien KI selbst ein Ding der Unmöglichkeit zu sein und nun ist sie überall. Warum sollte es so eine unüberwindliche Hürde sein, KI auf die richtige Art und Weise einzusetzen? An dieser Stelle kommt dann die Regulierung ins Spiel. Die Regierungen, die Europäische Kommission und das Europäische Parlament müssen die

großen Tech-Konzerne dazu verpflichten, die von ihnen verursachten Kohlendioxid-Emissionen nachzuverfolgen und zu messen. Die Konzerne müssen auch etwas an den Bedingungen ändern, unter denen sie die infrastrukturell notwendigen Mineralien beschaffen, auch wenn es schwierig ist. Wenn es nicht nachhaltig geht, dann dürfen wir es nicht machen. Aber wir können Innovationen finden, mit denen es auch nachhaltig geht.

Auf europäischer Ebene wird gerade eine mögliche KI-Verordnung diskutiert. Sie soll die mit KI verbundenen Risiken eindämmen und fundamentale Grundrechte stärken. Was halten Sie von der Verordnung in ihrer jetzigen Fassung?

Für Nachhaltigkeit bringt sie rein gar nichts. Mein größtes Problem mit der KI-Verordnung ist, dass sie ökologische Risiken gar nicht als Risiken begreift, die einer Risikoeinschätzung bedürfen. Die Nachhaltigkeit von KI muss transparenter werden. Dann könnten wir endlich über CO₂-Höchstgrenzen oder Grenzen für Trainingsdurchgänge sprechen, bei denen eine bestimmte Zeit lang nur eine gewisse Anzahl von GPUs für das Training der Algorithmen verwendet werden darf. Dafür brauchen wir aber Daten darüber, wieviel Strom beim Algorithmustraining verbraucht wird oder wieviel Wasser benötigt wird, um die Server zu kühlen. Ohne die entsprechenden Daten bewegen wir uns im Reich der Spekulation. Darum setze ich mich für verbindliche Mess- und Transparenzvorgaben ein.



Dimension:
Ökonomische
Nachhaltigkeit



Kriterium:
Arbeitsbedingungen
und Arbeitsplätze



Indikator:
Faire Löhne
entlang der
Wertschöpfungskette

Problematische Arbeitsbedingungen bei der Entwicklung von KI existieren nicht nur bei der Hardwareproduktion, sondern auch bei der Datenaufbereitung. Die Datensätze, die für das Training von KI-Systemen benötigt werden, müssen meist erst von sogenannten Crowd- oder Clickworkern *gelabelt*, das heißt klassifiziert und annotiert werden. Diese Menschen bearbeiten oft unter prekären Bedingungen massenhaft kleine Aufgaben am Computer (per Klick), ohne fest angestellt zu sein. Bei der Entwicklung als auch beim Einkauf von KI sollte darauf geachtet werden, dass die Arbeitsbedingungen entlang des gesamten Lebenszyklus einer KI fair sind. Das beinhaltet eine angemessene Entlohnung, gute Arbeitsbedingungen oder auch Aufstiegs- und Weiterbildungsmöglichkeiten – auch für Clickworker.

**PROF. DR. AIMEE
VAN WYNSBERGHE**



... forscht in den Bereichen Informations- und Kommunikationstechnologie und Robotik. Nach Assistenzprofessuren in Twente und der Technischen Universität Delft trat sie 2020 eine Professur für Ethik und Technologie an der TU Delft an. Für ihre Forschungsarbeiten und ihren Beitrag zum wissenschaftlichen Dialog wurde die Philosophin 2018 mit einem L'Oréal UNESCO For Women In Science-Award ausgezeichnet. Im Februar 2021 trat sie die von der Alexander von Humboldt-Stiftung neu konzipierte Professur *Applied Ethics of Artificial Intelligence* an.

Ein gefährliches Ausblenden von Umweltrisiken

Die KI-Verordnung der EU

Die EU hat die Notwendigkeit erkannt, die mit KI verbundenen Risiken auf politischer Ebene anzugehen. Doch wenn es um den Ressourcenverbrauch und die Umweltauswirkungen der Technologie geht, ist die KI-Verordnung auf beiden Augen blind.

Im April 2021 veröffentlichte die Europäische Kommission ihren Entwurf zur KI-Verordnung (engl: Artificial Intelligence Act, AI Act). Sie reagierte damit auf die immer dringlichere Notwendigkeit, Künstliche Intelligenz zu regulieren. Das Ziel der Verordnung ist es, auf die fundamentalen Risiken einzugehen, die beim Einsatz von KI für die Gesellschaft und die in ihr lebenden Menschen bestehen. Allerdings enttäuscht der Entwurf, was die Risiken für die Umwelt angeht. Dies wiegt umso schwerer, als die Kommission in ihrem der KI-Verordnung vorausgegangen Weißbuch hervorgehoben hatte, dass KI-Entwicklung umweltfreundlich stattfinden müsse.

Die KI-Verordnung schlägt eine Regulierung vor, die in der EU Regeln zum Schutz der Sicherheit, Gesundheit und Grundrechte beim Einsatz von KI vereinheitlichen soll. Gleichzeitig sollen Innovationen auf dem Gebiet gefördert werden. Dazu verfolgt die KI-Verordnung einen risikobasierten Ansatz. Sie schlägt unterschiedliche Regeln für KI-Systeme je nach ihrer Risikoeinstufung vor. Dabei berücksichtigt sie jedoch in keiner Weise die Umweltrisiken, die sich aus der Entwicklung, Inbetriebnahme und dem Einsatz von KI-Systemen ergeben.



In Art. 37 ist ausdrücklich festgelegt, dass die Europäische Union den Schutz der Umwelt bei ihrer Politikgestaltung berücksichtigen muss.

Daher sind wir der Meinung: Wenn die KI-Verordnung unsere Sicherheit, Gesundheit und Grundrechte schützen soll, würde die Europäische Union fahrlässig handeln, wenn sie den Umweltschutz vernachlässigt. Die europäischen Institutionen müssen sich darüber im Klaren sein, welche Umweltrisiken KI mit

sich bringt, denn die Beweislage ist erdrückend. Der Ressourcenverbrauch der Systeme ist teils extrem hoch, ebenso der Ressourcenverbrauch der für ihren Betrieb notwendigen Infrastruktur. In ihrem Regulierungsvorschlag wird die Kommission daher nicht der *Charta der Grundrechte der Europäischen Union* gerecht, die in Art. 37 ausdrücklich festlegt, dass die Europäische Union den Schutz der Umwelt bei ihrer Politikgestaltung berücksichtigen muss.

Im Licht des vielversprechenden Weißbuchs erscheint der Entwurf zur KI-Verordnung wie eine verpasste Gelegenheit. Dort wurde hervorgehoben, dass die Folgen, die der Einsatz von KI-Systemen auf die Umwelt hat, entlang der gesamten Lieferkette und über den gesamten Lebenszyklus hinweg gebührend berücksichtigt werden müssen – etwa in Bezug auf den Ressourcenverbrauch, der für das Trainieren der Algorithmen oder das Speichern von Daten notwendig ist. Im gegenwärtigen Entwurf zur KI-Verordnung sind keine verpflichtende Umweltauflagen für Hersteller und/oder Anwender zu finden. In der jetzigen Fassung dürfen Hersteller Verhaltenskodizes mit freiwilligen Selbstverpflichtungen formulieren und anwenden, unter anderem hinsichtlich Nachhaltigkeit und Umweltschutz. Aber freiwillige Selbstverpflichtungen können kaum als angemessene Reaktion auf eine zunehmend allgegenwärtige und ressourcenaufwendige Technologie wie KI angesehen werden.

Daher verpasst die KI-Verordnung bisher leider eine entscheidende Chance, die Entwicklung und den Einsatz von KI-Systemen nachhaltig, ressourcenschonend und unter Beachtung der planetaren Grenzen zu gestalten. Dies ist aber eine notwendige Maßnahme im kollek-

tiven Kampf gegen die Klimakrise. Daneben widerspricht der Ansatz der Kommission den Zielsetzungen des europäischen Green New Deal und ähnlicher Initiativen der EU. Die europäischen Institutionen verhandeln nach wie vor über die KI-Verordnung. Noch können sie angemessen auf eines der größten Risiken von KI-Technologien reagieren. Ein erster notwendiger Schritt wäre, die enormen Umweltrisiken von KI anzuerkennen. Die KI-Verordnung sollte sie zu einem relevanten Kriterium in der Risikoeinstufung von KI-Systemen machen. Darüber hinaus sollten Organisationen, die KI-Systeme entwickeln und implementieren, ihren KI-bezogenen Ressourcenverbrauch überwachen und die entsprechenden Daten der Öffentlichkeit zugänglich machen. Sie sollten außerdem dazu verpflichtet werden, wirksame Schritte zu einer umweltfreundlicheren KI-Entwicklung einzuleiten.

Wir brauchen mehr Erkenntnisse über den tatsächlichen Ressourcenverbrauch von KI, um zu einer stärker auf belastbaren Daten basierten Regulierung von KI-Technologien zu kommen. Die KI-Verordnung bietet hier eine Chance, die nicht verpasst werden sollte. Es liegt nun am Europäischen Parlament und an den EU-Mitgliedsstaaten, die unzureichende Risikobewertung der Kommission wieder rückgängig zu machen.

NIKOLETT ASZÓDI



... ist Policy & Advocacy Managerin bei AlgorithmWatch. Der Fokus in ihrer Arbeit liegt auf dem Einsatz von Systemen des automatisierten Entscheidens (ADM) im öffentlichen Sektor sowie auf horizontalen Regulierungen der EU im Bereich ADM – insbesondere die KI-Verordnung der EU (AI Act).

Eine ökologisch nachhaltige KI braucht Regulierung

Künstliche Intelligenz (KI) gilt als Schlüsseltechnologie des 21. Jahrhunderts. Sie kann unzähligen konkreten Zwecken dienen: Übersetzungen, medizinischen Diagnosen oder personalisierten Produktempfehlungen. KI dürfte daher nach und nach in nahezu alle Bereiche der Gesellschaft Einzug halten, nicht nur in Form neuer Produkte und Dienstleistungen, sondern auch durch die Verbesserung bestehender Leistungen, die damit „smarter“ werden. Eine steigende Durchdringung der Gesellschaft mit KI-basierten Lösungen bedeutet allerdings auch, dass der globale Energieverbrauch steigt. Dieser entsteht nicht nur durch den Betrieb von Endgeräten wie Smartphones. Ein wesentlicher Teil des Energieverbrauchs von KI-Anwendungen findet extern durch Datentransfers und Rechenzentren statt. Zwar nimmt auch die Energieeffizienz der Rechenzentren stetig zu. Bleibt dieser Sektor aber weiterhin weitgehend unreguliert, kann ein Rebound-Effekt eintreten, d. h. dass die mit den Energieeinsparungen verbundenen Kosteneinsparungen zum Beispiel zu einer intensiveren Nutzung und dadurch wiederum zu einem weiteren Anstieg des absoluten Energieverbrauchs führen. Die EU-Kommission schätzt, dass der Energieverbrauch von Datenzentren in der EU von 77 Terrawattstunden (2,7 Prozent des Gesamtstromverbrauchs) im

Jahr 2018 auf 99 Terrawattstunden (3,2 Prozent des Gesamtstromverbrauchs) im Jahr 2030 steigen wird. Kurzum: Mehr und mehr Datentransfers und Datenverarbeitung bedeuten auch mehr Energieverbrauch.

Es ist somit eine wichtige Frage, wie sich KI-Lösungen möglichst energieeffizient entwickeln, steuern und nutzen lassen. Ein möglicher Ansatz ist, die Konsument*innen über den ökologischen Fußabdruck von KI aufzuklären. Insbesondere Zertifikate und Labels, wie sie die Europäische Union für „Green IT“ vorsieht, könnten den Konsument*innen helfen, umweltfreundlichere Optionen zu erkennen und dann auch auszuwählen. Denkbar ist auch, für die Entwicklung und den Verkauf von KI-Lösungen und -Hardware sogenannte „Carbon Impact Assessments“ vorzuschreiben. Wenn der ökologische Fußabdruck von KI-Anwendungen auf diesem Wege beziffert und öffentlich gemacht wird, könnte er auch das Konsumverhalten beeinflussen.

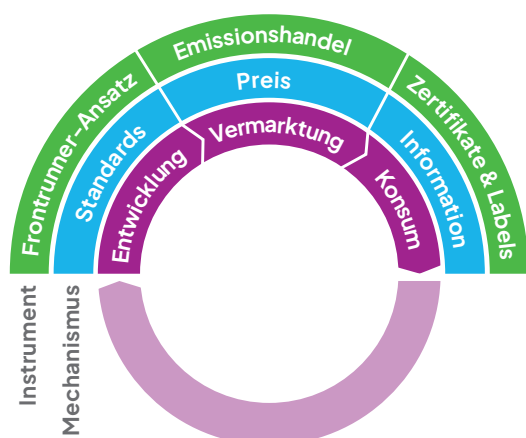
Allerdings ist zweifelhaft, wie wirksam Informationen alleine sind. Nicht nur entsteht damit ein weiterer Bereich, in dem sich Verbraucher*innen informieren müssen, um aufgeklärt zu entscheiden. Auch ist aus anderen Bereichen bekannt, dass selbst die Kombination aus ausreichendem Wissen und einer umweltfreundlichen Einstellung nicht unbedingt zu einem veränderten Verhalten führt. Außerdem dürften sich energieeffizientere Anwendungen kaum durchsetzen, wenn sie zugleich teurer oder weniger leistungsfähig sind, wie

DR. MARKUS B. SIEWERT



... ist Geschäftsführer des TUM Think Tank an der Hochschule für Politik München. Als Politikwissenschaftler interessiert er sich für sozio-politische Herausforderungen der digitalen Transformation, für Künstliche Intelligenz und Nachhaltigkeit sowie für die gesellschaftliche Governance dieser Themen.

die Ergebnisse der von uns durchgeführten Studie Consumers are willing to pay a price for explainable, but not for green AI. Evidence from a choice-based conjoint analysis nahelegen. Vor allem bei kostenlosen Apps dürften Aufpreise für höhere Energieeffizienzstandards wenig Akzeptanz finden. Umweltfreundliche KI dürfte sich daher kaum auf der Basis von Informationen und Labels alleine am Markt etablieren. Politisches Handeln und staatliche Regulierung müssen vielmehr zusätzlich an Stellen ansetzen, die den Konsumententscheidungen vorgelagert sind – wie es die nachfolgende Grafik verdeutlicht.



Finanzielle Instrumente können bereits bei der Vermarktung von Produkten ansetzen und Angebot und Nachfrage über den Preis steuern. Eine relativ einfache und schnell umsetzbare Lösung wäre etwa, sowohl die Produzenten von KI-Anwendungen und -Produkten als auch Rechenzentren in das schon existierende CO₂-Emissionshandelssystem einzubeziehen. Dies hätte stärkere Anreize zur Folge, möglichst energiespar-

DR. PASCAL D. KÖNIG



... ist Politikwissenschaftler an der TU Kaiserslautern und zurzeit Gastforscher am *Minda de Gunzburg Center for European Studies* der Harvard University. Sein wissenschaftliches Interesse gilt den Folgen der Datafizierung von Gesellschaften, der Steuerung durch Algorithmen (insbesondere ethischen und regulatorische Aspekten) sowie der Frage, wie politische Akteur*innen mit den Herausforderungen des digitalen Wandels umgehen.

same Anwendungen und Infrastrukturen für Information zu entwickeln und zu nutzen. Eine Verteuerung der extern verbrauchten Energie und der dabei entstehenden Emissionen von KI-Anwendungen würde Unternehmen dazu antreiben, bereits bei der Entwicklung ökologische Nachhaltigkeit zu berücksichtigen („Green AI“), statt nur auf die Leistungsfähigkeit zu schauen („Red AI“). Solche Maßnahmen können einen wichtigen Unterschied ausmachen. Ein OECD-Bericht zeigt, dass verschiedene technische Entscheidungen – etwa die Wahl der verwendeten Modelle, der Hardware und der Rechenzentren, inklusive ihrer Standorte – zu enormen Energieeinsparungen führen.

Schließlich kann Regulierung aber auch direkt bei der Entwicklung von KI-Anwendungen ansetzen. Dies könnte beispielsweise durch den Top-Runner-Ansatz geschehen, wie er in Japan eingesetzt wird: Alle Anbieter müssen jeweils innerhalb einer vorgegebenen Zeitspanne – etwa innerhalb eines Jahres – den vom führenden Anbieter erreichten höchsten Energieeffizienz-Standard einholen. Andernfalls riskieren sie – etwa infolge eines staatlichen Verbots – dass sie ihr Produkt nicht mehr am Markt anbieten dürfen. Solch eine harte Regulierung setzt neue Standards, indem sie sie dynamisch an den Marktwettbewerb und den jeweiligen Stand der Technik knüpft.

Der Einsatz all dieser Instrumente baut im Idealfall aufeinander auf und wird von einer kontinuierlichen Datenerhebung und Evaluation begleitet. Hierbei sollte ein handlungsfähiger

Staat nicht alleine den Berichten und Informationen von Unternehmen vertrauen müssen. Er muss auch Kapazitäten aufbauen, um Unternehmen effektiv kontrollieren sowie als letztverantwortliche Instanz im Zweifel regulatorisch eingreifen zu können.

PROF. DR. STEFAN WURSTER



... ist Professor für Policy Analysis an der TU München. Er befasst sich in seiner Forschung mit Policy-Vergleichen von Politikfeldern mit starkem Nachhaltigkeitsbezug, wie der Bildungs-, Forschungs-, Innovations-, Umwelt- und Energiepolitik. Seine Forschungsschwerpunkte sind unterschiedliche Instrumente politischer Steuerung sowie der Vergleich von Demokratie und Autokratie.

KI aus dem Baukasten

Das Software-Unternehmen elevait entwickelt KI-Produkte für die Automatisierung von Geschäftsprozessen. Durch den modularen Aufbau der Systeme werden Trainingszeiten reduziert und Ressourcen geschont. Außerdem wird dadurch mittelständischen Unternehmen der Einsatz von KI ermöglicht.

Wie kommt elevait zu dem Anspruch, nachhaltige KI zu entwickeln?

In den Jahren seit unserer Gründung haben wir festgestellt, wie die Umsetzung von KI nicht gut funktioniert: Die Entwicklung darf sich nicht an einzelnen Kundenprojekten mit speziellen Problemen orientieren. Derartige Problemlösungen können nicht auf weitere Projekte übertragen werden, so dass Zeit und Geld nicht gerade nachhaltig investiert worden sind. Wir liefern Firmen, die mit ihren Daten zu bestimmten Ergebnissen gelangen möchten, also nun passende KI-Bausteine, die die Kundendaten verarbeiten, ohne jedes Mal von Grund auf neu entwickelt werden zu müssen. Dieses Prinzip hat natürlich seine Grenzen, aber es hat sich als sehr praktikabel erwiesen und ist nicht nur ökonomisch, sondern auch ökologisch nachhaltig. Wenn für jedes KI-Projekt ein neues Modell trainiert werden muss, ist der zur Berechnung notwendige Server- und Energieaufwand enorm. Je weniger Training wir brauchen, desto weniger Ressourcen werden verbraucht, um die Modelle zu erstellen und somit die Lösung für den Kunden bereitzustellen.

Wie können wir uns dieses Bausteinprinzip vorstellen? Wie ist so eine KI aufgebaut, dass sie für unterschiedliche klein- und mittelständische Unternehmen anpassbar und einsetzbar ist?

Stellen Sie sich ein Projekt vor, bei dem Dokumente extrahiert werden müssen. Wenn wir dafür eine KI-Lösung en bloc entwickeln, müssten wir zum Beispiel ein Modell bauen, das die Art



der Dokumente erkennt, bestimmte Bereiche herauszieht und den entsprechenden Text liest. Bei unserem Ansatz werden die Aufgaben in diesem Prozess als modulare Bausteine behandelt. Für jeden Arbeitsschritt gibt es ein definiertes Modell, zum Beispiel für die Handschrifterkennung. Dieser Baustein hängt nicht bzw. nur wenig vom spezifischen Projektprofil der jeweiligen Kunden ab. Derartig trainierte Modelle können wir in allen Anwendungsfällen einsetzen, in denen sie gebraucht werden. Wir haben nach diesem Bausteinprinzip Modelle für einzelne kleine Aufgaben trainiert und orchestrieren diese Einzelmodule über einen Workflow, der sich aus den Projektanforderungen ergibt.

Bei diesem modularen Verfahren sind die Unternehmen nicht mehr in die KI-Entwicklung einbezogen. Kann es nicht auch problematisch sein, wenn die Unternehmen gar nicht wissen, was da wie passiert?

Wir haben gemerkt, dass wir eine KI-Software einzelnen Kunden nicht in der Erwartung anbieten können, dass wir sie verkaufen und damit unseren Beitrag geleistet haben. Die Anwendung von KI folgt bestimmten Workflows, die in den Unternehmen etabliert sind, das muss sich über einen gewissen Zeitraum hinweg entwickeln, auch in dem Sinne, dass die Modellqualität mit den trainierten Elementen immer besser wird. Bei den Kunden muss Vertrauen wachsen. Wir arbeiten mit transparenten Benchmarks und sagen ihnen zum Beispiel: Wir haben ein repräsentatives Datenset generiert und mit



den Modellen analysieren lassen, 80 Prozent der Dokumente werden erkannt und die Daten mit 95-prozentiger Genauigkeit eingelesen. Dann können uns die Kunden ein Feedback geben, ob diese Werte für sie okay sind. Wenn die Ergebnisse stimmen, spielt es letztlich keine Rolle, ob die KI-Lösung für sie eine Blackbox ist. Allerdings müssen wir erst einmal zu dem Status gelangen, dass sie dieser Blackbox vertrauen.

Mit welchen Anwendungsfällen kommen Unternehmen auf Sie zu?

Unser Kerngeschäft sind KI-Produkte für die Automatisierung von Geschäftsprozessen wie die automatische Verarbeitung von Dokumenten. Da geht es beispielsweise um eine automatisierte Erfassung und Weiterverarbeitung von Formularen zur Bestellung oder Inventur, Anamnesebögen im Krankenhaus, Rechnungen, Lieferscheinen, Service-Tickets, Bauplänen oder reinen Texten im weitesten Sinn.

Wie wirkt sich die Automatisierung im Unternehmen aus?

Einer unserer Kunden ist zum Beispiel ein Hersteller von orthopädischen Produkten. Individualbestellungen machen einen sehr großen Anteil seines Geschäfts aus. Die Patient*innen lassen also ihre Beine oder Arme in einem Sanitätshaus individuell abmessen, diese Maße wer-

den dem Anbieter übermittelt und dieser muss innerhalb von kürzester Zeit das Produkt herstellen und versenden. Dieser Vorgang bringt eine sehr enge Taktung und hohe Individualisierung mit sich. Es muss zudem alles genau passen. Bei der Bestellung müssen bis zu 40 Maße plus 40 bis 80 Konfigurationsoptionen berücksichtigt werden. Bislang wurde der Bestellvorgang noch zu 80 Prozent über ein Formular abgewickelt, das per E-Mail oder Fax verschickt wird. Unser Kunde erhält pro Tag mehrere Tausend Bestellformulare, die oft handschriftlich ausgefüllt sind. Früher tippeten täglich um die 40 Beschäftigte in zwei Schichten die Formulare ab. Es wurde für unseren Kunden immer schwieriger, dafür Mitarbeiter*innen zu finden, was seinen Wachstumsplänen entgegenlief. Das Unternehmen wollte mehr Bestellungen am Tag abarbeiten und das war mit dem bisherigen Prozess so nicht möglich. Unsere Aufgabe bestand darin, die Bestellungen teil- oder vollautomatisiert zu erkennen und ins System einzupflegen.



Dimension:
Ökologische Nachhaltigkeit



Kriterium:
Energieverbrauch



Indikator:
Berücksichtigung und Optimierung der Energieeffizienz

Bei der KI-Entwicklung und dem KI-Einsatz kann viel Energie eingespart werden, beispielsweise durch das Komprimieren von Modellen, durch ein effizientes Trainieren der Modelle, durch datenminimalistische Ansätze, durch die Nutzung vortrainierter Modelle, durch weniger komplexe Modelle oder durch eine effiziente Software- und Hardwareinfrastruktur. Es fehlt bisher an einem Bewusstsein für dieses Thema und an Expertise zu entsprechenden Methoden. Organisationen, die KI entwickeln oder einsetzen, sollten jedoch die Energieeffizienz eines KI-Systems zu einem zentralen Kriterium in ihren Entscheidungs- und Auswahlprozessen machen, um ökologisch nachhaltig zu handeln.

PG:	PG:
(bei Mose vordere Schrittänge)	
PF:	PF:
PE:	PE:
PD: 38	PD: 38
PC: 29.5	PC: 29.5
PB: 21.5	PB: 21.5
PA: 11	PA: 11

vorgestrichen in

☐ Anatomisch abg. Länge „IE“ bis K

☐ Futterstiel bei Länge „IE“ bis K

☐ Nähte nach out.

Ausstattung Leib

☐ Schrägvorsatz

☐ Schwangerschaft

☐ Links ☐ Rechts

☐ Gestrickabschl.

☐ Taillengürt

☐ Hand und Taillen

☐ Leibteil KKL 1

Im Sanitäthaus werden bei den Patient*innen die Maße von Beinen und Armen genommen und handschriftlich eingetragen. Diese Daten werden dann gelegentlich sogar noch per Fax übermittelt.

Wenn Automatisierungsprozesse eingeführt werden, steht immer die Befürchtung im Raum, dass damit Arbeitsplatzverluste einhergehen. Wie sind Sie damit umgegangen?

Wir müssen bedenken, dass unserer Gesellschaft ein demografischer Wandel bevorsteht, bei dem sehr viele Menschen in den nächsten zehn Jahren aus der Arbeitswelt ausscheiden werden. Gleichzeitig wird die Zahl der zu verarbeitenden Dokumente, Informationen und Daten stetig ansteigen. KI ist eine Lösung, um dieser immer größer werdenden Lücke zwischen zur Verfügung stehender Arbeitskraft und dem Bedarf an Datenverarbeitung zu begegnen. Es geht nicht darum, den Menschen die Jobs wegzunehmen, sondern darum, dass wir uns diesem Problem stellen. Außerdem soll KI die Arbeit erleichtern. Die Mitarbeiter*innen können dort eingesetzt werden, wo eine KI nicht unterstützen kann.

Dennoch verändert der Einsatz von KI die Arbeit der Mitarbeiter*innen. Wenn das Abtippen von Daten jetzt schneller geht, was machen die dafür angestellten Mitarbeiter*innen stattdessen?

Wenn man versucht, einen Prozess zu automatisieren, dann reicht es nicht, nur KI einzusetzen. Ein reibungsloser Prozessablauf hängt vom Spezial- und Erfahrungs-

wissen der Mitarbeiter*innen ab. Zum Beispiel kennen die Mitarbeiter*innen den Kunden X seit langer Zeit. Wenn dieser Kunde etwas bestellt, dann rechnen sie seit Jahren pauschal immer 10 Prozent auf die angegebenen Maße dazu. Sie wissen nämlich aus Erfahrung, dass die Bestellungen bei diesem Kunden immer zu knapp ausfallen und er sie ohne Anpassung später ohnehin ändern wird. Wenn ein Automatisierungsprozess erfolgreich durchgeführt werden soll, müssen auch solche Nu-

ancen ins Zielsystem aufgenommen werden. Eine reine Datenextraktion reicht da nicht aus. Daher stellen wir unseren Kunden mit dem KI-System auch einen Editor zur Verfügung, mit dem sie selber Regeln definieren können. Sie können dann im System festlegen: Wenn eine spezielle Kundennummer und folgendes Attribut erkannt wurde, dann lösche bitte diesen Wert oder rechne auf alle Maße 10 Prozent hinzu. Wir haben das Tool so gebaut, dass unsere Kund*innen das selbst übernehmen und überlegen können, was dem System noch fehlt. So können sie eigenständig Stück für Stück das vorhandene implizite Fachwissen der einzelnen Mitarbeiter*innen ins System einpflegen, um so den Automatisierungsgrad stetig zu erhöhen.

Inwiefern müssen Mitarbeiter*innen zu diesem Zweck weitergebildet werden?

Wir haben mit den Angestellten aus dem Kreis der Kundenservice-Mitarbeiter*innen, also den Perso-



Dimension:
Ökonomische
Nachhaltigkeit



Kriterium:
Arbeitsbedingungen
und Arbeitsplätze



Indikator:
Evaluation und
Optimierung der
Arbeitsbedingungen

KI-Systeme werden immer häufiger im Arbeitsleben eingesetzt, was die Arbeitsbedingungen sowohl verbessern als auch verschlechtern kann. Viele befürchten, dass der Einsatz von KI zum Abbau von Arbeitsplätzen führt. Organisationen sollten die Konsequenzen für Arbeitnehmer*innen abschätzen, bevor sie ein KI-System einführen. Die Systeme können zum Beispiel dazu führen, dass die Arbeit monotoner wird, dass Arbeitnehmer*innen stärker überwacht werden oder dass die Qualifikationen der Arbeitnehmer*innen für bestimmte Tätigkeiten ihren Wert verlieren. Hier sollte ein Interessensausgleich stattfinden, zum Beispiel durch Weiterbildungsangebote.



KI-Modelle sollten so entwickelt werden, dass sie weniger Rechenzeit und damit weniger Hardware benötigen. Das hat letztlich einen geringeren ökologischen Fußabdruck zur Folge.

nen, die bisher die Daten abgetippt hatten, einen Workshop durchgeführt, um zu erklären, wie das Regelwerk adaptiert werden kann. Sie pflegen heute das System – unter Aufsicht der Kundenservice-Leitung und eines IT-Verantwortlichen. Aber die Regeln kommen von den Mitarbeiter*innen. Der Fokus ihrer Arbeit verschiebt sich durch den Einsatz der KI zudem. Sie sind jetzt mehr mit dem Kontrollieren und Bearbeiten von Grenzfällen beschäftigt, die nicht durch das System, sondern nur persönlich durch die Mitarbeiter*innen abgearbeitet werden können. Dafür haben sie mehr Zeit, sich diesen Fällen individuell zu widmen.

Sie arbeiten mit dem nachhaltigen Rechenzentrum Cloud&Heat zusammen. Wie wichtig ist es, bereits in der KI-Entwicklung ressourcenschonend zu denken?

Da gibt es eine starke Wechselwirkung zwischen ökonomischen und ökologischen Faktoren. Für uns ist eine zentrale Frage, wie wir eine möglichst geringe Rechenzeit erreichen.

Schaffe ich es allein durch die Wahl der Architektur meines KI-Modells und der dahinterliegenden Software, die Rechenzeit zu reduzieren? Das ist ein ganz einfacher und wichtiger Faktor, um ökologisch nachhaltig zu sein – weil wir so weniger Strom verbrauchen. Das ist mit dem ökonomischen Vorteil verbunden, dass es weniger kostet. Wir dürfen auch die Hardware nicht vergessen, die für die Berechnungen notwendig ist. Allein der Bau der Hardware hinterlässt einen großen CO₂-Fußabdruck. Wir müssen unsere Modelle auf GPUs trainieren. Wenn ich 100 GPUs brauche, ist der Fußabdruck entsprechend groß. Wenn ich aber meine Modell-Architektur so intelligent wähle, dass ich ein ähnliches Ergebnis auf 10 GPUs berechnen kann, dann habe ich einen deutlich kleineren CO₂-Fußabdruck. Wir optimieren unsere Modelle aus intrinsischer Motivation, aber gleichzeitig gibt es auch viele monetäre Anreize.

**GREGOR
BLICHMANN**



... forschte nach seinem Studium in Computer Science als wissenschaftlicher Mitarbeiter an der Technischen Universität Dresden zu webbasierter Software- und Serviceentwicklung sowie semantischen Webtechnologien. Anschließend wechselte er als Software Engineer und -Architekt zum Vorgängerunternehmen von *elevait*, wo er die grundlegende Software-Architektur mit aufbaute. Mit dem Start von *elevait* und mit dem Wachstum der technischen Teams übernahm er zunehmend die Koordination zwischen den technischen Teams, bis er Chief Technology Officer (CTO) des Unternehmens wurde.

Wärme statt Überhitzung: Wie Rechenzentren zu nachhaltigen Heizkörpern werden

Das Dresdner Start-up *Cloud&Heat* hat sich zum Ziel gesetzt, Cloud-Infrastruktur nachhaltiger zu machen. Durch sein Rechenzentren-Kühlungssystem, das auf Direkt-Heißwasserkühlung setzt, kann das Unternehmen laut einer Modellrechnung bei einem Pilotprojekt in Frankfurt jährlich bis zu ca. 710 Tonnen Kohlendioxid einsparen – im Vergleich zu traditionellen, luftgekühlten Rechenzentren. Ronny Reinhardt, Team Lead Business Development bei *Cloud&Heat*, erklärt, wie sich die Abwärme aus Rechenzentren zum Heizen nutzen und damit eine nachhaltige Cloud betreiben lässt.



Welche Rolle spielen Rechenzentren für den Ressourcenverbrauch von KI?

KI verbraucht Rechenleistung und somit Energie. Die KI-Modelle laufen in der Regel auf GPU-Hardware, also Grafikkarten, die eine sehr hohe Leistungsaufnahme haben. Wenn wir auf Servern rechnen, dann wird energetisch gesehen eigentlich nur elektrische Energie in Wärme umgewandelt. Das hat wiederum zur Folge, dass zusätzlicher Energiebedarf besteht, um die Server zu kühlen. In Frankfurt allein könnte zum Beispiel mit der Wärme, die in den Rechenzentren erzeugt wird, theoretisch die gesamte Stadt geheizt werden. Rechenzentren als Basis für Cloud- und KI-Lösungen bilden eine stark wachsende Industrie, über die wir uns jetzt Gedanken machen müssen, um auch in Zukunft gut aufgestellt zu sein.



Server-Schränke von Cloud&Heat: Das Unternehmen versucht, den steigenden Ressourcenverbrauch der Cloud- und Rechenzentrumsbranche zu bremsen und ihren CO₂-Fußabdruck zu reduzieren.

Reicht es aus, wenn Rechenzentren auf Ökostrom oder erneuerbare Energien setzen und gegebenenfalls ihren CO₂-Austausch kompensieren?

Aus unserer Sicht ist das nicht ausreichend. Natürlich ist erneuerbare Energie besser als ein konventioneller Strommix. Aber grundsätzlich sollte unser Ziel sein, so wenig Energie wie möglich zu verbrauchen. Das fängt bei der Software-Entwicklung auf Applikationsebene an und zieht sich durch die verschiedenen Software-Ebenen hindurch bis in die Rechenzentren. Hier ist dann die Frage: Wie lässt sich ein Rechenzentrum möglichst effizient betreiben? Die Kühlung ist ein großer Faktor. Vor zehn Jahren war es nötig, die gleiche Menge an Energie, die der Server braucht, in die Kühlung zu stecken. Heute benötigen wir dafür nur noch um die 20 bis 30 Prozent. Dennoch müssen wir hier kontinuierlich besser werden. Das zweite große Thema ist die Abwärmenutzung. Wie gesagt ist es energetisch gesehen so, dass wir in Rechenzentren im Prinzip elektrische Energie reinstecken und Wärme rausbekommen – auch wenn wir natürlich noch nebenbei etwas berechnen. Diese Abwärme muss in Zukunft sinnvoll genutzt werden. Dafür gibt es viele

verschiedene Ansätze wie die Einspeisung in die Wärmeversorgung von Gebäuden oder die Anbindung an Fernwärmenetze.

Mit welchen Maßnahmen optimieren Sie die Energieeffizienz in den Rechenzentren?

Unser Fokus ist die sogenannte Direkt-Heißwasserkühlung. Wir leiten die Wärme, die die Bauteile eines Servers produzieren, der Prozessor oder die Grafikkarten, direkt ab, um sie für andere Zwecke zu nutzen – zum Beispiel für eine Gebäudeheizung, wie wir es bei unserem Pilotprojekt im ehemaligen Rechenzentrum der Europäischen Zentralbank in einem Hochhaus in Frankfurt tun. Das ist eine technische Herausforderung, weil wir die Wärme auf einem hohen

Temperaturniveau abgreifen müssen, sonst lässt sich damit wenig anfangen. Die Direkt-Heißwasserkühlung ist energieeffizienter als eine Luftkühlung, bei der Luft durch klassische Kälteanlagen gekühlt werden muss. Das verbraucht deutlich mehr Energie, als eine Pumpe zu betreiben, die Wasser durch einen Kreislauf führt und über die Server und die heißen Bauteile fließen lässt. Hier verbrauchen wir also schon weniger Energie. Wir nutzen zudem die Abwärme, indem wir sie in das Gebäude einspeisen – zum Beispiel in eine Heizungsanlage. Auf der Grundlage dieser Technologie setzen wir eine Open-Source-basierte Cloud-Lösung auf, die KI-Kunden nutzen können, um das Training und die Inferenz der KI-Modelle möglichst nachhaltig ausführen zu können.

Sie haben gemeinsam mit Vattenfall auch ein Pilotprojekt in Schweden gestartet, bei dem es darum geht, Abwärme für die Energieversorgung zu nutzen. Wie genau läuft das ab? Was ist das Ziel des Projekts?

Gemeinsam mit Vattenfall bauen wir eine Cloud für KI-Unternehmen oder andere Unternehmen auf, die eine hohe Rechenleistung brauchen und denen die Themen Nachhaltigkeit und Energieeffizienz wichtig sind. Wir haben auf dem Gelände eines Biomassekraftwerks unsere Rechenzentrumscontainer mit der Direkt-Heißwasserkühlung aufgestellt. Wir befinden uns also direkt an der Quelle von erneuerbaren Energien. Unsere Anlagen dort sind außerdem so effizient, dass wir für den Betrieb des Rechenzentrums nur ca. sieben Prozent zusätzliche Energie im Vergleich mit der unmittelbaren Serverleistung benötigen. Dort sind wir dann an das Fernwär-

menetz angebunden, in das wir die bei der Kühlung erzeugte Wärme einspeisen. Sie wird dann an umliegende Haushalte in der Nähe von Stockholm weitergeleitet.

Sie haben eingangs erwähnt, dass theoretisch der Wärmebedarf der Stadt Frankfurt mit der in Rechenzentren erzeugten Wärme abgedeckt werden könnte. Warum wird es dann nicht getan? Woran scheitert es konkret?

Ein klassisches Rechenzentrum wird auch heutzutage noch per Luftkühlung betrieben und das erschwert die Abwärmenutzung, weil das Temperaturniveau der Luft nicht besonders hoch und der Wärmetransport über größere Strecken schwierig ist. Um den Status quo mit der Luftkühlung zu ändern, müssten sich sehr viele Stakeholder koordiniert gemeinsam vorwärtsbewegen. Die großen Rechenzentren sind meistens sogenannte Colocation-Rechenzentren, in die sich Kunden einmieten. Die Kunden müssten also selbst wassergekühlte Hardware einbringen und die Rechenzentren die Struktur dafür schaffen. Das ist der Grund, warum *Cloud&Heat* in seiner Ausprägung existiert. Wir haben unser eigenes Rechenzentrum und wir können den

Kunden unsere eigene Cloud anbieten. Das größte Hindernis besteht darin, dass sehr viele Stakeholder auf dem Markt in dem Bereich eher konservativ denken. Um umzurüsten, müssen Rechenzentren erst einmal investieren und die Investitionskosten über Betriebskosteneinsparungen oder eine Vergütung für die Abwärme wieder reinholen.

Gibt es weitere Hindernisse, die verhindern, dass immer mehr Rechenzentren auf die Wasserkühlung oder andere innovative Kühlungssysteme setzen, die energiesparend sind?

Zunächst ist es auch eine Frage der Gewohnheit und des Angebots: Es gibt noch immer viel mehr luftgekühlte als wassergekühlte Hardware. Zwar steigt die Zahl der Anbieter und Modelle, aber die Angebotsbreite hat noch nicht das gleiche Niveau. Dann heißt es, die Hersteller müssten sich bewegen. Die Hersteller wiederum sagen, dass die Nachfrage nach wassergekühlten Systemen noch nicht ausreicht, um ihr Angebot umzustellen. Und so geht es teilweise nur in kleinen Schritten voran. Deshalb unterstützen wir an dieser Stelle und haben beispielsweise mit der Thomas-Krenn.AG einen wassergekühlten Server entwickelt, der als eines der ersten Systeme den Blauen Engel für Server erhalten hat.

Subventionen für das Bereitstellen von Abwärme

Um den Wandel zu einer nachhaltigeren Infrastruktur von Rechenzentren zu unterstützen, testen europäische Länder unterschiedliche Fördermodelle. Die Niederländische Unternehmensagentur (Rijksdienst voor Ondernemend Nederland) subventioniert beispielsweise im Auftrag des Niederländischen Ministeriums für Wirtschaft und Klimapolitik die Einführung von Technologien, die Treibhausgasemissionen reduzieren können. Wenn Unternehmen oder Organisationen Abwärme bereitstellen, können sie eine Subventionierung von 0,033 € bis 0,044 € pro Kilowattstunde erhalten. Solche Subventionierungen können wichtige Anreize für Rechenzentren setzen, Abwärme stärker zu nutzen.



Quelle: <https://english.rvo.nl/sites/default/files/2020/11/Brochure%20SDE%20plus%20plus%202020.pdf>

CLOUD&HEAT

EMISSIONSREDUZIERUNG / pro Jahr



8 KW

IT-Leistung des Rechenzentrums



äquivalent zu



900 x
Bäume²



≈ 1 ha

emittiert von



6 x
Autos³



*Quelle: 1) Cloud&Heat-Kühlungstechnologie im Vergleich zu herkömmlichen luftgekühlten Rechenzentren 2) <https://www.co2online.de/service/klima-orakel/beitrag/wie-viele-baeume-braucht-es-um-eine-tonne-co2-zu-binden-10658/> 3) VW Passat BlueMotion oder BMW 320d; Standardverbrauch von 1,92 CO₂t / Auto bei 15000km Jahreslaufleistung

Das Unternehmen hat berechnet, dass ein Rechenzentrum mit einer Gesamt-IT-Leistung von 8 kW ca. 11 Tonnen CO₂ pro Jahr einsparen kann, wenn es Wasser- statt Luftkühlung einsetzt. Dies entspricht den Emissionen, die sechs Autos pro Jahr ausstoßen. Um diese Emissionen zu kompensieren, wäre ein Hektar Wald (mit 10.000 m²) oder 900 Laubbäume erforderlich. Source: <https://www.cloudandheat.com/wp-content/uploads/2020/02/2020-CloudHeat-Whitepaper-Cost-saving-Potential.pdf>

Es wäre gesellschaftlich durchaus sinnvoll, einen Anreiz dafür zu schaffen, das Auskoppeln von CO₂-freier Wärme mit einer Prämie zu honorieren. Dann würden sich die notwendigen Technologien schneller entwickeln.

Der dritte Punkt ist die Transparenz Nutzer*innen gegenüber. Das betrifft einmal die Cloud-Nutzung von Unternehmen. Ihnen müsste transparent vermittelt werden, dass bestimmte Berechnungsschritte eine gewisse Menge an CO₂ verursachen. Aber auch den Endnutzer*innen fehlt noch jegliches Bewusstsein dafür, dass die Verwendung bestimmter Apps oder Dienste einen CO₂-Ausstoß zur Folge hat. In anderen Bereichen hat sich solch ein Bewusstsein für den CO₂-Fußabdruck inzwischen entwickelt.

So ein gesellschaftlich wichtiges Thema bräuchte zudem noch mehr politische Unterstützung. Es gibt Modelle aus anderen Ländern, in denen beispielsweise die aus solchen Prozessen abgeführte Abwärme vergütet wird. Im Prinzip haben wir es mit CO₂-freier Wärme zu tun, da das CO₂ ohnehin schon im Rechenschritt „verbraucht“ wurde. Wir können diese Wärme also entweder vernichten oder sinnvoll einsetzen. Es wäre gesellschaftlich durchaus sinnvoll, einen Anreiz dafür zu schaffen, das Auskoppeln von CO₂-freier Wärme mit einer Prämie zu honorieren. Dann würden sich die notwendigen Technologien schneller entwickeln.

Daten über den CO₂-Fußabdruck und den Energieverbrauch von KI-Systemen liegen in einem Unternehmen oft gar nicht vor. Es wird argumentiert, dass der Aufwand für die Erfassung zu hoch sei. Wie schwierig ist es, in dieser Hinsicht Transparenz herzustellen?

Es ist tatsächlich nicht einfach. Im Cloud-Kontext teilen sich verschiedene Kund*innen einen Server, da ist es nicht ohne weiteres möglich zu sagen, wie hoch der Verbrauch der Kund*in X an dem Server war oder wie viel Energie bei diesem oder jenem Rechenvorgang verbraucht wurde. Aber es gibt auch dafür Lösungen. Wir bringen genau solche Fragen gerade in ein großes europäisches

Forschungsprojekt (IPCEI-CIS) ein. Wichtig ist dabei, möglichst schnell zu einer Lösung zu kommen, selbst wenn nicht sofort eine hohe Genauigkeit erreicht werden kann. Auch die CO₂-Fußabdrücke, die auf Lebensmitteln stehen, sind nicht auf die Nachkommastelle genau. Darum geht es meist gar nicht. Wir müssen vielmehr ein fundiertes Gefühl dafür haben, dass bestimmte Lebensmittel zum Beispiel zu großen CO₂-Emissionen führen und andere nicht. Genauso müssen wir auch bei Cloud-Lösungen einen Schritt vorankommen, um mehr Transparenz zu schaffen und perspektivisch Rechenjobs dahin zu bringen, den CO₂-Ausstoß zu reduzieren.

Ist es auch ökonomisch gesehen sinnvoll, auf ökologisch nachhaltige Lösungen zu setzen?

Ich denke, das wird immer mehr konvergieren. Wir sehen gerade, dass es mit steigenden Strompreisen immer attraktiver wird, die Energieeffizienz zu steigern. In manchen Ländern sind die Stromkosten noch so niedrig, dass niemand sich darüber große Gedanken machen muss. Dort gehen Ökologie und Ökonomie auseinander. Natürlich besteht die Lösung nicht darin, den Strompreis als Innovationsanreiz für energieeffiziente Technologien zu erhöhen. Dann würden die Rechenzentren in andere Länder abwandern. Ökologie und Ökonomie sind aber längst keine Widersprüche mehr.

Im Koalitionsvertrag spricht die Bundesregierung davon, dass neue Rechenzentren ab 2027 klimaneutral betrieben werden sollen. Was halten Sie von dieser Ansage?

Es ist gut, dass das Problem auf der politischen Ebene erkannt wurde, es ist aber auch extrem wichtig, schnell zu handeln, weil jetzt Rechenzentren im Megawatt-Bereich geplant und gebaut werden. Wenn wir also nicht jetzt gegensteuern, werden die Rechenzentren jahrzehntlang so arbeiten, wie



sie geplant wurden. In Amsterdam wird schon jetzt kein neues Rechenzentrum mehr ohne Abwärmekonzept gebaut. Es gibt also schon Bewegung. Aber der Staat tut sich weiterhin schwer damit, Kriterien vorzugeben, die auf Nachhaltigkeit abzielen.

Was halten Sie von Siegeln für Rechenzentren?

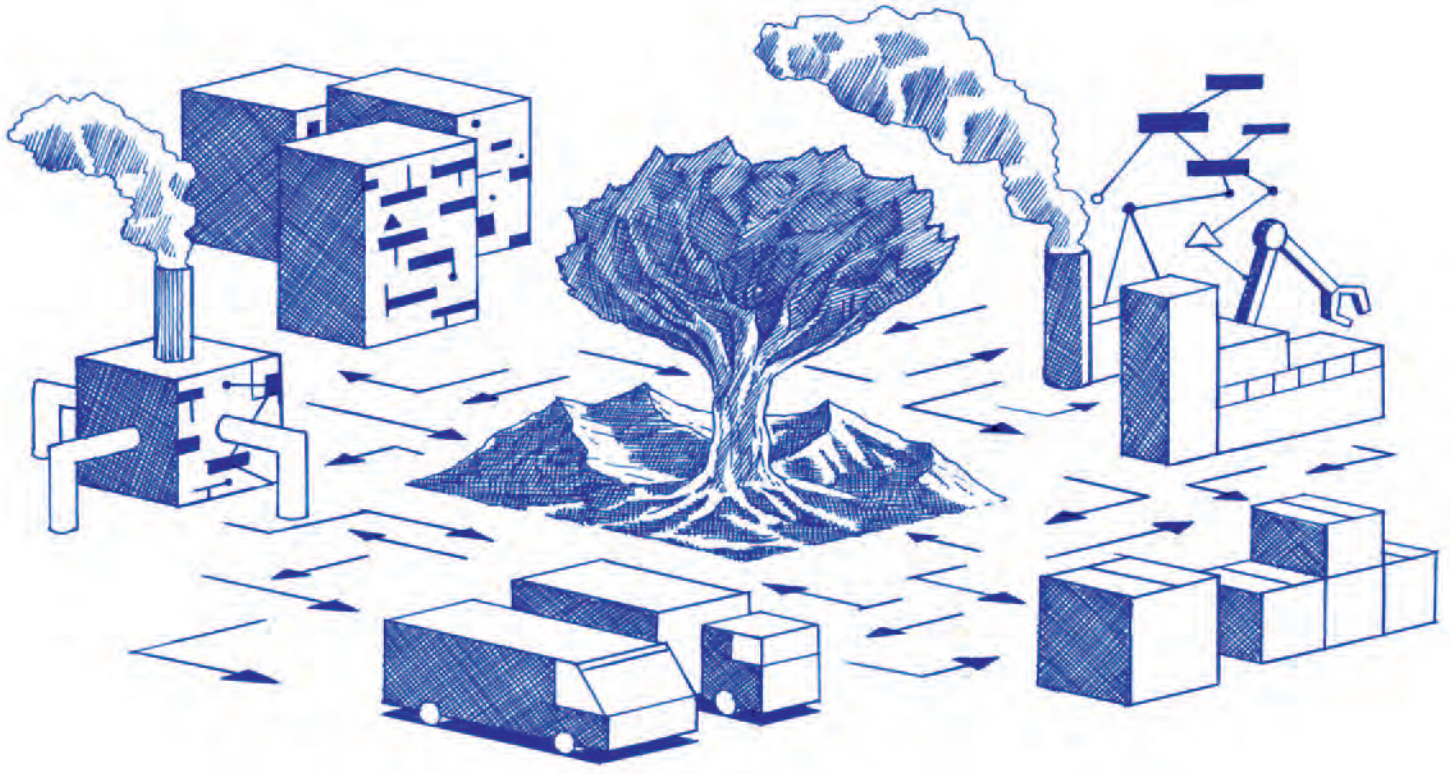
Das ist ein guter Ansatz, um ein gewisses Rahmenwerk zu etablieren, an das sich alle halten können. Aber ich glaube, es gibt noch nicht viele Rechenzentren, die die Voraussetzungen dafür erfüllen können. Wir orientieren uns allerdings an solchen ambitionierten Zielen. Darüber hinaus muss aber auch immer genügend Freiraum da sein, damit neue Lösungen, die noch nicht in einem Siegel abgebildet sind, ihre Wirkung entfalten können.

Wenn wir also nicht jetzt gegensteuern, werden die Rechenzentren jahrzehntlang so arbeiten, wie sie geplant wurden. In Amsterdam wird schon jetzt kein neues Rechenzentrum mehr ohne Abwärmekonzept gebaut.

DR. RONNY REINHARDT



... ist als Team Lead Business Development bei *Cloud&Heat Technologies* tätig. Er ist aktiv an den europäischen Cloud- und Dateninitiativen Gaia-X und IPCEI-CIS beteiligt. Bei Gaia-X war er Mitglied des technischen Komitees und für IPCEI-CIS koordiniert er das GREEN-CIS-Konsortium mit. Im KI-Verband gehört Ronny Reinhardt zur Arbeitsgruppe Klimawandel. Außerdem arbeitet er in der Initiative *Large European AI Models* (LEAM) mit. Zuvor forschte und lehrte er zum Technologie- und Innovationsmanagement an der FSU Jena, der University of Utah und der TU Dresden.



Es geht weiter mit sustain

In der nächsten Ausgabe sehen wir uns den Code genauer an.
Einige unserer Themen:

- Wie Code ressourcensparender werden kann
- Welche Methoden KI-Systeme fairer machen
- Wie KI dabei hilft, bei der Energieversorgung von Gebäuden Strom einzusparen

Wir präsentieren unsere Guidelines für nachhaltige KI und zeigen Schritt für Schritt, wie entlang des Lebenszyklus eines KI-Systems Nachhaltigkeit umgesetzt werden kann – von der ersten Idee bis zur Anwendung.

Impressum

Sustain Magazin #1

Juni 2022

Online verfügbar unter <https://algorithmwatch.org/de/sustain-magazin-2022>

Herausgeber

AW AlgorithmWatch gGmbH

Verantwortlicher (gemäß § 5 TMG und § 55 RStV): Matthias Spielkamp

Linienstr. 13

10178 Berlin

Deutschland

Inhaltliche Leitung

Anne Mollen

Redaktionelle Leitung

Waldemar Kesler

Inhaltliche Mitarbeit

Andreas Meyer (DAI Labor, TU Berlin)

Friederike Rohde (Institut für ökologische Wirtschaftsforschung)

Josephin Wagner (Institut für ökologische Wirtschaftsforschung)

Illustrationen

Juliette Baily

Kevin Lucbert

Presse- und Öffentlichkeitsarbeit

Marc Thümmeler

Gestaltung / Layout

Beate Autering (AD), Benito Felkel, Tiger Stangl

Entstanden im Rahmen des Projekts

„SustAI: Der Nachhaltigkeitsindex für Künstliche Intelligenz“



Projektpartner:



Diese Publikation steht unter der Lizenz Creative Commons Namensnennung 3.0 Deutschland

<https://creativecommons.org/licenses/by/3.0/de/legalcode>

Gefördert durch:



Bundesministerium
für Umwelt, Naturschutz, nukleare Sicherheit
und Verbraucherschutz

aufgrund eines Beschlusses
des Deutschen Bundestages



Projektpartner:

AlgorithmWatch

AlgorithmWatch ist eine gemeinnützige Forschungs- und Advocacy-Organisation mit dem Ziel, Systeme automatisierter Entscheidungsfindung (ADM) und deren Auswirkungen auf die Gesellschaft zu beobachten und zu analysieren. Eine Säule unserer Arbeit ist die Auseinandersetzung mit ADM-Systemen und Nachhaltigkeit. <https://algorithmwatch.org/de/>



Institut für ökologische Wirtschaftsforschung

Das IÖW ist ein führendes wissenschaftliches Institut auf dem Gebiet der praxisorientierten Nachhaltigkeitsforschung. Wir erarbeiten Strategien und Handlungsansätze für ein zukunftsfähiges Wirtschaften – für eine Ökonomie, die ein gutes Leben ermöglicht und die natürlichen Grundlagen erhält. Seit über 30 Jahren beschäftigen wir uns mit Zukunftsfragen und finden immer wieder neue, oft auch ungewöhnliche Antworten. <https://www.ioew.de/>



Distributed Artificial Intelligence Labor

Das DAI-Labor an der TU Berlin sieht sich als Mittler zwischen universitärer Forschung und industrieller Verwertung. Mit unserem interdisziplinären Team erzeugen wir Innovationen und überführen universitäre Forschung in Anwendungen des täglichen Lebens. Dies geschieht in enger Kooperation mit anderen wissenschaftlichen und industriellen Institutionen. <https://dai-labor.de/>

