

REPORT

AUTOMATING SOCIETY

2020



DEUTSCHE
AUSGABE



ALGORITHM
WATCH

| BertelsmannStiftung

IMPRESSUM

Automating Society Report 2020

Januar 2021

Online verfügbar unter: <https://automatingsociety.algorithmwatch.org>

Veröffentlicht von

AlgorithmWatch gGmbH
Linienstr. 13
10178 Berlin

Bertelsmann Stiftung
Carl-Bertelsmann-Str. 256
33311 Gütersloh

Herausgeber-innen

Fabio Chiusi
Sarah Fischer
Nicolas Kayser-Bril
Matthias Spielkamp

Projektmanagement

Fabio Chiusi

Koordination

Marc Thümmler

Comics

Samuel Daveti
Lorenzo Palloni
Alessio Ravazzani

Übersetzung

Katrin Harlaß

Gestaltung

Beate Autering

Redaktionsassistenz

Leonard Haas

Redaktionsschluss: 30. September 2020



Diese Publikation ist veröffentlicht unter der Lizenz Creative Commons Namensnennung 4.0 International
<https://creativecommons.org/licenses/by/4.0/legalcode.de>


Inhalt

Einleitung 4

Europäische Union 16

Deutschland 36

Team 53



Systeme zum automatisierten Entscheiden sind im Alltag angekommen. Wie gehen wir damit um?

Von Fabio Chiusi

Redaktionsschluss für diesen Bericht war der 30. September 2020.
Spätere Entwicklungen konnten nicht berücksichtigt werden.

Es war ein bewölkerter Augusttag in London, und die Schüler:innen waren wütend. Zu Hunderten strömten sie protestierend zum Parliament Square. Auf ihren Plakaten prangten Unterstützerslogans für ungewöhnliche Verbündete: ihre Lehrer:innen. Und daneben ein noch weit ungewöhnlicheres Angriffsziel: ein Algorithmus.

Wegen der Corona-Pandemie hatte das Vereinigte Königreich im März die Schulen geschlossen. Das Virus hielt auch den Sommer über ganz Europa in Atem, und den Schüler:innen war klar, dass ihre Abschlussprüfungen ausfallen und ihre Noten – irgendwie – anders ermittelt werden würden. Was sie sich allerdings nicht hatten vorstellen können, war, dass Tausende von ihnen schlechter abschneiden würden als erwartet.

Ihren Schildern und Sprechgesängen zufolge wussten die protestierenden Schüler:innen offenbar ganz genau, wer daran schuld war: das vom Office of Qualifications and Examinations Regulation (Ofqual) genutzte ADM-System (automated decision-making system). Die Behörde plante, die bestmöglichen datenbasierten Einschätzungen sowohl für General Certificates of Secondary Education (Mittlere Reife) als auch A-Level-Ergebnisse (Abiturzeugnisse) zu generieren, und zwar so, dass „die Verteilung der Noten einem Muster folgt, das dem vorangegangener Jahre ähnelt, so dass die Schüler:innen dieses Jahrgangs infolge der besonderen Umstände keine systematischen Nachteile erleiden“.

Die Regierung wollte überoptimistische¹ Benotungen, die aus einem ausschließlich menschlichen Urteil erwachsen, vermeiden: Ihrer eigenen Einschätzung zufolge wären die Abschlüsse, verglichen mit denen der Vergangenheit, zu gut ausgefallen. Doch dieser Versuch, „mit Schüler:innen, die in diesem Sommer ihre Prüfungen nicht ablegen konnten, so fair wie möglich umzugehen“, scheiterte spektakulär. Und an diesem grauen Protesttag im August riss der Strom der Protestierenden nicht ab, die weiterhin Sprechchöre bildeten und ihre Plakate hochhielten, um soziale Gerechtigkeit einzufordern. Manche waren außer sich, andere weinten.

„Stop stealing our future“, stand auf einem der Plakate, eine Anleihe bei den Protesten der Klimaaktivist:innen von Fridays for Future: Hört auf, uns unsere Zukunft wegzunehmen. Andere Slogans waren etwas genauer auf die Mängel

1 „Die Forschungsliteratur legt nahe, dass bei der Einschätzung von Noten, die Schüler:innen aller Wahrscheinlichkeit nach erreichen, Lehrer:innen zu Optimismus neigen (allerdings nicht in allen Fällen)“, schreibt Ofqual, vgl. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/909035/6656-2_-_Executive_summary.pdf

des ADM-Benotungssystems zugeschnitten: Sprüche wie „Benotet meine Arbeit, nicht meine Postleitzahl“ oder „Wir sind Schüler:innen, keine statistischen Größen“ prangerten die diskriminierenden Ergebnisse an, die es geliefert hatte².

Schließlich stimmte die Menge einen lauten Sprechchor an, der künftig bei Protesten zum festen Kanon gehören dürfte: „Fuck the algorithm“. Aus Angst, die Regierung automatisiere eben mal so – und auf undurchsichtige Weise – ihre Zukunft, ganz gleich, ob die Benotung ihre tatsächlichen Fähigkeiten und Anstrengungen widerspiegelte oder nicht, forderten die Schüler:innen das Recht ein, sich ihre Zukunftschancen nicht über Gebühr von einem fehlerhaften Code vermässeln zu lassen. Sie wollten mitreden, und was sie zu sagen hatten, sollte gehört werden.

Algorithmen sind weder „neutral“ noch „objektiv“, auch wenn wir das gerne glauben wollen. Sie reproduzieren die Annahmen und Überzeugungen jener, die sich dafür entscheiden, sie einzusetzen und zu programmieren. Aus diesem Grund sind für die Auswahl guter wie schlechter Algorithmen Menschen verantwortlich (oder sollten es sein) und nicht „Algorithmen“ oder ADM-Systeme. Maschinen mögen uns Angst machen, doch der Geist, der ihnen innewohnt, ist stets ein menschlicher. Und Menschen sind komplizierte Wesen, komplizierter noch als Algorithmen.

Ohnehin waren die protestierenden Schüler:innen nicht so naiv zu glauben, ihre Sorgen seien lediglich dem Versagen eines Algorithmus anzulasten. Ihre Sprechgesänge richteten sich nicht in einem Anfall technologischen Determinismus gegen „den Algorithmus“; ihre Motivation war vielmehr das dringende Bedürfnis, soziale Gerechtigkeit zu schützen und voranzutreiben. In dieser Hinsicht ähnelt ihr Protest viel eher dem der Ludditen. Ebenso wie den Anhänger:innen der Gewerkschaftsbewegung, die im 19. Jahrhundert mechanisierte Webstühle und Strickrahmen zerstörten, ist ihnen klar, dass ADM-Systeme etwas mit Macht zu tun haben und nicht irrtümlicherweise für eine angeblich objektive Technologie gehalten werden sollten. Also riefen sie „Gerechtigkeit für die Arbeiterklasse“, forderten den Rücktritt des Gesundheitsministers und brandmarkten ADM-Systeme als „Klassismus in Reinkultur“ und „eklatanten Klassismus“. Letztendlich hatten die Schüler:innen Erfolg. Sie erreichten eine Abschaffung des Systems, das ihren weiteren Bildungsweg und ihre Lebenschancen bedrohte: In einer spektakulären Kehrtwende verwarf die britische Regierung

2 Für weitere Details vgl. das Kapitel United Kingdom in der Gesamtausgabe des Automating Society Reports unter <https://automatingsociety.algorithmwatch.org/report2020/united-kingdom/>

das fehleranfällige ADM-System und wandte die von den Lehrer:innen prognostizierte Benotung an.

Doch hinter dieser Geschichte steckt mehr als nur die Tatsache, dass die Protestierenden am Ende ihr Ziel erreichten. Dieses Beispiel verdeutlicht, wie schlecht entwickelte, implementierte und überwachte ADM-Systeme, die menschliche Voreingenommenheit und Diskriminierung reproduzieren, dazu führen, dass das ihnen innewohnende Potenzial – etwa als Hebel zu dienen, Vergleichbarkeit und Fairness zu fördern – ungenutzt bleibt.

Stärker als viele Kämpfe der Vergangenheit macht dieser Protest bewusst, dass wir längst nicht mehr nur dabei sind, die Gesellschaft zu automatisieren. Wir haben sie bereits automatisiert – und endlich hat es jemand bemerkt.

/ Von der Automatisierung der Gesellschaft zur automatisierten Gesellschaft

Als wir die erste Auflage dieses Berichts publizierten, entschieden wir uns, ihn „Automating Society“ zu nennen, denn ADM-Systeme waren in Europa größtenteils neu, in Testphasen und noch unerforscht. Vor allem aber waren sie eher die Ausnahme als die Norm.

Das hat sich dramatisch verändert. Wie die zahlreichen Fälle, die von unserem Expertin:innen-Netzwerk in diesem Bericht zusammengetragenen wurden, deutlich zeigen, hat sich der Einsatz von ADM-Systemen in nur etwas mehr als einem Jahr um ein Vielfaches erhöht. ADM-Systeme wirken sich mittlerweile nicht nur auf alle möglichen menschlichen Aktivitäten aus, sondern auch, und das ist am bedeutsamsten, auf Dienstleistungen für Millionen EU-Bürger:innen – und deren Zugang zu ihren Rechten.

Während immer mehr ADM-Systeme eingesetzt werden, bleibt die Intransparenz, die sie umgibt, bestehen. Deshalb ist es dringender denn je, mehr Informationen ans Licht zu holen. Wir haben daher die Zahl der Länder, die wir untersucht haben, erhöht: Statt zwölf, wie in der ersten Fassung dieses Berichts, sind es nun 16; Estland, Griechenland, Portugal und die Schweiz sind hinzugekommen. Das ist zwar bei Weitem nicht erschöpfend, dennoch erlaubt es uns, ein umfassenderes Bild der derzeitigen Situation von ADM-Systemen in Europa zu zeichnen. Wir sind überzeugt, dass diese Arbeit unerlässlich ist. Denn diese Systeme können Auswirkungen nicht nur auf unser Alltagsleben haben. Sie werfen auch die Frage auf, wie sich Automatisierung auf

unsere Institutionen, unsere Normen und Werte, ja auf die Demokratie insgesamt auswirkt.

Dies gilt insbesondere während der COVID-19-Pandemie, in einer Zeit also, in der viele ADM-Systeme (zumeist übereilt) eingeführt wurden, die darauf ausgerichtet sind, mit Hilfe von datenbasierten Tools und Automatisierung zum Schutz der öffentlichen Gesundheit beizutragen. Wir hielten diese Entwicklung für derart bedeutend, dass wir uns entschieden haben, ihr im Rahmen des Projektes „Automating Society“ eine „Sonderausgabe“ zu widmen, die im September 2020 als „Vorabbericht“ veröffentlicht wurde³.

Selbst in Europa scheint es beim Einsatz von ADM-Systemen kein Halten mehr zu geben. Das wird an den in diesem Bericht aufgeführten Fällen deutlich, die zu den vielen, über die wir bereits in der [Vorgängerausgabe](#) berichtet haben, hinzugekommen sind – vom Wohlfahrtssektor über die Bildung und das Gesundheitswesen bis hin zur Justiz. Auf den folgenden Seiten geben wir zum ersten Mal ein Update zu den aktuellen Entwicklungen in diesen Fällen, und zwar auf dreierlei Weise. Erstens in journalistischen Stories; zweitens in Abschnitten, in denen verschiedene Beispiele katalogisiert werden; und drittens mit Hilfe von Comics. Unserer Auffassung nach sind diese ADM-Systeme in unser aller Leben bereits derart maßgebend – und werden es zunehmend sein – dass wir unbedingt versuchen müssen, alle Zielgruppen zu erreichen, und zwar auf eine neue Weise. Wir müssen erläutern, wie sie funktionieren und was sie *mit uns machen*. Denn letztlich wirken sich ADM-Systeme auf jeden und jede Einzelne von uns aus.

Oder sie könnten es zumindest. Wir haben zum Beispiel gesehen, wie in Estland ein neuer, automatisierter, proaktiver Service Familienzuschüsse verteilt. Eltern müssen nicht einmal mehr einen Antrag stellen: Von Geburt an sammelt der Staat alle Informationen zu jedem einzelnen Neugeborenen und dessen Eltern und führt sie in Datenbanken zusammen. Dies sorgt dafür, dass Eltern automatisch alle finanziellen Hilfen erhalten, die ihnen rechtlich zustehen.

In Finnland erfolgt mit Hilfe eines vom japanischen Großkonzern Fujitsu entwickelten Tools eine automatisierte Identifizierung individueller Risikofaktoren, die mit der gesellschaftlichen Stigmatisierung junger Erwachsener verbunden sind. In Frankreich ist es möglich, Daten aus

3 'Automated Decision-Making Systems in the COVID-19 Pandemic: A European Perspective', <https://algorithmwatch.org/en/project/automating-society-2020-covid19/>

sozialen Netzwerken zu sammeln und Algorithmen des Maschinellen Lernens mit ihnen zu füttern, um Steuerbetrug aufzudecken.

Italien experimentiert mit „voraussagender Rechtsprechung“. Hierbei kommt Automatisierung zum Einsatz, die Richter:innen aufzeigt, wie frühere Gerichtsurteile zum jeweiligen Streitgegenstand tendenziell ausgefallen sind. Und in Dänemark versuchte die Regierung, jede Tastatureingabe und jeden Mausklick zu überwachen, die Student:innen bei Prüfungen auf ihren Computern vornahmen. Dies löste – auch hier – massive Proteste der Studierenden aus, die dazu führten, dass das System vorerst wieder eingestellt wurde.

/ Es ist an der Zeit, Fehlentwicklungen bei ADM-Systemen zu korrigieren

ADM-Systeme haben prinzipiell das Potenzial, das Leben der Menschen positiv zu beeinflussen – indem sie riesige Datenmengen verarbeiten, Verantwortliche in Entscheidungsfindungsprozessen unterstützen und maßgeschneiderte Anwendungen bieten.

In der Praxis haben wir allerdings nur wenige überzeugende Beispiele für einen solchen positiven Einfluss gefunden.

In der Ausgabe von 2019 noch kaum erwähnt, wird Gesichtserkennung inzwischen mit alarmierender Geschwindigkeit in ganz Europa getestet und eingesetzt.

So ist etwa das System VioGén, das in Spanien seit 2007 im Einsatz ist, um Risiken in Fällen von häuslicher Gewalt abzuschätzen, zwar bei Weitem noch nicht perfekt, zeigt jedoch „vernünftige Leistungskennzahlen“ und hat bereits geholfen, viele Frauen vor Gewalt zu schützen.

In Portugal hat ein zentralisiertes automatisiertes System zur Aufdeckung von Betrug im Zusammenhang mit Rezeptverschreibungen die Betrugsfälle innerhalb eines einzigen Jahres **nachweislich** um 80 % reduziert. In Slowenien hat sich den Aussagen der Steuerbehörden zufolge ein ähnliches System als nützlich erwiesen, das verwendet wird, um Steuerbetrug zu bekämpfen.⁴

Doch die aktuelle Bestandsaufnahme zum Einsatz von ADM-Systemen in Europa zeigt, dass positive Beispiele, die echte Vorteile bringen, rar gesät sind. Im gesamten Bericht beschreiben wir, wie die überwältigende Mehrheit der aktuellen Nutzungen Menschen tendenziell eher einem Risiko aussetzt als ihnen zu helfen. Um aber die tatsächlichen positiven und negativen Effekte beurteilen zu können, brauchen wir mehr Transparenz im Hinblick auf die Ziele, mehr Daten über die Funktionsweise der ADM-Systeme, die derzeit getestet und eingesetzt werden.

Die Botschaft an die politischen Entscheidungsträger:innen könnte klarer nicht sein: Wollen wir ihr Potenzial wirklich ausschöpfen und zugleich dafür sorgen, dass Menschenrechte und Demokratie respektiert werden, dann ist es an der Zeit, diese Systeme transparent zu machen und die Fehlentwicklungen bei ADM zu korrigieren. Jetzt.

/ Gesichtserkennung, überall Gesichtserkennung

In verschiedenen Ländern kommen verschiedene Tools zum Einsatz. Eine Technologie ist heute aber beinahe flächendeckend zu finden: Gesichtserkennung. Hierbei handelt es sich ohne Zweifel um die neueste, rasanteste und besorgniserregendste Entwicklung, die in diesem Bericht beleuchtet wird. In der Ausgabe von 2019 noch kaum erwähnt, wird Gesichtserkennung inzwischen mit alarmierender Geschwindigkeit in ganz Europa getestet und eingesetzt. Innerhalb von wenig mehr als einem Jahr seit unserem letzten Bericht ist Gesichtserkennung in Schulen und Stadien, an Flughäfen, ja selbst in Spielcasinos zu finden.

⁴ Weitere Details siehe Kapitel Slovenia in der Gesamtausgabe des Automating Society Reports unter <https://automating.society.algorithmwatch.org/report2020/slovenia/>

Sie kommt ebenfalls für die vorhersagende Polizeiarbeit (Predictive Policing) zum Einsatz, um Kriminelle abzuschrecken, um gegen [Rassismus](#) vorzugehen sowie, im Hinblick auf die COVID-19-Pandemie, zur Durchsetzung des Social Distancing, und zwar sowohl in Form von Apps, wie auch durch „smarte“ Videoüberwachung.

Ständig kommen neue Einsatzgebiete hinzu, obwohl die Zahl von Belegen für ihre mangelnde [Genauigkeit zunimmt](#). Und gibt es Widerstand dagegen, dann versuchen Verfechter:innen dieser Systeme ganz einfach, ihn zu umgehen. In Belgien ist ein von der Polizei genutztes Gesichtserkennungssystem immer noch „teilweise aktiv“, obwohl die Aufsichtsbehörde für Polizeilichen Informationsaustausch ein befristetes Verbot erlassen hat. Und in Slowenien wurde die Nutzung von Gesichtserkennungstechnologie durch die Polizei fünf Jahre, nachdem sie erstmals eingesetzt wurde, legalisiert.

Stellt man sich diesem Trend nicht entgegen, besteht die Gefahr, dass die Vorstellung, ununterbrochen – und verdeckt – beobachtet zu werden, zur Normalität wird, wodurch sich ein neuer Status quo von flächendeckender Massenüberwachung verfestigt. Eben aus diesem Grund hätten sich viele Akteur:innen der Bürgerrechtsbewegung eine weitaus aggressivere politische Antwort von Seiten der EU-Institutionen auf diese Entwicklungen gewünscht.⁵

Im Rahmen eines Pilotprojektes, bei dem derzeit ein ADM-System in Banken in Polen zum Einsatz kommt, spielt sogar das Lächeln eine Rolle: Je mehr die Angestellten lächeln, desto höher ist ihr Bonus. Und es werden nicht nur Gesichter überwacht. In Italien gab es den Vorschlag, in allen Fußballstadien ein Geräuschüberwachungssystem zu installieren, um Rassismus zu bekämpfen.

/ Blackboxes sind immer noch Blackboxes

Zu den alarmierenden Ergebnissen dieses Berichts gehört, dass zwar immer mehr ADM-Systeme eingesetzt werden, aber die Transparenz nicht Schritt hält. Im Jahr 2015 prägte Frank Pasquale, Professor an der Brooklyn Law School, einen berühmt gewordenen Begriff. Er nannte eine Gesellschaft, die von Netzwerken durchzogen ist, die auf intransparenten algorithmischen Systemen basieren, eine „Black Box Society“. Fünf Jahre später trifft diese Metapher leider immer noch zu – und sie gilt ausnahmslos für alle Länder,

die wir für diesen Bericht untersucht haben: Es gibt im Hinblick auf ADM-Systeme viel zu wenig Transparenz, sowohl im öffentlichen wie auch im privaten Sektor. Polen verfügte Intransparenz sogar von ganz oben – mit dem Gesetz zur Einführung eines automatisierten Systems zum Aufspüren von Bankkonten, die für illegale Aktivitäten genutzt werden („STIR“). Es sieht vor, dass mit bis zu 5 Jahren Gefängnis bestraft werden kann, wer Algorithmen und Risikoindikatoren offenlegt.

Zwar lehnen wir die Vorstellung, dass alle solche Systeme von Natur aus schlecht sind, nachdrücklich ab und nehmen stattdessen eine evidenzbasierte Perspektive ein. Dennoch ist es ohne Zweifel höchst problematisch, nicht in der Lage zu sein, ihre Funktionsweise und Auswirkungen auf der Grundlage echter Sachkenntnis und korrekter Fakten zu beurteilen. Schon deshalb, weil Intransparenz das Sammeln von Beweisen, die nötig sind, um überhaupt ein informiertes Urteil zur Nutzung von ADM-Systemen abgeben zu können, ernsthaft behindert.

Nimmt man dann noch die Schwierigkeiten hinzu, die sowohl unsere Wissenschaftler:innen wie auch unsere Journalist:innen hatten, wenn sie Zugang zu aussagekräftigen Daten über diese Systeme haben wollten, dann zeichnet sich hier für alle ein beunruhigendes Szenario ab, die solche Systeme kontrollieren und sicherstellen wollen, dass ihr Einsatz mit grundlegenden Rechten, gesetzlichen Regelungen und der Demokratie in Einklang steht.

/ Den algorithmischen Status quo hinterfragen

Was tut die Europäische Union in dieser Frage? Auch wenn die von der EU-Kommission unter der Leitung von Ursula von der Leyen erarbeiteten Dokumente sich eher auf „Künstliche Intelligenz“ beziehen als ADM-Systeme direkt anzusprechen, enthalten sie lobenswerte Absichten: Eine „vertrauenswürdigen KI“ soll gefördert und umgesetzt werden, die „die Menschen in den Vordergrund stellt“⁶.

Allerdings gibt die EU, wie im EU-Kapitel beschrieben, der vermeintlichen kommerziellen und geopolitischen Notwendigkeit, die „KI-Revolution“ anzuführen, Vorrang davor, sicherzustellen, dass die Ergebnisse solcher KI demokratischer Kontrolle unterliegen.

⁵ Wie ausführlich im EU-Kapitel dargelegt.

⁶ Vgl. das EU-Kapitel, dort insbesondere den Abschnitt zum „Weißbuch KI“ der EU-Kommission.

Zu den alarmierenden Ergebnissen dieses Berichts gehört, dass zwar immer mehr ADM-Systeme eingesetzt werden, aber die Transparenz nicht Schritt hält.

Dieser Mangel an politischem Mut wird daran am deutlichsten, dass im Weißbuch zu KI jeder Hinweis auf ein Moratorium für Systeme getilgt wurde, die an öffentlichen Orten Technologien zur Gesichtserkennung in Echtzeit einsetzen. Er ist überraschend – besonders zu einer Zeit, in der sich viele Mitgliedsstaaten im Zusammenhang mit allzu hastig implementierten ADM-Systemen, die sich negativ auf die Rechte von Bürger:innen ausgewirkt haben, mit einer zunehmenden Anzahl gerichtlicher Klagen – und Niederlagen – konfrontiert sehen.

Ein besonders wegweisender Fall stammt aus den Niederlanden. Dort zogen Bürgerrechtsaktivist:innen gegen ein invasives und intransparentes automatisiertes System zur Aufdeckung von Sozialbetrug (SyRI) vor Gericht und gewannen. Das Gericht in Den Haag urteilte, dass das System eine Verletzung der Europäischen Menschenrechtskonvention darstelle, und stoppte es. Die Sache wurde auch zu einem Präzedenzfall: Dem Urteil zufolge haben Regierungen eine „besondere Verantwortung“, Menschenrechte zu schützen, wenn sie solche ADM-Systeme implementieren. Für die so dringend benötigte Transparenz zu sorgen, wird als essenzieller Bestandteil dieser Verantwortung angesehen.

Seit unserem ersten Bericht haben sich die Medien und die Aktivist:innen der Zivilgesellschaft als Triebkräfte etabliert, die die Rechenschaftspflicht von ADM-Systemen einfordern. So ist es zum Beispiel in Schweden Journalist:innen gelungen, die Offenlegung des Codes zu erzwingen, der hinter dem Trelleborg-System für vollautomatisierte Entscheidungen über Anträge auf Sozialhilfe steht. In Berlin scheiterte das am Bahnhof Südkreuz durchgeführte Pilotprojekt zur Gesichtserkennung, und es kam nirgendwo in Deutschland zu einer Implementierung des Systems. Dies war auf den lautstarken Widerspruch von Aktivist:innen zurückzuführen – so lautstark, dass es ihnen gelang, Par-

teipositionen zu beeinflussen und damit letztendlich die politische Agenda der Regierung.

Griechische Aktivist:innen von Homo Digitalis deckten auf, dass an Versuchen mit dem griechischen Pilotprojekt eines Systems namens ‚iBorderCtrl‘, einem von der EU finanzierten Projekt zum Einsatz von ADM bei Grenzkontrollen, kein einziger echter Reisender teilgenommen hatte und enthüllten damit, dass die Fähigkeiten vieler solcher Systeme häufig übertrieben dargestellt werden. In Dänemark wurde zwischenzeitlich dank der Arbeit von Akademiker:innen und Journalist:innen sowie der Datenschutzbehörde (DPA) ein Profiling-System gestoppt, das Risiken erkennen sollte, die mit vulnerablen Familien und Kindern in Verbindung gebracht werden (das sogenannte „Gladsaxe-Modell“).

Auch in anderen Ländern spielten die Datenschutzbehörden selbst eine wichtige Rolle. In Frankreich befand die Nationale Datenschutzbehörde sowohl ein Projekt zur Geräuschüberwachung wie auch eines zur Gesichtserkennung an Oberschulen für illegal. In Portugal weigerte sich die Datenschutzbehörde, den Einsatz eines Videoüberwachungssystems durch die Polizei in den Städten Leiria und Portimão zu genehmigen, da sie es für überdimensioniert hielt. Sein Einsatz hätte nicht nur zu einer „großflächigen und systematischen Überwachung und Verfolgung von Menschen, ihren Gewohnheiten und ihres Verhaltens“ geführt, sondern auch „zur Identifizierung von Menschen anhand von Daten im Zusammenhang mit physischen Merkmalen“. Und in den Niederlanden forderte die staatliche Datenschutzbehörde mehr Transparenz bei voraussagenden Algorithmen, die von Regierungsbehörden genutzt werden.

Einige Länder wandten sich schließlich mit der Bitte um Rat an eine Ombudsperson. In Dänemark leistete diese Form der Beratung einen Beitrag zur Entwicklung von Strategien und ethischen Richtlinien beim Einsatz von ADM-Systemen

im öffentlichen Sektor. In Finnland stufte die stellvertretende parlamentarische Ombudsperson automatisierte Steuerschätzungen als illegal ein.

Und dennoch fragt man sich angesichts des ungebremsten Einsatzes solcher Systeme in ganz Europa nach wie vor verwundert: Reicht dieses Maß an Kontrolle aus? Als die Ombudsperson in Polen die Legalität des in einer Bank genutzten (und oben erwähnten) Lächel-Detektors in Zweifel zog, verhinderte diese Entscheidung weder ein späteres Pilotprojekt in der Stadt Sopot, noch hielt es mehrere Unternehmen davon ab, ihr Interesse an einer Übernahme des Systems zu bekunden.

/ Es fehlt an allem: Kontrolle, Durchsetzung, Kompetenzen und Erläuterungen

Aktivismus ist ein größtenteils reaktives Unterfangen. Aktivist:innen können meist nur dann reagieren, wenn ADM-Systeme in der Erprobung sind oder bereits zum Einsatz kommen. In dem Zeitraum, den Bürger:innen brauchen, um Widerstand zu organisieren, könnten ihre Rechte bereits unnötigerweise verletzt worden sein. Und dies kann trotz der Schutzmechanismen geschehen, die in den meisten Fällen durch EU-Gesetze oder Gesetze der Mitgliedsstaaten garantiert sein sollten. Aus diesem Grund sind proaktive Maßnahmen zum Schutz von Bürgerrechten so überaus wichtig – bevor Pilotprojekte gestartet und Systeme eingesetzt werden.

Dennoch werden selbst in jenen Ländern, in denen Schutzgesetze existieren, diese schlicht nicht durchgesetzt. So ist etwa in Spanien „automatisiertes Verwaltungshandeln“ gesetzlich geregelt, und es gelten besondere Anforderungen im Hinblick auf Qualitätskontrolle und behördliche Aufsicht. Festgeschrieben ist ebenfalls eine Auditierung des Informationssystems und seines Quellcodes. Zudem hat Spanien ein Gesetz zur Informationsfreiheit. Ungeachtet dessen geben öffentliche Einrichtungen, so schreibt unser Autor, nur selten detaillierte Informationen zu den von ihnen genutzten ADM-Systemen frei. Ähnliches gilt für Frankreich. Dort gibt es seit 2016 ein Gesetz, das algorithmische Transparenz verpflichtend macht, doch auch hier ohne jede Wirkung.

Unter Umständen reicht noch nicht einmal die gerichtliche Klage gegen einen Algorithmus auf der Grundlage spezieller Transparenzregelungen aus, um Rechte von Nutzer:innen durchzusetzen und zu schützen. Wie der Fall des Parcour-

soup-Algorithmus in Frankreich zeigt, der von Universitäten genutzt wird, um Studienbewerber:innen zu sortieren, können nach Belieben Ausnahmen erarbeitet werden, um eine Verwaltung von der Rechenschaftspflicht auszunehmen.

Besonders beunruhigend ist dies dann, wenn es mit dem verbreiteten Mangel an Fähigkeiten und Kompetenzen rund um ADM-Systeme im öffentlichen Sektor einhergeht, wie er von vielen Fachleuten beklagt wird. Wie sollten Amtsträger:innen auch Auskunft geben oder für Transparenz sorgen im Zusammenhang mit Systemen, die sie nicht verstehen?

Einige Länder haben jüngst versucht, dieses Problem anzugehen. So wurde in Estland ein speziell auf ADM-Systeme ausgerichtetes Kompetenzzentrum gegründet. Dieses soll helfen, ein besseres Verständnis dafür zu entwickeln, wie solche Systeme bei der Weiterentwicklung öffentlicher Dienstleistungen zum Einsatz kommen können, und speziell dem Ministerium für Wirtschaft und Kommunikation sowie der Staatskanzlei Informationen und Grundlagen für ihr weiteres Vorgehen bei der Entwicklung von e-Government liefern. Ebenso hat die Schweiz im Rahmen ihrer breit angelegten nationalen Strategie für eine „Digitale Schweiz“ ein „Kompetenznetzwerk“ gefordert.

Ungeachtet dessen sind mangelhafte digitale Kompetenzen ein bekanntes Problem – mit Auswirkungen auf einen Großteil der Bevölkerung etlicher europäischer Staaten. Abgesehen davon ist es ziemlich schwer, die Durchsetzung von Rechten einzufordern, von denen man nicht einmal weiß, dass man sie hat. Die Proteste im Vereinigten Königreich und anderswo sowie die bekannten Skandale im Zusammenhang mit der Nutzung von ADM-Systemen⁷ haben mit Sicherheit das Bewusstsein für die Risiken und Chancen einer zunehmend automatisierten Gesellschaft erhöht. Allerdings ist dieses Bewusstsein, obgleich es zunimmt, in vielen Ländern nach wie vor unterentwickelt.

Die Ergebnisse unserer Recherchen sind eindeutig: ADM-Systeme nehmen bereits Einfluss auf alle möglichen Aktivitäten und Beurteilungen, doch ihr Einsatz erfolgt nach wie vor größtenteils ohne jede vernünftige demokratische Debatte. Darüber hinaus ist es eher die Regel als die Ausnahme, dass die Durchsetzungs- und Aufsichtsmechanismen

⁷ Vgl. das Kapitel zu France in der Gesamtausgabe des Automating Society Reports unter <https://automatingsociety.algorithmwatch.org/report2020/france/>

– so sie denn überhaupt existieren – dieser Entwicklung hinterherhinken.

Selbst der Zweck dieser Systeme wird den betroffenen Bürger:innen gegenüber nicht generell gerechtfertigt oder erläutert, ganz zu schweigen von den Vorteilen, die sie bringen sollen. Man denke nur an den proaktiven Service „AuroraAI“ in Finnland: Dieser soll, wie unsere finnischen Autor:innen berichten, automatisch „Lebensereignisse“ identifizieren; nach den Vorstellungen seiner Befürworter:innen ist er dazu gedacht, als eine Art „Kinder mädchen“ zu fungieren, das Bürger:innen hilft, spezielle Anforderungen staatlicher Dienstleister im Zusammenhang mit bestimmten Lebensumständen zu erfüllen, wie etwa einem Umzug, veränderten Familienverhältnissen usw. Es könne hier „Nudging“ am Werk sein, schreiben unsere Autor:innen weiter. Damit meinen sie, dass das System, anstatt Individuen zu ermächtigen, am Ende zum genauen Gegenteil führen könnte, dass es nämlich aufgrund seines speziellen Designs und seiner Architektur bestimmte Entscheidungen suggerieren oder die Optionen einer Person beschränken könnte.

Und deshalb ist es umso wichtiger zu wissen, was genau im Hinblick auf staatliche Dienstleistungen eigentlich „optimiert“ werden soll: „Wird die Servicenutzung maximiert, werden Kosten minimiert, oder wird das Wohlergehen der Bürger:innen verbessert?“, so die Frage der Forscher:innen. „Auf welchen Kriterien basieren diese Entscheidungen, und wer trifft sie?“ Die bloße Tatsache, dass wir auf diese fundamentalen Fragen keine Antwort haben, spricht Bände über den Grad an erlaubter Beteiligung und Transparenz, selbst für ein potenziell invasives ADM-System wie dieses.

/ Die Falle des „technologischen Solutionismus“

Für all dies gibt es eine übergreifende ideologische Rechtfertigung. Sie wird „Technologischer Solutionismus“ genannt und beeinflusst nach wie vor sehr stark die Art und Weise, wie viele der von uns untersuchten ADM-Systeme entwickelt werden. Auch wenn der Begriff schon lange angeprangert wird, weil er eine fehlerhafte Ideologie bezeichnet, die jedes gesellschaftliche Problem als „Fehler“ begreift, der mit Hilfe von Technologie „repariert“ werden muss⁸, findet diese Rhetorik in den Medien und Kreisen der

8 Man denke an das Debakel um den „Buona Scuola“-Algorithmus in Italien; vgl. das Kapitel Italy in der Gesamtausgabe des Automating Society Reports unter <https://automatingsociety.algorithmwatch.org/report2020/italy/>

Politik nach wie vor breite Anwendung, um die unkritische Übernahme automatisierter Technologien ins öffentliche Leben zu rechtfertigen.

Werden sie als „Lösung“ verkauft, steuern ADM-Systeme unverzüglich in das Territorium hinein, das in Arthur C. Clarkes Drittem Gesetz beschrieben wird: Magie. Magie zu regulieren, ist schwierig, wenn nicht gar unmöglich. Und diesbezüglich für Transparenz zu sorgen und Erläuterungen zu geben, noch weitaus schwieriger. Man sieht zu, wie die Hand in den Hut fasst und ein Kaninchen hervorzieht, doch der Vorgang selbst ist eine „Blackbox“ und *soll es auch bleiben*.

Diesen Umstand prangerten zahlreiche am Projekt „Automating Society“ beteiligte Wissenschaftler:innen als fundamentalsten Mangel in den Argumentationen an, die hinter vielen der von ihnen beschriebenen ADM-Systeme stehen. Wie im Kapitel über Deutschland dargelegt, impliziert dies auch, dass ein Großteil der an solchen Systemen geäußerten Kritik als pauschale Ablehnung von „Innovationen“ verleumdet und Befürworter:innen digitaler Rechte als „Neo-Ludditen“ (Neue Maschinenstürmer:innen) dargestellt werden. Damit werden nicht nur die historischen Realitäten der Maschinenstürmerbewegung ignoriert, die sich auf Arbeitsmarktpolitik bezog und nicht auf Technologien per se. Nein, es gefährdet auch – und dies ist vielleicht noch wichtiger – die Effektivität möglicher Aufsichts- und Durchsetzungsmechanismen.

Zu einer Zeit, da die „KI“-Branche das Entstehen einer „lebhaften“ Lobby-Industrie erlebt, insbesondere im Vereinigten Königreich, könnte dies dazu führen, dass Leitlinien für „ethics-washing“ und anderen politischen Reaktionen verabschiedet werden, die ebenso ineffektiv wie strukturell inadäquat sind, um die Folgen von ADM-Systemen für die Menschenrechte zu bewältigen. Diese Sichtweise gipfelt letztlich in der Prämisse, dass nicht ADM-Systeme an demokratische Gesellschaften angepasst werden sollten, sondern vielmehr wir Menschen uns den ADM-Systemen anpassen sollten.

Um diesem Narrativ entgegenzutreten, sollten wir uns nicht scheuen, einige grundsätzliche Fragen zu stellen: nämlich, ob ADM-Systeme demokratiekompatibel sein und zum Wohl der Gesellschaft als Ganzes zum Einsatz kommen können, anstatt nur in Teilbereichen. So könnte es zum Beispiel sein, dass bestimmte menschliche Aktivitäten – etwa jene, die die Sozialfürsorge betreffen – nicht der Automatisierung unterworfen werden sollten, oder dass bestimmte

Technologien – namentlich Gesichtserkennung in Echtzeit an öffentlichen Plätzen – nicht beim endlosen Streben nach „KI-Leadership“ gefördert, sondern stattdessen verboten werden sollten.

Noch wichtiger ist, dass wir jedes ideologische Framing, das uns hindert, solche Fragen zu stellen, ablehnen sollten. Im Gegenteil: Was es jetzt braucht, ist ein echter Politikwandel, um für eine strengere Überprüfung dieser Systeme zu sorgen. Im folgenden Abschnitt listen wir die Kernforderungen auf, die sich aus unseren Rechercheergebnissen ableiten. Wir hoffen, dass sie breit diskutiert und am Ende umgesetzt werden.

Nur durch eine informierte, inklusive und evidenzbasierte demokratischen Debatte wird es uns gelingen, die richtige Balance zu finden zwischen den Vorteilen, die ADM-Systeme im Hinblick auf Schnelligkeit, Effizienz, Fairness, bessere Prävention und besseren Zugang zu staatlichen Dienstleistungen bieten können – und dies auch tun – und den Gefahren, die sie für die Rechte von uns allen darstellen.

Handlungsempfehlungen

Basierend auf den Forschungsergebnissen unseres Berichts „Automating Society 2020“, empfehlen wir den Entscheidungsträger:innen im EU-Parlament und den Parlamenten der Mitgliedsstaaten, der EU-Kommission, den Nationalregierungen, Wissenschaftler:innen, Organisationen der Zivilgesellschaft (Interessenvertretungen, Stiftungen, Gewerkschaften usw.) sowie dem privatwirtschaftlichen Sektor (Unternehmen und Branchenverbänden) die folgenden Maßnahmen. Diese Empfehlungen sollen helfen, sicherzustellen, dass ADM-Systeme, die derzeit in ganz Europa im Einsatz sind oder eingeführt werden sollen, tatsächlich mit Menschenrechten und Demokratie vereinbar sind:

1. **Transparenz von ADM-Systemen erhöhen**

Ohne die Möglichkeit, genau zu wissen, wie, warum und zu welchem Zweck ADM-Systeme zum Einsatz kommen, sind alle anderen Anstrengungen, solche Systeme mit Grundrechten in Einklang zu bringen, zum Scheitern verurteilt.

/ Öffentliche Register für ADM-Systeme einführen, die von der öffentlichen Hand genutzt werden

Mitgliedsstaaten sollten durch EU-weite Regelungen verpflichtet werden, öffentliche Register für ADM-Systeme einzurichten, die im öffentlichen Sektor genutzt werden.

Damit verbunden sein muss die gesetzliche Verpflichtung, dass die für ein ADM-System verantwortlichen Stellen bzw. Personen den Zweck des Systems offenlegen und dokumentieren, ergänzt um eine Erläuterung des Modells (einschließlich seiner Logik) sowie Informationen darüber, wer das System entwickelt hat. Diese Informationen müssen auf einfach zu verstehende und leicht zugängliche Weise verfügbar gemacht werden, einschließlich strukturierter digitaler Daten, die auf einem standardisierten Protokoll basieren.

Staatliche Behörden haben eine besondere Verantwortung, die Betriebseigenschaften von ADM-Systemen, die in der öffentlichen Verwaltung zum Einsatz kommen, transparent zu machen. Unterstrichen wird dies von der Entscheidung zu einer kürzlich eingereichten Verwaltungsbeschwerde in Spanien. Dort heißt es, dass „jedes von der staatlichen Verwaltung genutzte ADM-System standardmäßig öffentlich gemacht werden sollte“. Hat dieses Urteil Bestand, könnte es zu einem Präzedenzfall für ganz Europa werden.

Für ADM-Systeme, die im öffentlichen Sektor zum Einsatz kommen, sollten Offenlegungsmechanismen in allen Fällen verpflichtend sein. Für den Einsatz von ADM-Systemen durch private Unternehmen sollten sie gelten, sofern ein KI-/ADM-System signifikante Auswirkungen auf eine Einzelperson, eine bestimmte Gruppe oder die Gesellschaft als Ganzes hat.

/ Rechtlich verbindliche Rahmenbedingungen für den Zugang zu Daten einführen, um Forschung im öffentlichen Interesse zu unterstützen und zu ermöglichen

Verstärkte Transparenz erfordert neben der Offenlegung von Informationen über Zweck, Logik und Entwickler:in eines Systems auch die Möglichkeit, dessen Inputs und Outputs einer tiefgreifenden Analyse und Prüfung zu unterziehen. Darüber hinaus müssen die Trainingsdaten und Datenergebnisse unabhängigen Wissenschaftler:innen, Journalist:innen und Aktivist:innen der Zivilgesellschaft für Forschung und Recherche im öffentlichen Interesse zugänglich gemacht werden.

Aus diesem Grund empfehlen wir die Einführung robuster, rechtlich verbindlicher Rahmenbedingungen für den Zugang zu Daten, die explizit darauf ausgerichtet sind, Forschung im öffentlichen Interesse zu ermöglichen und zu unterstützen, und die die Regelungen zum Datenschutz sowie die Gesetze zum Schutz der Privatsphäre respektieren.

In Anlehnung an bestehende Best Practices auf nationaler und EU-Ebene sollten solche abgestuften Rahmenwerke sowohl Sanktionssysteme wie auch Kontrollmechanismen und regelmäßige Überprüfungen einschließen. Wie private Datensharing-Partnerschaften gezeigt haben, bestehen berechtigte Bedenken im Hinblick auf die Privatsphäre der Nutzer:innen und die mögliche Deanonymisierung bestimmter Arten von Daten.

Politische Entscheidungsträger:innen sollten sich Regelungen für die gemeinsame Nutzung von Gesundheitsdaten zum Vorbild nehmen, um den privilegierten Zugang zu bestimmten Arten granularer Daten zu ermöglichen und gleichzeitig sicherzustellen, dass persönliche Daten ausreichend geschützt sind (z. B. durch sichere Betriebsumgebungen).

Rechenschaftspflichten durchzusetzen ist nur möglich, wenn es einen Zugang zu Daten der Plattformen gibt. Das ist zugleich eine Voraussetzung für die effektive Umsetzung zahlreicher Ansätze zur Überprüfung von Systemen.

2. Einen eindeutigen Rahmen für die Rechenschaftspflicht von ADM-Systeme schaffen

Wie Forschungsergebnisse aus Spanien und Frankreich gezeigt haben, führen gesetzlich verankerte Transparenzfordernisse und/oder die Offenlegung von Informationen nicht automatisch zu verantwortungsvollem Handeln. Weitere Schritte sind notwendig, um sicherzustellen, dass Gesetze und Vorschriften auch tatsächlich durchgesetzt werden können.

/ Ansätze für die effektive Auditierung algorithmischer System entwickeln und etablieren

Um für echte Transparenz zu sorgen, muss der erste Schritt (die Einrichtung öffentlicher Register), um Prozesse ergänzt werden, die eine effektive Auditierung algorithmischer Systeme garantieren.

Der Begriff „Auditing“ ist weit verbreitet, doch gibt es keine allgemeinverbindliche Definition. Wir verstehen Auditing in diesem Kontext gemäß der ISO-Definition als einen „systematischen, unabhängigen und dokumentierten Prozess, um objektive Nachweise zu erlangen und diese objektiv zu bewerten um festzustellen, inwiefern die Prüfkriterien erfüllt sind“.⁹

Wir haben bisher noch keine zufriedenstellenden Antworten auf die komplexen Fragen¹⁰, die von der Auditierung algorithmischer Systeme aufgeworfen werden. Allerdings deuten unsere Ergebnisse klar darauf hin, dass diese Antworten gefunden werden müssen, und zwar im Rahmen einer breiten Debatte, an der alle Interessengruppen beteiligt sind, sowie durch gründliche, engagierte Forschung.

Es sollten sowohl Auditkriterien, als auch angemessene Auditprozesse entwickelt werden. Dabei ist ein Ansatz zu verfolgen, der nicht nur möglichst viele Beteiligte und Betroffene einbezieht, sondern auch die unverhältnismäßigen Auswirkungen von ADM-Systemen auf vulnerable Gruppen berücksichtigt und deren Mitsprache sichert.

Aus diesem Grund fordern wir die politischen Entscheidungsträger:innen auf, solche Stakeholder-Prozesse zu initiieren, um die angesprochenen Fragen zu klären, sowie Finanzierungsquellen zur Verfügung zu stellen, die die Teil-

⁹ <https://www.iso.org/obp/ui/#iso:std:iso:19011:ed-3:v1:en>

¹⁰ Beim Nachdenken über mögliche Modelle für algorithmisches Auditing tauchen diverse Fragen auf. 1) Wer/Was (Dienstleistungen/Plattformen/Produkte) soll auditiert werden? Wie können Auditing-Systeme auf die betreffende Art von Plattform/Dienstleistung zugeschnitten werden? 2) Wann sollte das Auditing von einer öffentlichen Stelle vorgenommen werden (auf EU-Ebene, nationaler Ebene, lokaler Ebene), und wann kann es von privaten Stellen/Expert:innen übernommen werden (Unternehmen, Zivilgesellschaft, Wissenschafter:innen)? 3) Wie stellt man die Unterscheidung zwischen der Beurteilung von Auswirkungen ex-ante (d. h. in der Designphase/Entwicklungsphase) und ex-post (d. h. im Betrieb) sowie im Hinblick auf die betreffenden Herausforderungen klar? 4) Wie können Kompromisse bei den verschiedenen Vor- und Nachteilen der Überprüfbarkeit bewertet werden? (So könnten etwa Simplität, Allgemeingültigkeit, Anwendbarkeit, Präzision, Flexibilität, Interpretierbarkeit, Datenschutz und Leistungsfähigkeit eines Auditing-Prozesses in einem Spannungsverhältnis stehen). 5) Welche Informationen müssen verfügbar sein, damit ein Audit effektiv und verlässlich ist (z. B. Quellcode, Trainingsdaten, Dokumentation)? Brauchen Prüfende physischen Zugang zu den Systemen, während sie in Betrieb sind, um einen effektiven Audit durchzuführen? 6) Welche Nachweispflichten für Händler:innen/Service-Anbieter:innen sind nötig und angemessen? 7) Wie können wir sicherstellen, dass das Auditing überhaupt möglich ist? Sollten Auditanforderungen bereits bei der Entwicklung algorithmischer Systeme Berücksichtigung finden („prüffähiges Design“)? 8) Vorschriften für die Publikation der Ergebnisse: Fällt ein Audit negativ aus, und die Probleme sind nicht behoben: Wie sollten sich Prüfende verhalten; auf welche Weise kann publik gemacht werden, dass ein Versagen vorliegt? 9) Wer prüft die Prüfenden? Wie stellt man die Rechenschaftspflicht der Prüfenden sicher?

nahme bisher nicht angemessen repräsentierter Interessengruppen sichern.

Weiterhin fordern wir die Bereitstellung ausreichender Mittel zur Unterstützung/Finanzierung von Forschungsprojekten, die sich mit der Entwicklung von Modellen zur effektiven Auditierung algorithmischer Systeme beschäftigen.

/ Zivilgesellschaftliche Organisationen als Watchdogs von ADM-Systemen unterstützen

Unsere Erkenntnisse zeigen deutlich, dass für eine effektive kritische Hinterfragung undurchsichtiger ADM-Systeme die Arbeit zivilgesellschaftlicher Organisationen unabdingbar ist. Mittels Forschung und Interessenvertretung, häufig auch in Kooperation mit Journalist:innen und dem akademischen Bereich, haben sie sich in den letzten Jahren wiederholt in politische Debatten um diese Systeme eingemischt. Dabei ist es ihnen in einigen Fällen gelungen, effektiv sicherzustellen, dass öffentliche Interessen und Grundrechte sowohl vor als auch nach der Implementierung solcher Systeme in vielen europäischen Ländern angemessen berücksichtigt werden.

Zivilgesellschaftliche Akteur:innen sollten daher als Watchdogs der zunehmend automatisierten Gesellschaft unterstützt werden. Als solche gehören sie zu den integralen Bestandteilen jedes effektiven Rahmenwerks für die Rechenschaftspflicht von ADM-Systemen.

/ Gesichtserkennung verbieten, die den Weg für Massenüberwachung ebnet

Nicht alle ADM-Systeme sind gleichermaßen gefährlich, und ein risikobasierter Regulierungsansatz, wie etwa in Deutschland und der EU, spiegelt diese Tatsache korrekt wider. Um jedoch eine praktikable Rechenschaftspflicht für Systeme zu gewährleisten, die als risikobehaftet eingestuft werden, müssen effektive Aufsichts- und Durchsetzungsmechanismen geschaffen werden. Dies ist umso wichtiger für solche Systeme, denen ein „hohes Risiko“ zur Verletzung von Nutzer:innenrechten bescheinigt wird.

Ein vordringliches Beispiel, das sich aus unseren Erkenntnissen ergeben hat, ist die Gesichtserkennung. Wie sich gezeigt hat, stellen ADM-Systeme, die auf biometrischen Technologien beruhen, darunter Gesichtserkennung, eine ernsthafte Bedrohung für das Gemeinwohl und die Grund-

rechte dar, denn sie ebnen den Weg zu undifferenzierter Massenüberwachung – insbesondere angesichts der Tatsache, dass sie trotzdem weiter verbreitet werden und auf intransparente Art und Weise zum Einsatz kommen.

Wir fordern, dass alle öffentlichen Nutzungen von Gesichtserkennung, die sich zu Massenüberwachung auswachsen könnten, auf EU-Ebene bis auf Weiteres verboten werden, und zwar so schnell wie möglich.

Solche Technologien könnten innerhalb der EU sogar jetzt schon als illegal gelten, zumindest für bestimmte Anwendungen, wenn ihr Einsatz ohne „ausdrückliche Zustimmung“ der gescannten Personen erfolgt. Diese juristische Auslegung wurde von den staatlichen Behörden in Belgien vorgeschlagen, die den Einsatz von Gesichtserkennung in ihrem Land in einem bahnbrechenden Urteil mit einer Strafzahlung belegten.

3. Algorithmische Kompetenzen verbessern und die öffentliche Debatte um ADM-Systeme stärken

Mehr Transparenz von ADM-Systemen ist nur dann von echtem Nutzen, wenn diejenigen, die mit ihnen konfrontiert werden, wie etwa Aufsichtsbehörden, Regierungen und Standardisierungsgremien, in der Lage sind, auf verantwortungsvolle und umsichtige Art und Weise mit diesen Systemen und ihren Auswirkungen umzugehen. Darüber hinaus müssen jene, die von diesen Systemen betroffen sind, verstehen können, wo, warum und wie diese Systeme zum Einsatz kommen. Aus diesem Grund müssen wir die Stärkung algorithmischer Kompetenz auf allen Ebenen vorantreiben, bei wichtigen Akteur:innen ebenso wie in der breiten Öffentlichkeit, und vielfältigere öffentliche Debatten über ADM-Systeme und deren Auswirkungen auf die Gesellschaft fördern.

/ Unabhängige Kompetenzzentren für ADM einrichten

Parallel zu unserer Forderung, das Auditing von Algorithmen und entsprechende Forschung zu unterstützen, plädieren wir dafür, unabhängige Kompetenzzentren für ADM auf nationaler Ebene einzurichten. Diese Zentren sollen dazu dienen, Algorithmen zu kontrollieren und zu bewerten, Forschung zu betreiben und Berichte zu erstellen. In Koordination mit Regulierungsbehörden, der Zivilgesellschaft

und der Wissenschaft sollten sie Regierung wie Industrie zu den Auswirkungen des Einsatzes von ADM-Systemen auf die Gesellschaft und Menschenrechte beraten. Die übergreifende Rolle dieser Zentren besteht darin, ein sinnvolles System der Rechenschaftspflicht zu entwickeln sowie die entsprechenden Kapazitäten zu schaffen.

Um Vertrauen, Transparenz und Kooperation zwischen allen Beteiligten aufzubauen, sollten die nationalen Kompetenzzentren zivilgesellschaftliche Organisationen, Interessengruppen sowie bereits bestehende Durchsetzungsstellen wie etwa Datenschutzbehörden und nationale Menschenrechtsgruppen einbeziehen.

Als unabhängige Körperschaften des öffentlichen Rechts käme den Kompetenzzentren eine zentrale Rolle zu: Sie würden die Entwicklung politischer Konzepte und nationaler Strategien im Umgang mit ADM koordinieren sowie Kompetenzen und Fähigkeiten bei bereits existierenden Regulierungsbehörden sowie Regierungen und Standardisierungsgremien aufbauen, um auf die zunehmende Nutzung von ADM-Systemen zu reagieren.

Diese Zentren sollten keine Regulierungsbefugnisse haben, sondern unverzichtbare Expertise einbringen, die die Frage beantworten helfen, wie Grundrechte geschützt und kollektiver und gesellschaftlicher Schaden abgewendet werden können. So sollten sie etwa kleine und mittlere Unternehmen (KMU) bei der Erfüllung ihrer Sorgfaltspflichten im Hinblick auf Menschenrechtsfragen unterstützen, einschließlich der Durchführung von Folgenabschätzungen für Menschenrechte (Human Rights Impact Assessments) oder Folgenabschätzungen zum Einsatz von Algorithmen (Algorithmic Impact Assessments), sowie bei der Registrierung von ADM-Systemen in dem oben dargestellten öffentlichen Register.

/ Eine inklusive und vielfältige demokratische Debatte rund um ADM-Systeme fördern

Es müssen nicht nur Kenntnisse und Kompetenzen bei denjenigen gestärkt werden, die ADM-Systeme zum Einsatz bringen. Ebenso unerlässlich ist es, mit Hilfe einer breiteren Debatte und vielfältiger Programme die algorithmischen Kompetenzen der Allgemeinheit zu fördern.

Unsere Erkenntnisse legen nahe, dass ADM-Systeme der breiteren Öffentlichkeit nicht nur intransparent bleiben,

wenn sie im Einsatz sind, sondern dass selbst die Entscheidung, ob ein ADM-System überhaupt zum Einsatz kommen soll oder nicht, für gewöhnlich ohne Wissen oder Beteiligung der Öffentlichkeit getroffen wird.

Es besteht daher die dringende Notwendigkeit, von Anfang an die Öffentlichkeit (das öffentliche Interesse) einzubeziehen, wenn darüber entschieden wird, ADM-Systeme einzusetzen.

Oder, um es etwas allgemeiner auszudrücken: Wir brauchen eine vielfältigere öffentliche Debatte zu den Auswirkungen von ADM. Wir müssen aufhören, ausschließlich Expert-innengruppen anzusprechen; wir müssen dafür sorgen, dass die breitere Öffentlichkeit einen besseren Zugang zu diesen Themen bekommt. Dies bedeutet, eine andere Sprache zu sprechen als einen technisch-juristischen Slang, um öffentliche Aufmerksamkeit zu erzeugen und Interesse zu wecken.

Zu diesem Zweck sollten in Ergänzung zu den bereits genannten Maßnahmen detaillierte Programme zum Aufbau und zur Förderung digitaler Selbstbestimmung etabliert werden. Wenn unser Ziel darin besteht, eine verstärkte, informierte öffentliche Debatte zu führen und den Bürger:innen der EU Selbstbestimmung zu ermöglichen, dann müssen wir damit beginnen, Kenntnisse zur Digitalisierung zu vermitteln und voranzutreiben und dabei besonderes Augenmerk auf die sozialen, ethischen und politischen Konsequenzen des Einsatzes von ADM-Systemen legen.

Europäische Union: Weichen- stellung für die Zukunft von **ADM** in **Europa**



Systeme zum automatisierten Entscheiden nehmen bei der Verteilung von Rechten und Dienstleistungen in Europa eine zentrale Rolle ein, und immer mehr Institutionen auf dem Kontinent erkennen deren Bedeutung für das öffentliche Leben an – sowohl die Chancen als auch die Herausforderungen.

Von Kristina Penner und Fabio Chiusi



Seit unserem ersten Bericht im Januar 2019 – und ungeachtet der Tatsache, dass die EU nach wie vor in einer breit geführten Debatte um „vertrauenswürdige“ Künstliche Intelligenz befangen ist – haben diverse Gremien, vom EU-Parlament bis zum Europarat, Dokumente veröffentlicht, die helfen sollen, für die kommenden Jahre, wenn nicht Jahrzehnte, einen gemeinsamen Kurs im Umgang mit ADM zu finden.

Im Sommer 2019 [versprach](#) die neugewählte Präsidentin der EU-Kommission, Ursula von der Leyen, eine selbsterklärte Technikoptimistin, innerhalb von 100 Tagen nach ihrem Amtsantritt „Rechtsvorschriften mit einem koordinierten europäischen Konzept für die menschlichen und ethischen Aspekte der künstlichen Intelligenz“ und die Regulierung von Künstliche Intelligenz (KI) vorzuschlagen. Stattdessen veröffentlichte die Europäische Kommission im Februar 2020 ein [„Weißbuch“ zur Künstlichen Intelligenz](#). Das Strategiepaket enthält „Ideen und Maßnahmen“, um die Bürger:innen der EU zu informieren und den Weg für künftige Gesetzgebungsverfahren zu ebnen. Darüber hinaus macht es sich für die „technologische Unabhängigkeit“ Europas stark: In von der Leyens eigenen [Worten](#) bedeutet dies „die Fähigkeit, die Europa haben muss, um seine eigenen Entscheidungen zu treffen, die auf seinen eigenen Werten basieren und seine eigenen Regeln respektieren.“

Dies werde dabei helfen, uns alle zu Tech-Optimist:innen zu machen.

Ein zweites grundlegendes Vorhaben, das Auswirkungen auf ADM in Europa hat, ist der in von der Leyens „Agenda für Europa“ angekündigte Digital Services Act (DSA). Das Gesetz soll die E-Commerce-Direktive ersetzen, die seit 2000 in Kraft ist. Es zielt darauf ab, „unsere Haftungs- und Sicherheitsregelungen für digitale Plattformen, Dienstleistungen und Produkte zu modernisieren und unseren Digitalen Binnenmarkt zu vollenden“ – was fundamentale Debatten um die Rolle von ADM bei politischen Entscheidungen zum Thema Content Moderation, Zwischenhändlerhaftung und Meinungsfreiheit im Allgemeinen auslöst¹¹.

Eine explizite Fokussierung auf ADM-Systeme findet sich in einer vom Ausschuss für Binnenmarkt und Verbraucherschutz des EU-Parlaments [verabschiedeten](#) Resolution sowie in einer [Empfehlung](#) „zu den menschenrechtlichen Auswirkungen algorithmischer Systeme“ des Ministerkomitees des Europarates.

¹¹ Detaillierte Anmerkungen und Empfehlungen rund um ADM-Systeme im Kontext des DSA finden sich in den Veröffentlichungen zum Projekt 'Governing Platforms' von AlgorithmWatch.

VIELE BEOBACHTER:INNEN SEHEN IN DER ART UND WEISE, WIE EU-INSTITUTIONEN UND INSBESONDERE DIE KOMMISSION IHRE ÜBERLEGUNGEN UND VORSCHLÄGE ZU KI UND ADM FORMULIEREN, EINE GRUNDLEGENDE SPANNUNG ZWISCHEN WIRTSCHAFTLICHEN IMPERATIVEN UND DEN ERFORDERNISSEN ZUR WAHRUNG VON GRUNDRECHTEN.

Es hat sich gezeigt, dass insbesondere der Europarat im Verlauf des Jahres 2019 eine zunehmend wichtige Rolle in der politischen Debatte um KI eingenommen hat. Zwar bleibt abzuwarten, wie stark sein tatsächlicher Einfluss auf die Regulierungsbemühungen sein wird, doch spricht einiges dafür, dass er als „Wächter“ der Menschenrechte fungieren kann. Besonders deutlich wird dies in der Empfehlung „[Unboxing Artificial Intelligence: 10 steps to protect Human Rights](#)“ der Europarat-Kommissarin für Menschenrechte Dunja Mijatović sowie in der Tätigkeit des im September 2019 gegründeten Ad-hoc-Ausschusses für künstliche Intelligenz (CAHAI).

Viele Beobachter:innen sehen in der Art und Weise, wie EU-Institutionen und insbesondere die Kommission ihre Überlegungen und Vorschläge zu KI und ADM formulieren, eine grundlegende Spannung zwischen wirtschaftlichen Imperativen und den Erfordernissen zur Wahrung von Grundrechten. Auf der einen Seite will Europa „die Nutzung von und die Nachfrage nach Daten und datenbasierten Produkten und Dienstleistungen auf dem gesamten Binnenmarkt erhöhen“, um zu einem „führenden Akteur“ bei industriellen Anwendungen von KI zu werden und angesichts des wachsenden Drucks von Rivalen wie den USA und China die Wettbewerbsfähigkeit europäischer Unternehmen zu stärken. Das gilt in besonderer Weise für ADM, denn hier liegt die Annahme zugrunde, dass die EU mit Hilfe dieser „datenagilen“ Wirtschaft zum „Vorbild für eine durch Daten ermächtigte Gesellschaft werden [kann], die bessere Entscheidungen trifft – in der Geschäftswelt wie im staatlichen Bereich“. Im Weißbuch zur Künstlichen Intelligenz heißt es sinngemäß: Daten sind das Lebenselixier wirtschaftlicher Entwicklung.

Auf der anderen Seite kann die automatische Verarbeitung von Daten in den Bereichen Gesundheit, Arbeit und Sozialwesen jedoch Entscheidungen herbeiführen, die diskriminierende und unfaire Ergebnisse zur Folge haben. Diese „dunkle Seite“ des Einsatzes von Algorithmen bei Prozessen der Entscheidungsfindung geht die EU mit einer Reihe von Prinzipien an. Im Fall von risikoreichen Systemen sollen Regeln garantieren, dass automatisierte Entscheidungsprozesse mit den Menschenrechten und sinnvollen demokratischen Kontrollen vereinbar sind. Dieser Ansatz wird von EU-Institutionen unter dem Begriff „menschenzentriert“ zusammengefasst; er sei einzigartig und stünde in fundamentalem Gegensatz zu den in den USA (profitorientiert) und China (bestimmt von nationalen Sicherheitsinteressen und Massenüberwachung) verfolgten Ansätzen.

Es sind jedoch Zweifel laut geworden, ob Europa beide Ziele gleichzeitig erreichen kann. Ein wichtiges Beispiel ist die Gesichtserkennung: Obwohl uns – wie dieser Bericht zeigt – inzwischen viele Belege für den Einsatz ungeprüfter und intransparenter Gesichtserkennungssysteme in den meisten EU-Mitgliedsstaaten vorliegen, hat es die EU-Kommission bisher versäumt, schnell und entschlossen zu handeln, um die Rechte der EU-Bürger:innen zu schützen. Wie geleakte Entwürfe zum Weißbuch KI der Europäischen Kommission enthüllten¹², war die EU im Begriff, „biometrische Fernidentifikation“ an öffentlichen Orten zu verbieten, scheute aber in letzter Sekunde davor zurück und warb stattdessen für eine „breite Debatte“ zu diesem Thema.

Unterdessen wird die Nutzung kontrovers diskutierter Anwendungen von ADM bei Grenzkontrollen, sogar einschließlich Gesichtserkennung, im Rahmen EU-finanzierter Projekte weiterhin vorangetrieben.

Politische Strategien und Debatten

/ Das Europäische Datenstrategiepaket und das Weißbuch zur Künstlichen Intelligenz

Die versprochene umfassende Gesetzgebung mit einem „koordinierten europäischen Konzept für die menschlichen und ethischen Aspekte der künstlichen Intelligenz“, die von der Leyen in ihrer „Agenda für Europa“ angekündigt hatte, wurde nicht in den „ersten 100 Tagen“ ihrer Amtszeit vorgelegt. Die EU-Kommission hat jedoch eine Reihe von Dokumenten veröffentlicht, denen sich ein Paket von Prinzipien und Konzepten entnehmen lässt, die als Grundlage dienen können.

Am 19. Februar 2020 wurden zeitgleich die „[Europäische Datenstrategie](#)“ und das „[Weißbuch Zur künstlichen Intelligenz](#)“ publiziert, in denen die wichtigsten Prinzipien des strategischen Konzeptes der EU für KI dargelegt werden (inklusive ADM-Systeme, obwohl diese nicht explizit genannt werden). Zu diesen Prinzipien zählen: „Der Mensch zuerst“ (Technologie, die für die Menschen arbeitet), technologische Neutralität (keine Technologie ist per se gut oder

¹² <https://www.politico.eu/article/eu-considers-temporary-ban-on-facial-recognition-in-public-spaces/>

schlecht; dies wird ausschließlich durch deren Nutzung bestimmt) und natürlich Unabhängigkeit und Optimismus. Wie von der Leyen [es ausdrückt](#): „Wir wollen unsere Unternehmen, unsere Wissenschaftler:innen, Innovator:innen, Unternehmer:innen ermutigen, künstliche Intelligenz zu entwickeln. Und wir wollen unsere Bürger:innen ermutigen, auf diese zu vertrauen und sie zu nutzen. Wir müssen dieses Potenzial freisetzen.“

All dem liegt die Auffassung zugrunde, dass neue Technologien keine neuen Werte mit sich bringen sollten. Die „neue digitale Welt“, die von der Leyens Administration anstrebt, soll Bürger:innen und Menschenrechte umfassend schützen. „Exzellenz“ und „Vertrauen“ – wie bereits im Titel des Weißbuchs hervorgehoben, werden als die tragenden Säulen angesehen, auf denen ein europäisches KI-Modell ruhen kann und sollte, wodurch es sich von den Strategien sowohl der USA wie auch Chinas unterscheidet.

In den Details des Weißbuchs sucht man diesen Anspruch allerdings vergeblich. So wird darin ein risikobasierter Ansatz für die Regulierung von KI dargelegt, bei dem der Grad an Regulierung den Auswirkungen von „KI“-Systemen auf das Leben der Bürger:innen entspricht. „Bei hohem Risiko wie etwa im Gesundheitswesen, bei der Polizei oder im Verkehrssektor“, heißt es in der [Presseerklärung](#), „sollten KI-Systeme transparent und nachvollziehbar sein sowie menschlicher Aufsicht unterliegen“. Zu den geforderten Schutzmechanismen gehören auch die Prüfung und Zertifizierung angewandeter Algorithmen; diese sollten ebenso weit Verbreitung finden wie bei „Kosmetikartikeln, Autos oder Spielzeug“. Für „KI-Anwendungen ohne hohes Risiko“ soll dagegen lediglich ein freiwilliges Kennzeichnungssystem gelten: „Die KI-Anwendungen der betreffenden Wirtschaftsakteure würden dann ein Gütesiegel erhalten.“

Kritiker:innen haben jedoch [angemerkt](#), dass bereits die in dem Papier vorgenommene Definition von „Risiko“ nicht nur ein Zirkelschluss, sondern auch zu vage ist, was dazu führen könnte, dass diverse wirkungsvolle ADM-Systeme

durch das Raster des vorgeschlagenen Rahmenwerks fallen¹³.

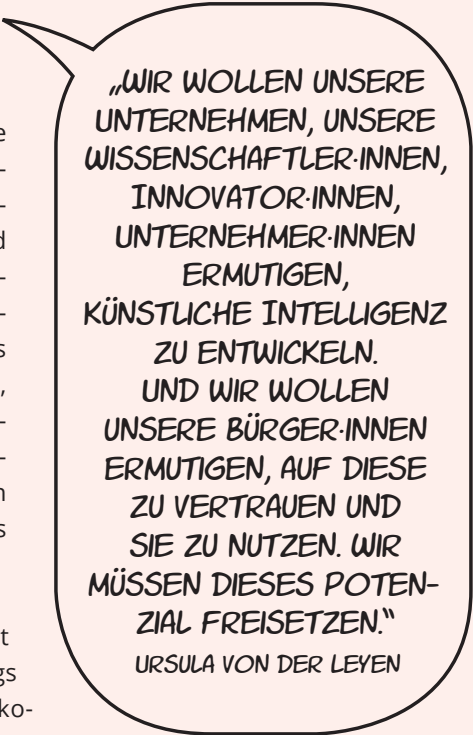
Die im Verlauf der öffentlichen Konsultation zwischen Februar und Juni 2020 eingegangenen Stellungnahmen¹⁴ machen deutlich, wie kontrovers dieses Konzept diskutiert wird. 42,5 % der Wortmeldungen stimmten darin überein,

dass „verpflichtende Anforderungen“ auf „KI-Anwendungen mit hohem Risiko“ beschränkt werden sollten, während 30,6 % einer solchen Beschränkung kritisch gegenüberstanden.

Davon abgesehen gibt es keinerlei Beschreibung klarer Durchsetzungsmechanismen für solche Anforderungen. Und ebenso wenig wird ein Prozess beschrieben, wie man zu solchen gelangen könnte.

Für biometrische Technologien, insbesondere Gesichtserkennung, sind die daraus resultierenden Konsequenzen sofort ersichtlich. Hierzu schlägt das Weißbuch eine Unterscheidung vor zwischen „biometrischer Authentifizierung“, die als nicht kontrovers betrachtet wird (z. B. Gesichtserkennung zur Entsperrung eines Smartphones), und „biometrischer Fernidentifikation“ (wie etwa der Einsatz an öffentlichen Orten, um die Identität von Protestierenden festzustellen),

die ernste Bedenken im Hinblick auf die Einhaltung von Menschenrechten und die Verletzung der Privatsphäre aufwerfen könnten.



„WIR WOLLEN UNSERE UNTERNEHMEN, UNSERE WISSENSCHAFTLER:INNEN, INNOVATOR:INNEN, UNTERNEHMER:INNEN ERMUTIGEN, KÜNSTLICHE INTELLIGENZ ZU ENTWICKELN. UND WIR WOLLEN UNSERE BÜRGER:INNEN ERMUTIGEN, AUF DIESE ZU VERTRAUEN UND SIE ZU NUTZEN. WIR MÜSSEN DIESES POTENZIAL FREISETZEN.“

URSULA VON DER LEYEN

13 „Um zwei Beispiele zu nennen: VioGén, ein ADM-System zur Vorhersage geschlechtsspezifischer Gewalt, und Ghostwriter, eine Anwendung zur Aufdeckung von Betrugsversuchen bei Prüfungen, würden mit hoher Wahrscheinlichkeit durchs Regulierungsraster fallen, obwohl sie mit extremen Risiken verbunden sind“ (<https://algorithmwatch.org/en/response-european-commission-ai-consultation/>)

14 „Insgesamt gingen 1.215 Stellungnahmen ein, davon 352 von Unternehmen und Unternehmer-/Branchenverbänden, 406 von Bürger:innen (92 % EU-Bürger:innen), 152 von akademischen Einrichtungen/Forschungsinstituten und 73 von staatlichen Behörden. Die Zivilgesellschaft war durch 160 Wortmeldungen vertreten (darunter 9 von Verbraucherschutzorganisationen, 129 von Nichtregierungsorganisationen und 22 von Gewerkschaften). 72 Stellungnahmen fielen unter „Sonstige“. Die Stellungnahmen kamen „aus aller Welt“, darunter „Indien, China, Japan, Syrien, Irak, Brasilien, Mexiko, Kanada, die USA und das Vereinigte Königreich“. (aus dem zusammenfassenden Bericht über die Konsultationen, siehe folgender Link: <https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence>)

Gemäß dem von der EU vorgeschlagenen System wären nur die Fälle der zweitgenannten Kategorie problematisch. In den [FAQ](#) zum Weißbuch heißt es: „Dies ist die Form der Gesichtserkennung, die den stärksten Eingriff in die Privatsphäre darstellt; sie ist in der EU grundsätzlich verboten“, es sei denn, es besteht ein „wesentliches öffentliches Interesse“ an ihrem Einsatz.

In dem erläuternden Dokument wird behauptet, die Genehmigung von Gesichtserkennung [ist] derzeit die Ausnahme. Die Erkenntnisse in diesem Bericht belegen jedoch unzweifelhaft das Gegenteil: Gesichtserkennung wird mit rasender Geschwindigkeit zur Norm. Eine geleakte Entwurfsfassung des Weißbuchs erkannte die Dringlichkeit des Problems offenbar an. Dort war der Vorschlag enthalten, ein drei bis fünf Jahre dauerndes Moratorium für den Einsatz von Gesichtserkennung an öffentlichen Orten zu verfügen, bis – und falls – ein Weg gefunden werden könne, sie mit demokratischen Kontrollmechanismen in Einklang zu bringen.

Unmittelbar vor der offiziellen Veröffentlichung des Weißbuchs [forderte](#) sogar EU-Kommissarin Margrethe Vestager, eine „Aussetzung“ dieser Nutzungen.

Kurz nach Vestagers Forderung fügten offizielle Vertreterinnen der Kommission jedoch hinzu, dass diese „Aussetzung“ nationale Regierungen nicht daran hindern würde, Gesichtserkennung gemäß den existierenden Regelungen einzusetzen. Letztendlich wurde in der finalen Fassung des Papiers jede Erwähnung eines Moratoriums gestrichen; stattdessen enthielt es die Ankündigung „einer breit angelegte[n] europäische[n] Debatte über die besonderen Umstände, die eine solche Nutzung rechtfertigen könnten“. Dazu gehört, so steht es im Weißbuch, dass der Einsatz hinreichend begründet und verhältnismäßig ist sowie demokratische Schutzmechanismen und die Achtung der Menschenrechte garantiert sind.

Das gesamte Dokument hindurch werden mit KI-basierten Technologien verbundene Risiken ganz allgemein als „potenziell“ bezeichnet, die Vorteile dagegen als sehr real und unmittelbar dargestellt. Dies veranlasste viele¹⁵ Menschenrechtsaktivist:innen zu der Aussage, das dem Weißbuch zugrundeliegende Narrativ spiegele eine besorg-

niserregende Umkehrung der Prioritätensetzung der EU wider; globale Wettbewerbsfähigkeit würde über den Schutz der Grundrechte gestellt.

Dennoch werden in den Dokumenten einige grundsätzliche Probleme angesprochen. So zum Beispiel die Interoperabilität solcher Lösungen und die Schaffung eines Netzwerks von Forschungszentren, die sich mit Anwendungsmöglichkeiten für KI beschäftigen und die auf „Exzellenz“ und den Aufbau von Kompetenzen ausgerichtet sind.

Das Ziel ist, „in der EU in den nächsten zehn Jahren insgesamt mehr als 20 Mrd. EUR an KI-Investitionen pro Jahr zu mobilisieren“.

Das Weißbuch scheint zudem von einem gewissen technologischen Determinismus geprägt zu sein. „Es ist äußerst wichtig“ heißt es darin, „dass öffentliche Verwaltungen, Krankenhäuser, Versorgungsbetriebe und Verkehrsdienste, Finanzaufsichtsbehörden und andere Bereiche von öffentlichem Interesse rasch mit der Einführung KI-gestützter Produkte und Dienstleistungen beginnen. Ein besonderer Schwerpunkt wird auf den Bereichen Gesundheitsfürsorge und Verkehr liegen, in denen die Technologien so weit ausgereift sind, dass sie in großem Maßstab eingesetzt werden können.“

Ob die Empfehlung zu einem raschen Einsatz von KI-Lösungen in allen Sphären menschlicher Aktivität mit den Anstrengungen der EU-Kommission vereinbar ist, die strukturellen Herausforderungen anzugehen, die ADM-Systemen im Hinblick auf Grundrechte und Fairness mit sich bringen, bleibt allerdings abzuwarten.

/ Entschließung des EU-Parlaments zu ADM und Verbraucherschutz

Eine vom EU-Parlament im Februar 2020 verabschiedete [Entschließung](#) ging das Thema ADM-Systeme im Kontext des Verbraucherschutzes gezielter an. Die Entschließung hob richtigerweise hervor, dass in den Bereichen „komplexe Algorithmensysteme und automatisierte Entscheidungsfindungsprozesse rasch technologische Fortschritte erzielt werden“ und dass „diese Technologien zahlreiche Anwendungen, Chancen und Herausforderungen bieten und praktisch alle Bereiche des Binnenmarkts betreffen“. Darüber hinaus betont der Text, dass „der derzeitige Rechtsrahmen der EU [...] überprüft werden muss, um zu sehen, ob damit auf das Entstehen von KI und automatisierter Entscheidungsfindung reagiert [...] werden kann“.

15 Unter ihnen: Access Now (https://www.accessnow.org/cms/assets/uploads/2020/05/EU-white-paper-consultation_AccessNow_May2020.pdf), AI Now (<https://ainowinstitute.org/ai-now-comments-to-eu-whitepaper-on-ai.pdf>), EDRI (<https://edri.org/can-the-eu-make-ai-trustworthy-no-but-they-can-make-it-just/>) und AlgorithmWatch (<https://algorithmwatch.org/en/response-european-commission-ai-consultation/>).

DAS GESAMTE DOKUMENT HINDURCH WERDEN MIT KI-BASIERTE TECHNOLOGIEN VERBUNDENE RISIKEN GANZ ALLGEMEIN ALS „POTENZIELL“ BEZEICHNET, DIE VORTEILE DAGEGEN ALS SEHR REAL UND UNMITTELBAR DARGESTELLT

Die Entschließung fordert einen „gemeinsamen Ansatz der EU für die Entwicklung automatisierter Entscheidungsfindungsprozesse“ und trifft detaillierte Aussagen zu diversen Bedingungen, die solche Systeme erfüllen sollten, um im Einklang mit europäischen Werten zu stehen. Verbraucher:innen sollten „angemessen darüber informiert werden“, auf welche Weise ihr Leben von Algorithmen beeinflusst wird, und sie sollten Zugang zu Personen mit Entscheidungsbefugnis haben, um Entscheidungen überprüfen und, falls nötig, korrigieren lassen zu können. Außerdem sollten sie darüber informiert werden, „wenn die Preise von Waren oder Dienstleistungen auf der Grundlage automatisierter Entscheidungsfindung und der Erstellung von Profilen des Verbraucherverhaltens personalisiert wurden“.

Die Entschließung erinnert die EU-Kommission daran, dass ein sorgfältig ausgearbeiteter risikobasierter Regulierungsansatz erforderlich ist und betont, Produktsicherheitsvorschriften müssten in Betracht ziehen, dass ADM-Systeme sich „weiterentwickeln und in einer Art und Weise handeln können, die beim ersten Inverkehrbringen nicht vorgesehen ist“ und „dass es problematisch ist, die Haftung in Fällen zu bestimmen, in denen die Schädigung der Verbraucher auf autonome Entscheidungsfindungsprozesse zurückzuführen ist“.

Mit der Feststellung, dass „der Mensch letztlich für Entscheidungen verantwortlich und in der Lage sein muss, sich über Entscheidungen hinwegzusetzen, die im Zusammenhang mit freiberuflichen Dienstleistungen wie den medizinischen, juristischen und Buchhaltungsberufen sowie für den Bankensektor getroffen werden“, wenn „berechtigte

öffentliche Interessen auf dem Spiel stehen“, nimmt die Entschließung Bezug auf [Art. 22 DSGVO](#). Insbesondere sollten „vor der Automatisierung professioneller Dienstleistungen“ die Risiken „ordnungsgemäß“ bewertet werden.

Am Ende listet die Entschließung detaillierte Anforderungen an die Qualität und Transparenz beim Datenmanagement. Unter anderem wird betont, wie wichtig es sei, „nur hochwertige und tendenzfreie Datensätze zu verwenden, um die Leistung algorithmischer Systeme zu verbessern und das Vertrauen und die Akzeptanz der Verbraucher zu stärken“; „verständliche und tendenzfreie Algorithmen“ zu verwenden; „Überprüfungsstrukturen“ zu schaffen, die Verbraucher in die Lage zu versetzen, „eine Überprüfung endgültiger und dauerhafter automatisierter Entscheidungen durch Menschen sowie gegebenenfalls eine Entschädigung zu verlangen“.

/ Das „Initiativrecht“ des EU-Parlaments optimal nutzen

In ihrer Antrittsrede brachte Ursula von der Leyen sehr deutlich ihre Unterstützung eines „Initiativrechts“ für das Europäische Parlament [zum Ausdruck](#). „Wenn das Parlament mehrheitlich Entschließungen annimmt, in denen die Kommission zur Vorlage von Legislativvorschlägen aufgefordert wird, sage ich zu, darauf – unter uneingeschränkter Wahrung der Grundsätze der Subsidiarität, Verhältnismäßigkeit und besserer Rechtsetzung – in Form eines Rechtsakts zu [reagieren](#).“

Sollte es sich bei „KI“ tatsächlich um eine Revolution handeln, die ein entsprechend auf sie abgestimmtes Paket

gesetzlicher Vorschriften erfordert, das mutmaßlich im Laufe des 1. Quartals 2021 kommen wird, dann wollen die gewählten Repräsentant:innen mitreden. Dies, gepaart mit von der Leyens ausdrücklich erklärter Absicht, deren gesetzgeberische Kompetenzen zu stärken, könnte sogar in etwas gipfeln, das Politico als „parlamentarischen Moment“ [bezeichnet](#) hat: in dessen Folge die Parlamentsausschüsse damit beginnen, viele verschiedene Berichte zu schreiben.

Jeder dieser Berichte untersucht spezifische Aspekte der Automatisierung in der öffentlichen Politik, die, ungeachtet der Tatsache, dass sie die kommende „KI“-Gesetzgebung prägen sollen, relevant für ADM sind.

So [fordert](#) etwa der Rechtsausschuss in seinem „Entwurf eines Berichts an die Kommission mit Empfehlungen zu einem Rahmen für die ethischen Aspekte von künstlicher Intelligenz, Robotik und damit zusammenhängenden Technologien“ die Einrichtung einer [„Europäischen Agentur für künstliche Intelligenz“](#) und zugleich ein Netzwerk nationaler Aufsichtsbehörden in den einzelnen Mitgliedsstaaten, um sicherzustellen, dass Entscheidungen mit ethischen Implikationen, bei denen Automatisierung zum Einsatz kommt, ethisch einwandfrei sind und bleiben.

In seinem Berichtsentwurf über die [„Rechte geistigen Eigentums bei der Entwicklung von KI-Technologien“](#) legt derselbe Ausschuss seine Sichtweise in Bezug auf das künftige Verhältnis zwischen [geistigem Eigentum und Automatisierung](#) dar. Zunächst trifft der Berichtsentwurf die Aussage, dass „mathematische Methoden nicht patentierbar sind, es sei denn, es handelt sich um Erfindungen technischer Art“, um zugleich mit Bezug auf algorithmische Transparenz zu behaupten, dass „Reverse Engineering eine Ausnahme von Geschäftsgeheimnissen darstellt“.

Der Bericht geht sogar so weit, Betrachtungen darüber anzustellen, wie der Schutz „der durch KI geschaffenen technischen und künstlerischen Schöpfungen“ sicherzustellen sei, „um diese Form des Schaffens zu fördern“ und „ist der Ansicht, dass bestimmte durch KI erzeugte Werke mit geistigen Werken vergleichbar sind und daher urheberrechtlich geschützt werden könnten“.

Und schließlich legt der Ausschuss in einem dritten [Dokument](#) („Artificial Intelligence and Civil Liability“) einen detaillierten „Risk Management Approach“ für die Haftpflicht in Zusammenhang mit KI-Technologien vor. Demgemäß „haftet als zentraler Einstiegspunkt für Streitverfahren aus-

schließlich die Partei, die am besten in der Lage ist, ein technologiebasiertes Risiko zu überwachen und zu managen“.

Wichtige Prinzipien, die den Einsatz von ADM im Strafrecht betreffen, finden sich im [„Berichtsentwurf über künstliche Intelligenz im Strafrecht und ihre Verwendung durch die Polizei und Justizbehörden in Strafsachen“](#) des Ausschusses für bürgerliche Freiheiten, Justiz und Inneres. Im Anschluss an eine detaillierte Auflistung tatsächlicher, gegenwärtiger Nutzungen von „KI“ – also eigentlich ADM-Systemen – durch die Polizei¹⁶ hält es der Ausschuss „für notwendig, eine klare und faire Regelung für die Zuweisung der rechtlichen Verantwortung für die möglichen nachteiligen Folgen zu schaffen, die durch diese fortgeschrittenen digitalen Technologien verursacht werden“.

Anschließend werden einige Eigenschaften solcher Regelungen beschrieben: keine vollautomatischen Entscheidungen¹⁷, algorithmische Erklärbarkeit, die „für die Nutzer verständlich“ ist sowie „eine obligatorische Bewertung der Auswirkungen auf die Grundrechte [...] vor der Einführung oder dem Einsatz von KI-Systemen für die Strafverfolgung oder die Justiz“, sowie, ergänzend, „eine periodische obligatorische Prüfung aller KI-Systeme, die von Strafverfolgungs- und Justizbehörden verwendet werden, um die algorithmischen Systeme zu prüfen und zu bewerten, sobald diese in Betrieb sind“.

Darüber hinaus fordert der Berichtsentwurf „ein Moratorium für den Einsatz von Gesichtserkennungssystemen für die Strafverfolgung, bis die technischen Standards als vollständig grundrechtskonform angesehen werden können, die erzielten Ergebnisse nicht diskriminierend sind und die Öffentlichkeit Vertrauen in die Notwendigkeit und Verhältnismäßigkeit des Einsatzes solcher Technologien hat“.

16 Auf S. 5 stellt der Berichtsentwurf fest, dass „KI-Anwendungen, die von den Strafverfolgungsbehörden genutzt werden, Anwendungen wie Gesichtserkennungstechnologien, automatische Nummernschilderkennung, Sprecheridentifizierung, Spracherkennung, Lippenlesetechnologien, akustische Überwachung (d.h. Schusserkennungsalgorithmen), autonome Forschung und Analyse identifizierter Datenbanken, Vorhersage (präventive Polizeiarbeit und Kriminalitäts-Hotspot-Analyse), Verhaltenserkennungswerkzeuge, autonome Werkzeuge zur Erkennung von Finanzbetrug und Terrorismusfinanzierung, Überwachung sozialer Medien (Scraping und Data Harvesting zum Aufspüren von Zusammenhängen), IMSI-Catcher und automatisierte Überwachungssysteme mit unterschiedlichen Erkennungsfähigkeiten (wie Herzschlagerkennung und Wärmebildkameras)“ umfassen.

17 „... dass in Gerichts- und Strafverfolgungskontexten die endgültige Entscheidung immer von einem Menschen getroffen werden muss“ (S. 6)

Ziel ist letztlich die Durchsetzung einer größeren allgemeinen Transparenz solcher Systeme, verbunden mit einer Forderung an die Mitgliedsstaaten, für „ein umfassendes Verständnis“ der bei der Strafverfolgung und in den Justizbehörden zum Einsatz kommenden KI-Systemen zu sorgen sowie – nach Maßgabe der Leitlinien eines „[öffentlichen Registers](#)“ – eine detaillierte Beschreibung „der Art der verwendeten Instrumente, der Arten von Straftaten, auf die sie angewendet werden, und der Unternehmen, deren Instrumente eingesetzt werden“.

Während der vorliegende Bericht verfasst wird, [arbeiten](#) der Ausschuss für Kultur und Bildung sowie der Ausschuss für Industrie, Forschung und Energie jeweils an ihren eigenen Berichtsentwürfen.

Alle diese Initiativen mündeten in die [Einsetzung](#) eines „Sonderausschusses zu künstlicher Intelligenz im digitalen Zeitalter“ (AIDA) am 18. Juni 2020. Mit 33 Mitgliedern und einem ersten Mandat für zwölf Monate wird er eine „Analyse der künftigen Auswirkungen“ der künstlichen Intelligenz auf die Wirtschaft in der EU sowie „insbesondere in den Bereichen Kompetenzen, Beschäftigung, Finanztechnologie, Bildung, Gesundheit, Verkehr, Tourismus, Landwirtschaft, Umwelt, Verteidigung, Industrie, Energie und E-Government“ [vornehmen](#).

/ Hochrangige Expertengruppe für KI & KI-Allianz

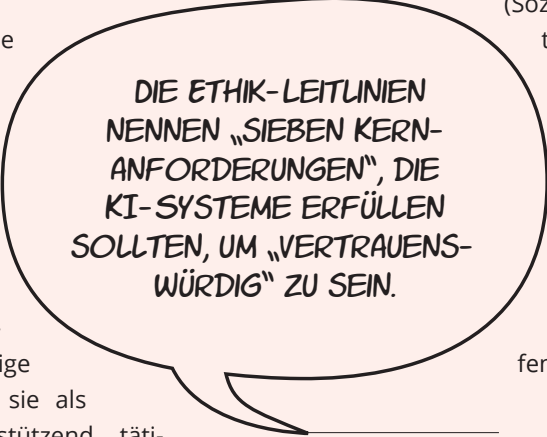
Die EU-Kommission setzte 2018 eine Hochrangige Expertengruppe für Künstliche Intelligenz (High-Level Expert Group on AI, HLEG) ein. Sie besteht aus 52 Expert:innen und soll die Implementierung der europäischen KI-Strategie unterstützen, indem sie Prinzipien identifiziert, die befolgt werden sollten, um „vertrauenswürdige KI“ zu erreichen. Außerdem soll sie als Steuerungsausschuss der unterstützend tätigen KI-Allianz eine offene Plattform für verschiedene Interessenvertreter:innen schaffen (die zum Zeitpunkt der Verfassung dieses Berichts bereits mehr als 4.000 Mitglieder hat), um einen breiteren Input für die Arbeit der Hochrangigen Expertengruppe zu gewährleisten.

Nach Veröffentlichung des ersten Entwurfs zu den Ethik-Leitlinien für eine vertrauenswürdige KI im Dezember 2018

wurde nach Stellungnahme von mehr als 500 Beteiligten im April 2019 eine überarbeitete Fassung vorgelegt. Sie macht sich für einen „menschenzentrierten Ansatz“ stark, um während des gesamten Lebenszyklus des KI-Systems Gesetzeskonformität, Ethik und Robustheit zu erreichen. Dennoch bleibt dies ein freiwilliges Rahmenwerk ohne konkrete, anwendbare Empfehlungen für seine Operationalisierung, Implementierung und Durchsetzung.

Zivilgesellschaftliche sowie Verbraucherschutz- und Menschenrechtsorganisationen haben sich dazu geäußert und verlangt, diese Richtlinien in greifbare Menschenrechte zu überführen¹⁸. So forderte etwa Access Now, eine gemeinnützige Organisation für digitale Rechte und selbst Mitglied der HLEG, die Europäische Kommission dringend zu einer Klarstellung dahingehend auf, wie in einem nächsten Schritt verschiedene Interessengruppen „Vertrauenswürdige KI“ testen, anwenden, verbessern, befürworten und stärken können, und zugleich die Notwendigkeit anzuerkennen, Europas rote Linien festzulegen. In [einem Gastbeitrag](#) schrieben zwei andere Mitglieder der HLEG, die Gruppe hätte „anderthalb Jahre gearbeitet“, um jetzt zu erleben, wie „ihre detaillierten Vorschläge“ von der Europäischen Kommission im Weißbuch „zum größten Teil ignoriert oder nur im Vorbeigehen erwähnt“ worden seien.¹⁹ Außerdem, so argumentierten sie, hätten Mitglieder der Gruppe, die ursprünglich die Aufgabe gehabt habe, Risiken und „rote Linien“ für KI zu identifizieren,

auf autonome Waffensysteme, Citizen-Scoring (Sozialkredit-Systeme) und automatisierte Identifizierung von Einzelpersonen mit Hilfe von Gesichtserkennung hingewiesen und diese als Einsatzgebiete von KI benannt, die es zu vermeiden gelte. Allerdings ist es den Industrievertreter:innen, die den Ausschuss dominieren²⁰, gelungen, diese Prinzipien streichen zu lassen, bevor der Entwurf veröffentlicht wurde.



DIE ETHIK-LEITLINIEN NENNEN „SIEBEN KERNANFORDERUNGEN“, DIE KI-SYSTEME ERFÜLLEN SOLLTEN, UM „VERTRAUENSWÜRDIG“ ZU SEIN.

18 Z.B. der Europäische Verbraucherverband (BEUC): https://www.beuc.eu/publications/beuc-x-2020-049_response_to_the_ecs_white_paper_on_artificial_intelligence.pdf

19 Mark Coeckelbergh und Thomas Metzinger: Mehr Mut zur KI-Ethik, <https://background.tagesspiegel.de/digitalisierung/mehr-mut-zur-ki-ethik>

20 Die Gruppe setzte sich aus 24 Unternehmensvertreter:innen, 17 Akademiker:innen, 5 Organisationen der Zivilgesellschaft und 6 sonstigen Mitgliedern, u. a. der Agentur der Europäischen Union für Grundrechte, zusammen.

Diese Unausgewogenheit, die Potenziale von ADM, lässt sich auch durchgängig im zweiten Ergebnispapier der Expertengruppe feststellen. In ihrem Bericht „[Policy and investment recommendations for trustworthy AI](#)“, veröffentlicht im Juni 2019, finden sich 33 Empfehlungen, die dazu gedacht sind, „Vertrauenswürdige KI in Richtung Nachhaltigkeit, Wachstum, Wettbewerbsfähigkeit und Inklusion zu steuern, während sie gleichzeitig Menschen ermächtigt, schützt und ihnen Nutzen bringt.“ Das Dokument ist im Wesentlichen ein Aufruf zur Ankurbelung der Einführung und Skalierung von KI im privaten und öffentlichen Sektor mit Hilfe von Investitionen in Tools und Anwendungen, um „vulnerable Bevölkerungsgruppen zu unterstützen“ und „niemanden zurückzulassen“.

Ungeachtet dessen und trotz aller legitimer Kritik bringen beide Richtlinien entscheidende Besorgnisse und Forderungen in Bezug auf Systeme zur automatisierten Entscheidungsfindung zum Ausdruck. So [nennen](#) etwa die Ethik-Leitlinien „sieben Kernanforderungen“, die KI-Systeme erfüllen sollten, um „vertrauenswürdig“ zu sein. Anschließend werden die Voraussetzungen aufgelistet, denen die praktische Umsetzung jeder dieser Anforderungen genügen sollte: Vorrang menschlichen Handelns und menschliche Aufsicht; technische Robustheit und Sicherheit; Schutz der Privatsphäre und Datenqualitätsmanagement; Transparenz; Vielfalt, Nichtdiskriminierung und Fairness; gesellschaftliches und ökologisches Wohlergehen; Rechenschaftspflicht. Darüber hinaus ist die Pilotversion einer konkreten „Bewertungsliste für vertrauenswürdige KI“ enthalten, die der praktischen Umsetzung dieser hochrangigen Grundsätze dienen soll. Ziel ist es, dass diese Bewertungsliste immer dann zur Anwendung kommt, „wenn KI-Systeme entwickelt, eingeführt und genutzt werden“, und dass sie „zukünftig auf den spezifischen Anwendungsfall des KI-Systems zugeschnitten“ wird.

Die Liste enthält zahlreiche Aspekte, die mit dem Risiko der Verletzung von Menschenrechten durch ADM-Systeme in Zusammenhang stehen, darunter fehlende menschliche Handlungsmacht und fehlende menschliche Aufsicht; mangelnde technische Robustheit und Sicherheit; die Unfähigkeit, unfaire Verzerrungen zu vermeiden oder einen gleichberechtigten und universellen Zugang zu solchen Systemen zu gewährleisten; sowie das Fehlen eines vernünftigen Zugangs zu den in sie eingespeisten Daten.

Im Zusammenhang bietet die Pilotversion der Bewertungsliste in den Leitlinien nützliche Fragen, die allen helfen können, die ADM-Systeme einzusetzen. So fordert sie zum

**ANGESICHTS DER TATSACHE,
DASS ES FÜR EINE EINZELPERSON
UNMÖGLICH IST, SICH SOLCHEN PROZESSEN
ZU VERWEIGERN, ZUMINDEST NICHT, OHNE
MIT NEGATIVEN KONSEQUENZEN RECHNEN
ZU MÜSSEN, MACHT DER EINSATZ VON
ADM IN REGIERUNG UND VERWALTUNG
VORSICHTSMASSNAHMEN
UND SCHUTZMECHANISMEN
ERFORDERLICH.**

Beispiel „eine Folgenabschätzung für Anwendungsfälle, bei denen die Möglichkeit einer Beeinträchtigung der Grundrechte besteht“. Ebenso enthalten ist die Frage, ob „für den Fall eines selbstlernenden oder autonomen KI-Systems“ „spezifischere Vorkehrungen zur Kontrolle und Aufsicht“ vorhanden sind und „Prozesse zur Gewährleistung der Qualität und Integrität Ihrer Daten“ eingeführt wurden.

Detaillierte Anmerkungen betreffen ebenfalls grundlegende Probleme im Zusammenhang mit ADM-Systemen wie deren Transparenz und Erklärbarkeit. Unter anderem wird gefragt, „inwieweit die Entscheidungen und damit das Ergebnis des KI-Systems nachvollziehbar sind“ und „inwieweit die Entscheidung des Systems die Entscheidungsprozesse der Organisation beeinflusst?“ Diese Fragen sind höchst relevant, um die Risiken, die mit der Einführung solcher Systeme einhergehen, beurteilen zu können.

Um Verzerrungen und diskriminierende Ergebnisse zu vermeiden, lenken die Leitlinien die Aufmerksamkeit auf „Aufsichtsverfahren, [...] mit deren Hilfe der Zweck, die Einschränkungen, Anforderungen und Entscheidungen des Systems klar und transparent analysiert und angegangen werden könnten“ und fordern zugleich eine Beteiligung der Interessenträger während des gesamten Lebenszyklus von KI-Systemen.

Darüber hinaus beinhalten die Politik- und Investitionsempfehlungen die Festlegung roter Linien durch einen institutionalisierten „Dialogs über die Verwendung und die Auswirkungen von KI-Systemen“ mit betroffenen Interessenvertretern, einschließlich Fachleuten der Zivilgesellschaft. Weiterhin mahnen sie dringend ein „Verbot KI-gestützten massenhaften Scorings von Einzelpersonen gemäß der Definition der Ethik-Leitlinien“ an und fordern „sehr klare und strenge Regeln für die Überwachung zu nationalen Sicher-

heitszwecken und anderen Zwecken, die angeblich dem öffentlichen oder nationalen Interesse dienen“. Dieses Verbot würde auch Technologien zur biometrischen Identifikation und zum biometrischen Profiling einschließen.

Das Dokument sagt auch, dass „eine eindeutige Definition, ob, wann und wie KI zur automatisierten Personenerkennung verwendet werden darf [...] für die Schaffung einer vertrauenswürdigen KI in Zukunft entscheidend“ ist und warnt: „Jede Form der Bürgerbewertung kann zum Verlust [von] Autonomie führen und den Grundsatz der Nichtdiskriminierung gefährden“ und sollte daher „nur dann eingesetzt werden, wenn es eindeutig gerechtfertigt ist und die Maßnahmen verhältnismäßig und fair sind“. Weiterhin wird betont, dass „Transparenz weder Diskriminierung verhindern noch Fairness gewährleisten“ kann. Dies bedeutet, dass die Möglichkeit bestehen muss, sich einem solchen Bewertungsverfahren entziehen zu können, und zwar idealerweise so, dass den Betroffenen daraus keine Nachteile entstehen.

Auf der einen Seite heißt es im Dokument, dass es „anzuerkennen und zu berücksichtigen“ gelte, „dass KI-Systeme dem Einzelnen und der Gesellschaft zwar einen erheblichen Nutzen bringen, gleichzeitig jedoch bestimmte Risiken bergen und möglicherweise negative, mitunter schwer absehbare, erkennbare oder messbare Auswirkungen (z.B. im Hinblick auf Demokratie, Rechtsstaatlichkeit, Verteilungsgerechtigkeit oder den menschlichen Geist als solchen) haben können“. Auf der anderen Seite ist die Expertengruppe jedoch der Auffassung, dass „unnötig präskriptive Regulierung vermieden werden sollte“.

Im Juli 2020 stellte die HLEG im Ergebnis eines Pilotprozesses, an dem sich 350 Interessenvertreter:innen beteiligten, außerdem ihre abschließende Bewertungsliste für vertrauenswürdige künstliche Intelligenz vor ([Assessment List for Trustworthy Artificial Intelligence, ALTAI](#)) vor.

Ziel der Checkliste, deren Nutzung vollständig freiwillig ist und die keinerlei regulatorische Konsequenzen hat, ist es, die sieben in den Ethik-Leitlinien zu KI der HLEG dargelegten Kernanforderungen in die Praxis umzusetzen. Im Kern geht es darum, allen, die KI-Lösungen implementieren möchten, die mit den Grundwerten der EU vereinbar sind – etwa Designer:innen und Entwickler:innen von KI, Datenwissenschaftler:innen, Beschaffungsbeamt:innen oder -spezialist:innen sowie juristische Berater:innen/ Compliance-Beauftragte – ein Instrumentarium zur Selbstbewertung zur Verfügung zu stellen.

/ Europarat: Zum Schutz von Menschenrechten bei ADM

Das Ministerkomitee des Europarates²¹ hat ergänzend zur Einrichtung des Ad-hoc-Ausschusses für künstliche Intelligenz (CAHAI) im September 2019 ein substantielles und überzeugendes Rahmenwerk veröffentlicht.

Als standardisierendes Instrument gedacht, beschreibt die „[Recommendation to Member states on the human rights impacts of algorithmic systems](#)“²² „fundamentale Herausforderungen“, die das Aufkommen und unsere „zunehmende Abhängigkeit“ von solchen Systemen mit sich bringen und die relevant sind „für Demokratie und Rechtsstaatlichkeit“.

Das Rahmenwerk, zu dem es eine Phase öffentlicher Konsultationen mit ausführlichen [Beiträgen](#) von Organisationen der Zivilgesellschaft gab, geht im Hinblick auf den Schutz von Grundwerten und Menschenrechten über das Weißbuch der EU-Kommission hinaus.

Die Empfehlung analysiert die Auswirkungen und sich entwickelnden Konfigurationen algorithmischer Systeme (Anhang A) und untersucht dazu alle Phasen des Prozesses, der zur Schaffung eines Algorithmus führt, das heißt Auftragsvergabe, Design, Entwicklung und laufender Einsatz.

Generell nimmt sie den in den Leitlinien der HLEG verfolgten „menschenzentrierten“ KI-Ansatz auf, skizziert aber auch einklagbare „Verpflichtungen der Mitgliedsstaaten“ (Anhang

21 Der Europarat ist zweierlei: „ein Regierungsorgan, in dem die nationalen Ansätze zu Problemen von europäischer Tragweite auf Augenhöhe diskutiert werden, und ein Forum zur Erarbeitung kollektiver Antworten auf diese Herausforderungen“. Zu seinen Arbeitsfeldern gehören „die politischen Aspekte der europäischen Integration, der Schutz demokratischer Institutionen und der Rechtsstaatlichkeit sowie der Menschenrechte – mit anderen Worten, alle Probleme, die einer abgestimmten europaweiten Lösung bedürfen“. Obgleich die Empfehlungen für die Regierungen der einzelnen Mitgliedsstaaten nicht bindend sind, kann das Ministerkomitee in bestimmten Fällen die Regierungen einzelner Mitgliedsstaaten ersuchen, ihm mitzuteilen, was sie auf diese Empfehlungen hin veranlasst haben. (Art. 15b der Satzung). Die Beziehungen zwischen dem Europarat und der Europäischen Union sind niedergelegt im (1) [Compendium of Texts governing the relations between the Council of Europe and the European Union](#) (Sammlung von maßgebenden Texten zur Regelung der Beziehungen zwischen dem Europarat und der Europäischen Union) sowie (2) [Memorandum of Understanding between the Council of Europe and the European Union](#) (Gemeinsame Absichtserklärung zwischen dem Europarat und der Europäischen Union).

22 Unter der Aufsicht des Lenkungsausschusses für Medien und Informationsgesellschaft (CDMSI) und erstellt vom Expertenausschuss für die menschenrechtlichen Dimensionen automatisierter Datenverarbeitung und verschiedener Formen Künstlicher Intelligenz (MSI-AUT)

B) sowie Verantwortlichkeiten der Akteure des privatwirtschaftlichen Sektors (Anhang C). Hinzu kommen Prinzipien wie das der „informationellen Selbstbestimmung“²³ sowie eine Auflistung detaillierter Vorschläge für Rechenschaftslegungsmechanismen und effektive Rechtsbehelfe und die Forderung nach einer Folgenabschätzung für grundlegende Menschenrechte.

Zwar erkennt das Dokument das „signifikante Potenzial digitaler Technologien im Umgang mit gesellschaftlichen Herausforderungen und für gesellschaftlich vorteilhafte Innovationen und wirtschaftliche Entwicklungen“ an, mahnt jedoch gleichzeitig dringend zur Vorsicht. So soll sichergestellt werden, dass diese Systeme weder absichtlich noch zufällig „Ungerechtigkeiten aufgrund von ethnischer Herkunft und Geschlecht sowie andere gesellschaftliche und arbeitsmarktrelevante Ungerechtigkeiten“ fortschreiben, „die unsere Gesellschaften noch immer prägen“.

Im Gegenteil, der Einsatz algorithmischer Systeme sollte auf proaktive und sensible Art und Weise erfolgen, um diese Ungerechtigkeiten anzugehen und „Aufmerksamkeit auf die Bedürfnisse und Stimmen vulnerabler Gruppen zu lenken“.

Am bemerkenswertesten aber ist, dass die Empfehlung das potenziell höhere Risiko für die Verletzung von Menschenrechten aufzeigt, wenn algorithmische Systeme von Mitgliedsstaaten zur Bereitstellung öffentlicher Dienstleistungen und bei politischen Maßnahmen eingesetzt werden. Angesichts der Tatsache, dass es für eine Einzelperson unmöglich ist, sich solchen Prozessen zu verweigern, zumindest nicht, ohne mit negativen Konsequenzen rechnen zu müssen, macht der Einsatz von ADM in Regierung und Verwaltung Vorsichtsmaßnahmen und Schutzmechanismen erforderlich.

Außerdem spricht die Empfehlung die Konflikte und Herausforderungen an, die aus Public-Private-Partnerships bei einer Vielzahl von Anwendungen erwachsen („weder klar staatlich noch klar privat“).

²³ „Staaten sollten sicherstellen, dass alle Prozesse im Zusammenhang mit Design, Entwicklung und kontinuierlicher Einführung algorithmischer Systeme den betroffenen Einzelpersonen Wege eröffnen, im Vorhinein Kenntnis über die damit verbundene vorgesehene Verarbeitung ihrer Daten zu erlangen (einschließlich des Zwecks und der möglichen Ergebnisse) sowie ihre Daten, auch mit Hilfe von Interoperabilität, zu kontrollieren“, heißt es in Anhang B, Abschnitt 2.1.

Zu den Empfehlungen für die Regierungen der Mitgliedsstaaten gehören unter anderem: Beendigung von Prozessen und Verweigerung des Einsatzes von ADM-Systemen, wenn „menschliche Kontrolle und Aufsicht sich als undurchführbar erweisen“ oder Menschenrechte bedroht sind; Einsatz von ADM-Systemen nur dann, wenn „auf allen Ebenen des Prozesses“ Transparenz, Rechenschaftspflicht, Rechtmäßigkeit und der Schutz der Menschenrechte garantiert werden können. Darüber hinaus soll die Überwachung und Bewertung dieser Systeme „kontinuierlich“ erfolgen, „inklusiv und transparent“ sein und sowohl einen Dialog mit allen relevanten Interessengruppen als auch eine Analyse der ökologischen Auswirkungen sowie weiterer potenzieller externer Effekte auf „Bevölkerungen und Umgebungen“ beinhalten.

In Anhang A definiert der Europarat zudem Algorithmen mit „hohem Risiko“, die anderen Gremien als Anregung dienen können. Genauer ist dort ausgeführt, dass „der Begriff „hohes Risiko“ dann Anwendung findet, „wenn es um die Nutzung algorithmischer Systeme in Prozessen oder bei Entscheidungen geht, die ernsthafte Konsequenzen für Einzelpersonen haben können oder in Situationen, wo das Fehlen von Alternativen eine besonders hohe Wahrscheinlichkeit der Verletzung von Menschenrechten mit sich bringt, etwa indem sie Verteilungsgerechtigkeiten erzeugen oder verstärken“.

Das Dokument, zu dessen Annahme es keines einstimmigen Votums der Mitglieder bedurfte, ist nicht bindend.

/ Regulierung terroristischer Online-Inhalte

Nach einer langen Phase, in der es nur schleppend voranging, nahm die Erarbeitung der [„Verordnung des Europäischen Parlaments und des Rates zur Verhinderung der Verbreitung terroristischer Online-Inhalte“](#) 2020 Fahrt auf. Sollte das verabschiedete Regelwerk noch immer automatisierte und proaktive Tools zur Erkennung und Entfernung von Online-Inhalten enthalten, würden diese mit hoher Wahrscheinlichkeit unter Art. 22 der DSGVO fallen.

Der Europäische Datenschutzbeauftragte (EDPS) drückt es so aus: „Da das von dem Vorschlag vorgesehene automatisierte Werkzeug nicht nur zur Entfernung und Vorratsdatenspeicherung von Inhalten (und zugehörigen Daten) im Zusammenhang mit dem Hochlader führen könnte, sondern letztendlich auch zu strafrechtlichen Untersuchungen dieses Hochladers, würden diese Werkzeuge bedeutende

Auswirkungen auf diese Person haben, sich erheblich nachteilig auf ihre Meinungsfreiheit auswirken und bedeutsame Risiken für ihre Rechte und Freiheiten darstellen“ und somit unter Art. 22(2) DSGVO fallen.

Ebenfalls, und zwar unverzichtbar, würde es substanziellere Schutzmechanismen erfordern, als die von der Kommission derzeit vorgesehenen. Das Advocacy Netzwerk European Digital Rights (EDRi) erläutert dazu: „Die vorgeschlagene Verordnung zur Verhinderung der Verbreitung terroristischer Online-Inhalte bedarf einer grundlegenden Überarbeitung, um den Werten der Europäischen Union zu entsprechen und die fundamentalen Rechte und Freiheiten ihrer Bürger:innen zu schützen.“

Ein frühzeitiger Sturm der Kritik von Seiten zivilgesellschaftlicher Gruppen und Ausschüssen des Europäischen Parlaments (EP), darunter Meinungen und Analysen der Europäischen Agentur für Menschenrechte (FRA) und von EDRi, sowie auch ein gemeinsamer kritischer Bericht dreier Sonderberichterstatter:innen der Vereinten Nationen als Reaktion auf den ursprünglichen Vorschlag hob die Gefahren für das Recht auf Meinungs- und Informationsfreiheit, die Freiheit und Pluralität der Medien, die unternehmerische Freiheit sowie des Rechtes auf Schutz der Privatsphäre und des Rechtes zum Schutz personenbezogener Daten hervor.

Kritische Punkte sind unter anderem eine unzureichende Definition terroristischer Inhalte, der Geltungsbereich der Verordnung (zum gegenwärtigen Zeitpunkt sind auch Inhalte für Bildungs- und journalistische Zwecke erfasst), die zuvor bereits genannte Forderung nach „proaktiven Maßnahmen“, das Fehlen effektiver richterlicher Aufsicht, unzureichende Berichtsverpflichtungen für Exekutivbehörden und fehlende Schutzmechanismen für „Fälle, in denen vernünftige Gründe für die Annahme vorliegen, dass grundlegende Rechte betroffen sind“ (EDRi 2019).

Wie der Datenschutzbeauftragte (EDPS) betont, sollten zu solchen „angemessenen Schutzmechanismen“ der Anspruch auf direktes Eingreifen einer Person und das Anrecht auf Erläuterung einer mit automatisierten Mitteln getroffenen Entscheidung gehören (EDRi 2019).

Vorgeschlagene bzw. geforderte Schutzmechanismen fanden zwar Eingang in den Berichtsentwurf des EP zu dem Vorschlag, dennoch bleibt abzuwarten, wer in der letzten Runde vor der finalen Abstimmung den längeren Atem hat. In Trialogen zwischen dem EU-Parlament, der neuen EU-Kommission und dem Europarat (die im Oktober 2019 be-

gannen und hinter verschlossenen Türen stattfinden) sind einem geleakten Dokument zufolge nur noch geringfügige Änderungen möglich.

Aufsicht und Regulierung

/ Erste Entscheidungen zur Konformität von ADM-Systemen mit der DSGVO

„Obwohl es im Verlauf der Verhandlungen zur DSGVO und der Datenschutzrichtlinie für Strafverfolgungsbehörden keine große Debatte um das Thema Gesichtserkennung gab, wurden die gesetzlichen Regelungen so ausgestaltet, dass sie im Laufe der Zeit entsprechend den technologischen Entwicklungen angepasst werden können. [...] Jetzt, da die EU über die Ethik von KI und die Notwendigkeit von Regulierungen diskutiert, ist der Moment gekommen festzustellen, ob Gesichtserkennungstechnologien in einer demokratischen Gesellschaft überhaupt erlaubt werden können. Nur dann, wenn die Antwort auf diese Frage „Ja“ lautet, wenden wir uns der Frage zu, auf welche Weise Schutzmechanismen und Rechenschaftspflichten gesetzlich zu verankern sind.“ – EDPS, Wojciech Wiewiórowski

„Gesichtserkennung ist ein besonders intrusiver biometrischer Mechanismus, der für die betroffenen Personen erhebliche Risiken der Verletzung ihrer Privatsphäre oder ihrer Bürgerrechte mit sich bringt.“ – (CNIL 2019)

Seit dem letzten „Automating Society“-Bericht hat es auf Grundlage der DSGVO die ersten Strafzahlungen und Entscheidungen gegeben, die in Zusammenhang mit einer Verletzung der von den nationalen Datenschutzbehörden (DPAs) erlassenen gesetzlichen Vorschriften standen. Die folgenden Fallstudien zeigen allerdings die praktischen Grenzen der DSGVO auf, wenn es um Art. 22 geht, der sich auf ADM-Systeme bezieht, und wie sie Datenschutzbehörden darauf zurückwirft, Fall-zu-Fall-Bewertungen vorzunehmen.

In Schweden kam man zu dem Schluss, dass ein für einen begrenzten Zeitraum in einer einzigen Schulklasse durchgeführtes Testprojekt zur Gesichtserkennung diverse Verpflichtungen gemäß der Datenschutzgrundverordnung

(insbes. Art. 2(14) und Art. 9 (2)) verletzt. (Europäischer Datenschutzausschuss 2019)

Ein ähnlicher Fall liegt auf Eis, nachdem die französische Commission Nationale de l'Informatique et des Libertés (CNIL) ihre Besorgnis geäußert hatte, als zwei Oberschulen planten, im Rahmen einer Partnerschaft mit dem US-amerikanischen Tech-Konzern Cisco Gesichtserkennungstechnologie einzusetzen. Die Stellungnahme hat keine rechtliche Wirkung, und die eingereichte Klage läuft noch.²⁴

Da die Zustimmung der Nutzer:innen allgemein als ausreichend gilt, um biometrische Daten zu verarbeiten, ist eine Ex-ante-Genehmigung durch Datenschützer nicht erforderlich, um solche Testläufe durchzuführen. In Schweden war das allerdings anders. Der Grund: das bestehende Machtungleichgewicht zwischen dem Daten-Controller und den Datensubjekten. Stattdessen wurden eine angemessene Folgenabschätzung sowie Vorabkonsultationen mit der Datenschutzbehörde für notwendig erachtet.

Der Europäische Datenschutzbeauftragte (EDPS) [bestätigte dies](#):

„Eine Zustimmung muss explizit und aus freien Stücken gegeben werden sowie informiert und bezogen auf den speziellen Anlass erfolgen. Allerdings steht es außer Zweifel, dass Personen ihre Zustimmung nicht verweigern und noch viel weniger ihre Zustimmung erteilen können, wenn sie Zugang zu öffentlichen Orten haben müssen, die von Überwachungsmechanismen zur Gesichtserkennung abgedeckt sind. [...] Und schließlich ist höchst fraglich, ob die Technologie verpflichtenden Prinzipien wie Datenminimierung und Datenschutz durch Technikgestaltung und datenschutzfreundliche Voreinstellungen (data protection by design) folgt. Gesichtserkennungstechnologie hat bisher noch nie vollständig korrekt funktioniert, und dies hat ernsthafte Konsequenzen für alle Personen, die fälschlicherweise identifiziert werden, sei es als Kriminelle oder anderweitig. [...] Allerdings wäre es ein Fehler, den Fokus ausschließlich auf Fragen des Datenschutzes zu richten. Vielmehr handelt es sich hier um eine für eine demokratische Gesellschaft fundamentale ethische Frage.“ (EDPS 2019)

DIE ENTSCHEIDUNG GILT ALS BEDEUTSAMERPRÄZEDENZFALL ZU EINEM HEISS DISKUTIERTEN THEMA, UND DAS URTEIL FAND GROSSE AUFMERKSAMKEIT SOWOHL BEI AKTEUR:INNEN DER ZIVILGESELLSCHAFT WIE AUCH BEI RECHTSGELEHRTEN IN GANZ EUROPA UND DARÜBER HINAUS.

Access Now kommentierte:

„Im Zuge der zunehmenden Entwicklung von Projekten zur Gesichtserkennung sehen wir bereits, dass die DSGVO nützliche Vorschriften zum Schutz von Menschenrechten enthält, die gerichtlich gegen die unrechtmäßige Sammlung und Nutzung sensibler Daten wie etwa biometrischer Daten durchgesetzt werden können. Allerdings könnten der unverantwortliche und häufig unbegründete Hype um die Effizienz solcher Technologien und die dahinterstehenden ökonomischen Interessen dazu führen, dass Zentral- und Lokalregierungen sowie private Unternehmen versuchen, das Gesetz zu umgehen.“

/ Von der Polizei in Südwaales eingesetzte automatisierte Gesichtserkennung für rechtswidrig erklärt

Im Laufe des Jahres 2020 erlebte das Vereinigte Königreich den ersten prominenten Anwendungsfall der Law Enforcement Directive²⁵ (Durchsetzungsrichtlinie) zum Einsatz von Gesichtserkennungstechnologie an öffentlichen Orten durch die Polizei. Die Entscheidung gilt als bedeutsamer Präzedenzfall zu einem heiß diskutierten Thema, und das Urteil fand große Aufmerksamkeit sowohl bei Akteur:innen der Zivilgesellschaft wie auch bei Rechtsgelehrten in ganz Europa und darüber hinaus.²⁶

Der Kläger war Ed Bridges, ein 37 Jahre alter Mann aus Cardiff. Inhalt seiner Klage war der [Vorwurf](#), sein Gesicht sei

²⁴ Siehe Kapitel France in der Gesamtausgabe des Automating Society Reports unter <https://automatingsociety.algorithmwatch.org/report2020/france/> sowie (Kayalki 2019)

²⁵ Die Durchsetzungsrichtlinie, in Kraft seit Mai 2018, „regelt den Umgang mit der Verarbeitung persönlicher Daten durch Datenverantwortliche zum Zwecke der Durchsetzung von Recht und Gesetz – welche/was nicht unter den Geltungsbereich der DSGVO fallen/fällt“. <https://www.dataprotection.ie/en/organisations/law-enforcement-directive>

²⁶ Die Entscheidung wurde am 4. September 2019 vom Obersten Gericht in Cardiff in der Sache Bridges ./. Die Polizei von Südwaales getroffen. (High Court of Justice 2019)

DIE EU BETEILIGT SICH MIT 8.199.387,75 EURO AN DEM PROJEKT ZUR ENTWICKLUNG „VERBESSERTER METHODEN ZUR GRENZÜBERWACHUNG“ ZUR BEKÄMPFUNG IRREGULÄREER MIGRATION.

ohne seine Einwilligung sowohl während der Weihnachtseinkäufe 2017 als auch bei einem friedlichen Protest gegen den Einsatz von Waffen ein Jahr später gescannt worden.

In seinem ersten Urteil hatte das Gericht den Einsatz von Technologie zur automatischen Gesichtserkennung ([Automated Facial Recognition Technology](#) AFR) durch die Polizei von Südwales noch für rechtmäßig und angemessen erklärt. Doch die Bürgerrechtsorganisation „Liberty“ legte Berufung gegen das Urteil ein, und das Berufungsgericht für England und Wales entschied gegen die Abweisung der Klage durch das Oberste Gericht und erklärte die Technologie am 11. August 2020 für illegal.

In der Urteilsbegründung gegen die Polizei von Südwales in drei von fünf Anklagepunkten [führte](#) das Berufungsgericht aus, das bestehende Rahmenwerk zum Einsatz von AFR weise „fundamentale Defizite“ auf und sein Einsatz entspreche nicht dem Grundsatz der „Verhältnismäßigkeit“; zudem war keine angemessene Datenschutzfolgeabschätzung (Data Protection Impact Assessment, DPIA) durchgeführt worden, da mehrere entscheidende Schritte hierfür fehlten.

Das Gericht urteilte allerdings nicht, dass das System diskriminierende Ergebnisse aufgrund von ethnischer Zu-

gehörigkeit oder Geschlecht erbringe, da die Polizei von Südwales keine ausreichenden Belege darüber gesammelt habe, um darüber gerichtlich befinden zu können.²⁷ Unabhängig davon war das Gericht der Auffassung, die folgende beachtenswerte Anmerkung hinzufügen zu müssen: „Da es sich bei AFR um eine neuartige und kontroverse Technologie handelt, hoffen wir, dass alle Polizeibehörden, die planen, sie in der Zukunft einzusetzen, den Wunsch haben mögen, sich dahingehend rückzuversichern, dass alles vernünftigerweise Mögliche unternommen wurde, um sicherzustellen, dass die verwendete Software keine geschlechtsspezifischen oder ethnischen Verzerrungen aufweist.“

Nach dem Urteil [forderte](#) Liberty die Polizei von Südwales und andere Polizeibehörden auf, die Nutzung von Gesichtserkennungstechnologie zu beenden.

²⁷ Die Polizeibehörde behauptete, keinen Zugang zur demografischen Zusammensetzung des für den Test verwendeten Datenpakets für den angewendeten Algorithmus „Neoface“ gehabt zu haben. Das Gericht stellte fest, es bleibe „dennoch die Tatsache, dass sich die SWP zu keinem Zeitpunkt ausreichend versichert hat, und zwar weder selbst noch mit Hilfe einer unabhängigen Überprüfung, dass das Softwareprogramm in diesem Fall keinen inakzeptablen geschlechtsspezifischen oder rassistischen Bias“ habe.

ADM in der Praxis: Grenzverwaltung und Überwachung

Während die EU-Kommission und ihre Interessenvertreterinnen darüber debattierten, ob Gesichtserkennungstechnologien reguliert oder verboten werden sollten, liefen in ganz Europa bereits Pilotprojekte zum testweisen Einsatz solcher Systeme.

Dieser Abschnitt beleuchtet eine fundamentale und häufig übersehene Verbindung, nämlich die zwischen Biometrie und den Grenzmanagementsystemen der EU. Hier zeigt sich deutlich, wie Technologien, die diskriminierende Ergebnisse erbringen können, auf Einzelindividuen – z. B. Migrantinnen – angewandt werden könnten, die bereits am stärksten von Diskriminierung betroffen sind.

/ Gesichtserkennung und die Nutzung biometrischer Daten in Politik und Praxis der EU

Im Laufe des Jahres 2019 zogen Gesichtserkennung und andere biometrische Identifikationstechnologien eine Menge Aufmerksamkeit von Regierungen, der EU, der Zivilgesellschaft sowie Bürgerinnen – und Menschenrechtsorganisationen auf sich, insbesondere in den Bereichen Strafverfolgung und Grenzsicherung.

Die Europäische Kommission beauftragte ein Konsortium staatlicher Agenturen, „den aktuellen Stand des Einsatzes von Gesichtserkennung bei Ermittlungsverfahren in allen EU-Mitgliedsstaaten festzustellen“ um auf einen „möglichen Austausch von Gesichtsdaten“ hinzuarbeiten. Sie beauftragte die Beraterfirma Deloitte mit der Durchführung einer Machbarkeitsstudie zur Ausweitung des Prüm-Systems zur Abfrage von Gesichtsbildern. Das EU-weit eingesetzte System ermöglicht den wechselseitigen automatisierten Zugriff auf DNA-Profile, daktyloskopische Daten und Daten aus nationalen Fahrzeugregistern. Es besteht die Sorge, dass eine europaweite Datenbank mit Gesichtsbildern allgegenwärtige, ungerechtfertigte oder illegale Überwachung zur Folge haben könnte.

/ Grenzenlose Grenzverwaltungssysteme

Wie schon in der vorherigen Ausgabe von Automating Society berichtet, ist die Implementierung von übergreifenden, interoperablen Grenzmanagementsystemen, die auf einen Vorschlag der EU-Kommission von 2013 zurückgehen, bereits europaweit im Gange. Obwohl die angekündigten neuen Systeme (EES, ETIAS²⁸, ECRIS-TCN²⁹) erst 2022 in Betrieb gehen sollen, hat die Verordnung über das Einreise-/Ausreisensystem (EES) Gesichtsbilder als biometrische Identifikatoren eingeführt und die Nutzung von Gesichtserkennungstechnologie zu Verifizierungszwecken erstmals in einem EU-Gesetz verankert.³⁰

Die Änderungen wurden von der Europäischen Agentur für Menschenrechte (FRA) bestätigt: „Es ist zu erwarten, dass in den auf EU-Ebene eingesetzten umfänglichen IT-Systemen, die für Asylverfahren, Migration und Sicherheitszwecke genutzt werden, eine systematischere Einführung der Verarbeitung von Gesichtsbildern stattfinden wird [...], sobald die dafür notwendigen juristischen und technischen Schritte vollzogen sind“.

Ana Maria Ruginis Andrei von der Europäischen Agentur für das Betriebsmanagement von IT-Großsystemen im Raum der Freiheit, der Sicherheit und des Rechts (eu-LISA) zufolge wurde diese erweiterte Interoperabilitätsarchitektur „aufgebaut, um den perfekten Motor für den erfolgreichen Kampf gegen die Bedrohungen der inneren Sicherheit und eine effektive Kontrolle von Migration zu schaffen und blinde Flecken beim Identitätsmanagement zu tilgen“. In der Praxis bedeutet dies die „Vorhaltung der Fingerabdrücke, Gesichtsbilder und anderer persönlicher Daten von bis zu 300 Millionen Nicht-EU-Bürgern durch Zusammenführung

28 ETIAS (EU Travel Information and Authorisation System) ist das von eu-LISA entwickelte neue „Visa Waiver“-System für das EU-Grenzmanagement. „Die bei Antragstellung eingegebenen Informationen werden automatisch mit bereits existierenden EU-Datenbanken (Eurodac, SIS und VIS), künftig auch den Systemen EES und ECRIS-TCN sowie mit den relevanten Interpol-Datenbanken abgeglichen. So wird eine verbesserte Überprüfung im Hinblick auf Sicherheitsrisiken, Risiken im Zusammenhang mit illegaler Einwanderung und Gesundheitsrisiken erreicht.“ (ETIAS 2019)

29 Das Europäische Strafregisterinformationssystem – Drittstaatsangehörige (ECRIS-TCN), das von eu-LISA entwickelt werden soll, wird eine Erweiterung des bereits bestehenden EU-Strafregisterinformationssystems (ECRIS) in Form eines zentralisierten „hit/no-hit“-Systems sein, das Informationen über Nicht-EU-Bürger abrufbar macht, sofern sie in einem EU-Staat vorbestraft sind.

30 EES wird den Betrieb im 1. Quartal 2022 aufnehmen, ETIAS wird bis Ende 2022 folgen. Beide Systeme sollen „im Zuständigkeitsbereich des Rates für Justiz und Inneres (JI-Rat) als Game Changer“ fungieren.

der Daten fünf voneinander getrennter Systeme.“ (Campbell 2020)

/ ETIAS: Automatisierte Screenings bei Grenzübertritt

Das [Europäische Reiseinformations- und Genehmigungssystem](#) (ETIAS), das zum Zeitpunkt der Erstellung dieses Berichts noch immer nicht in Betrieb ist, wird verschiedene Datenbanken nutzen, um Reisende aus Nicht-EU-Staaten (die kein Visum bzw. keinen Visa-Waiver benötigen) vor ihrer Einreise nach Europa einer automatisierten digitalen Sicherheitskontrolle zu unterziehen.

Das System wird künftig Daten für die verbesserte „Identifizierung potenzieller Sicherheitsrisiken oder Risiken irregulärer Einwanderung“ sammeln und analysieren (ETIAS 2020). Ziel ist „die Stärkung von Grenzkontrollen; die Vermeidung bürokratischer Hürden und Verspätungen für Reisende, die an den Grenzen vorstellig werden; und die Sicherstellung einer koordinierten und harmonisierten Risikobewertung von Einreisenden aus Drittstaaten“. (ETIAS 2020)

Ann-Charlotte Nygård, Leiterin der Abteilung Technical Assistance and Capacity Building bei der FRA, sieht bei ETIAS zwei spezielle Risiken: „erstens die Nutzung von Daten, die zu einer unbeabsichtigten Diskriminierung bestimmter Gruppen führen könnte, etwa dann, wenn Antragsteller:innen einer bestimmten ethnischen Gruppe mit einem hohen Immigrationsrisiko angehören; zweitens in Zusammenhang mit einem Sicherheitsrisiko, dass auf der Basis im Herkunftsland erhaltener Vorstrafen bewertet wird. Manche dieser früheren Verurteilungen könnten von Europäern als unangemessen angesehen werden, so zum Beispiel Urteile gegen Angehörige der LGBT-Community in bestimmten Ländern. Um dies zu verhindern, [...] müssen Algorithmen auditiert werden, damit sichergestellt ist, dass sie nicht diskriminierend wirken, und diese Art des Auditing muss interdisziplinäre Expert:innen einbeziehen.“ (Nygård 2019)

/ iBorderCtrl: Gesichtserkennung und Risikoeinstufung an den Grenzen

iBorderCtrl war ein Projekt, an dem Sicherheitsbehörden aus Ungarn, Lettland und Griechenland beteiligt waren und das zum [Ziel](#) hatte „schnellere und genauere Grenzkontrollen für Drittstaatenangehörige beim Grenzeintritt in EU-Mitgliedsstaaten an Landübergängen zu ermöglichen“.

iBorderCtrl nutzte Gesichtserkennungstechnologie, einen Lügendetektor sowie ein Scoring-System. Auf Basis der gewonnenen Daten teilte es menschlichen Grenzbeamten mit, ob es jemanden für gefährlich hielt oder ob es die Berechtigung einer Person zum Grenzübertritt in Zweifel zog.

Das Projekt lief Ende August 2019 aus, und die Ergebnisse – für jedwede potenzielle EU-weite Implementierung des Systems – waren widersprüchlich.

Obwohl „noch definiert werden muss, inwieweit das System oder Teile davon zum Einsatz kommen werden“, sieht die Projekt-Website unter „Ergebnisse“ „die Möglichkeit, die ähnlichen Funktionalitäten des neuen ETIAS-Systems zu integrieren“ und die „Fähigkeiten“ dahingehend zu erweitern, dass „das Grenzübertrittsverfahren dorthin gebracht werden kann, wo die Reisenden sind (Bus, Auto, Zug etc.)“.

Allerdings wurden weder die Module näher spezifiziert, auf die sich dies bezieht, noch die ADM-bezogenen Tools einer öffentlichen Bewertung unterzogen.

Gleichzeitig findet sich in den [FAQ](#) zum Projekt die Bestätigung, dass das getestete System „derzeit für einen Einsatz an der Grenze als nicht geeignet“ angesehen wird, und zwar, „zum einen, weil es sich um einen Prototyp handelt, zum anderen wegen der technologischen Infrastruktur auf EU-Ebene“. Dies bedeutet, dass „für eine Nutzung durch die Grenzbehörden weitere Entwicklungsschritte und eine Integration in die bestehenden EU-Systeme erforderlich“ wären.

Insbesondere – und zwar ungeachtet der Tatsache, dass das iBorderCtrl-Konsortium in der Lage war, die Funktionsfähigkeit einer solchen Technologie für Grenzkontrollen im Prinzip aufzuzeigen – ist ebenso klar, dass vor jedem tatsächlichen Einsatz ethischen, rechtlichen und gesellschaftlichen Auflagen Rechnung getragen werden muss.

/ Vergleichbare Horizon2020-Projekte

Im Rahmen des Programms Horizon2020 nahmen diverse Anschlussprojekte die Testung und Entwicklung neuer Systeme und Technologien für Grenzmanagement und -kontrolle in den Fokus. Diese sind auf der Webseite des Forschungs- und Entwicklungsinformationsdienstes der Europäischen Gemeinschaft (CORDIS) aufgelistet, wo Informationen zu allen EU-finanzierten Forschungsaktivitäten in diesem Zusammenhang vorgehalten werden.

Die Seite [verzeichnet](#) derzeit unter dem Programm/Thema „H2020-EU.3.7.3. – Strengthen security through border management“ (Verstärkte Sicherheit durch Grenzmanagement) der Europäischen Union 38 laufende Projekte. Das übergeordnete Programm „Sichere Gesellschaften – Schutz der Freiheit und Sicherheit Europas und seiner Bürger“, verfügt über ein Gesamtbudget von knapp 1,7 Milliarden Euro und finanziert 350 Projekte. Es bekämpft nach eigener Aussage „Unsicherheit, egal, ob sie nun aus Verbrechen, Gewalt, Terrorismus, Natur- oder menschengemachten Katastrophen, Cyber-Attacken, Verletzungen der Privatsphäre oder anderen Formen gesellschaftlicher und ökonomischer Verwerfungen erwächst, von denen Bürgerinnen und Bürger zunehmend betroffen sind“, und zwar mit Hilfe von Projekten, die vorrangig der Entwicklung neuer, auf KI und ADM basierender technologischer Systeme dienen.

Einige Projekte sind bereits beendet worden und/oder ihre Anwendungen sind schon im Einsatz, so zum Beispiel FastPass, ABC4EU, MOBILEPASS und EFFISEC – sie alle dienen der Untersuchung von Anforderungen an „integrierte, interoperable, automatisierte Grenzkontrollen (ABC)“, Identifikationssysteme und „smarte“ Schleusen an diversen Grenzübergängen.

TRESSPASS ist ein laufendes Projekt, das im Juni 2018 gestartet wurde und im November 2021 beendet sein wird. Die EU ist mit einem Beitrag in Höhe von knapp zehn Millionen Euro an dem Projekt beteiligt. Ziel der Koordinatorinnen von iBorderCRL (wie auch von FLYSEC und XP-DITE) ist es, die von iBorderCRL „implementierten und getesteten Ergebnisse und Konzepte wirksam einzusetzen“ und „sie innerhalb eines starken gesetzlichen und ethischen Rahmenwerks“ zu einer „multimodalen, risikobasierten Grenzübertrittslösung [zu] erweitern.“ (Horizon2020 2019)

Das Projekt ist nach eigener Aussage darauf ausgerichtet, die alte, unmoderne „regelbasierte“ Strategie für Sicherheitskontrollen an Grenzübergangspunkten in eine neue „risikobasierte“ umzuwandeln. Dies schließt die Anwendung biometrischer und sensorischer Technologien sowie ein risikobasiertes Managementsystem und sachbezogene Modelle für die Bewertung von Identität, Hab und Gut, Fähigkeiten und Absichten ein. Es zielt darauf ab, mit Hilfe von „Links zu Altsystemen und externen Datenbanken wie VIS/SIS/PNR“ Kontrollen zu ermöglichen und sammelt zu Sicherheitszwecken Daten aus all den vorgenannten Datenquellen.

Ein anderes Pilotprojekt ist FOLDOUT. Es startete im September 2018 und wird im Februar 2022 beendet sein. Der Beitrag der EU zu dem Projekt, das sich mit der Entwicklung „verbesserter Methoden zur Grenzüberwachung“ befasst, um illegaler Einwanderung entgegenzuwirken, und dabei einen speziellen Fokus auf „das Aufspüren von Menschen in dichter Vegetation in extremen Klimaverhältnissen“ legt, beläuft sich auf 8.199.387,75 Euro. Durch Kombination „verschiedener Sensoren und Technologien sowie einer intelligenten Einbindung derselben in eine effektive und robuste Detektionsplattform“ sollen Reaktionsszenarien vorgeschlagen werden. Pilotversuche dazu laufen in Bulgarien, flankiert von Demonstrationsmodellen in Griechenland, Finnland und Französisch-Guayana.

MIRROR (Migrationsrisiken aufgrund von Missverständnissen im Hinblick auf Chancen und Anforderungen) startete Anfang Juni 2019 und wird Ende Mai 2022 beendet sein. Der Beitrag der EU zu dem Projekt, das darauf abzielt zu „verstehen, wie Europa im Ausland wahrgenommen wird; Diskrepanzen zwischen Vorstellung und Realität aufzudecken; Fälle von gezielter Manipulation durch Medien aufzuspüren; und Fähigkeiten zu entwickeln, solchen Missverständnissen und den daraus erwachsenden Sicherheitsbedrohungen entgegenzuwirken“, beläuft sich auf rund 5,2 Millionen Euro. Basierend auf „wahrnehmungsspezifischer Bedrohungsanalyse, wird das MIRROR-Projekt Methoden der automatisierten Text-, Multimedia- und soziale Netzwerk-Analyse für verschiedene Arten von Medien (einschließlich sozialer Medien) mit empirischen Studien verknüpfen“, um „Technologien“ zu entwickeln und „handlungsorientierte Erkenntnisse“ zu gewinnen, „die [...] von Grenzbehörden und politischen Entscheidungsträgern sorgfältig validiert sind, z. B. über Pilotprojekte“.

Zu weiteren, bereits abgeschlossenen Projekten, die jedoch Erwähnung finden sollen, gehört Trusted Biometrics under Spoofing Attacks (TABULA RASA), das im November 2010 startete und im April 2014 beendet wurde. Es analysierte „die Schwächen von Software für biometrische Identifikationsprozesse in Bezug auf ihre Anfälligkeit für Spoofing, die zu einer verminderten Effizienz biometrischer Geräte führt“. Ein weiteres Projekt, Bodega, das im Juni 2015 startete und im Oktober 2018 beendet wurde, untersuchte, auf welche Weise sich „menschlicher Sachverstand“ nutzen lässt, wenn es um die „Einführung smarterer Grenzkontrollsysteme wie etwa biometriebasierte automatisierte Schleusen und Selbstbedienungssysteme“ geht.

Quellen:

Access Now (2019): Comments on the draft recommendation of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems <https://www.accessnow.org/cms/assets/uploads/2019/10/Submission-on-CoE-recommendation-on-the-human-rights-impacts-of-algorithmic-systems-21.pdf>

AlgorithmWatch (2020): Our response to the European Commission's consultation on AI <https://algorithmwatch.org/en/response-european-commission-ai-consultation/>

Campbell, Zach/Jones, Chris (2020): Leaked Reports Show EU Police Are Planning a Pan-European Network of Facial Recognition Databases <https://theintercept.com/2020/02/21/eu-facial-recognition-database/>

CNIL (2019): French privacy regulator finds facial recognition gates in schools illegal <https://www.biometricupdate.com/201910/french-privacy-regulator-finds-facial-recognition-gates-in-schools-illegal>

Coeckelbergh, Mark / Metzinger, Thomas(2020): Europe needs more guts when it comes to AI ethics <https://background.tagesspiegel.de/digitalisierung/europe-needs-more-guts-when-it-comes-to-ai-ethics>

Committee of Ministers (2020): Recommendation CM/Rec(2020)1 of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems https://search.coe.int/cm/pages/result_details.aspx?objectId=09000016809e1154

Commissioner for Human Rights (2020): Unboxing artificial intelligence: 10 steps to protect human rights <https://www.coe.int/en/web/commissioner/-/unboxing-artificial-intelligence-10-steps-to-protect-human-rights>

Committee on Legal Affairs (2020): Draft Report: With recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies https://www.europarl.europa.eu/doceo/document/JURI-PR-650508_EN.pdf

Committee on Legal Affairs (2020): Artificial Intelligence and Civil Liability [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU\(2020\)621926_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf)

Committee on Legal Affairs (2020): Draft Report: On intellectual property rights for the development of artificial intelligence technologies https://www.europarl.europa.eu/doceo/document/JURI-PR-650527_EN.pdf

Committee on Civil Liberties, Justice and Home Affairs (2020): Draft Report: On artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters https://www.europarl.europa.eu/doceo/document/LIBE-PR-652625_EN.pdf

Delcker, Janosch(2020): Decoded: Drawing the battle lines – Ghost work – Parliament's moment https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-drawing-the-battle-lines-ghost-work-parliaments-moment/?utm_source=POLITICO.EU&utm_campaign=5a7d137f82-EMAILCAMPAIN_2020_09_09_08_59&utm_medium=email&utm_term=0_10959edeb55a7d137f82-190607820

Data Protection Commission(2020): Law enforcement directive <https://www.dataprotection.ie/en/organisations/law-enforcement-directive>

EDRi (2019): FRA and EDPS: Terrorist Content Regulation requires improvement for fundamental rights <https://edri.org/our-work/fra-edps-terrorist-content-regulation-fundamental-rights-terreg/>

GDPR (Art 22): Automated individual decision-making, including profiling <https://gdpr-info.eu/art-22-gdpr/>

European Commission (2018): White paper: On Artificial Intelligence – A European approach to excellence and trust https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

European Commission (2020): A European data strategy https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

European Commission (2020): Shaping Europe's digital future – Questions and Answers https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_264

European Commission (2020): White Paper on Artificial Intelligence: Public consultation towards a European approach for excellence and trust <https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence>

European Commission (2018): Security Union: A European Travel Information and Authorisation

System – Questions & Answers
https://ec.europa.eu/commission/presscorner/detail/en/MEMO_18_4362

European Data Protection Board (2019): Facial recognition in school renders Sweden's first GDPR fine
https://edpb.europa.eu/news/national-news/2019/facial-recognition-school-renders-swedens-first-gdpr-fine_en

European Parliament (2020): Artificial intelligence: EU must ensure a fair and safe use for consumers
<https://www.europarl.europa.eu/news/en/press-room/20200120IPR70622/artificial-intelligence-eu-must-ensure-a-fair-and-safe-use-for-consumers>

European Parliament (2020): On automated decision-making processes: ensuring consumer protection and free movement of goods and services
https://www.europarl.europa.eu/doceo/document/B-9-2020-0094_EN.pdf

European Data Protection Supervisor (2019): Facial recognition: A solution in search of a problem?
https://edps.europa.eu/press-publications/press-news/blog/facial-recognition-solution-search-problem_de
 ETIAS (2020): European Travel Information and Authorisation System (ETIAS)
https://ec.europa.eu/home-affairs/what-we-do/policies/borders-and-visas/smart-borders/etias_en

ETIAS (2019): European Travel Information and Authorisation System (ETIAS)
<https://www.eulisa.europa.eu/Publications/Information%20Material/Leaflet%20ETIAS.pdf>

Horizon2020 (2019): robuT Risk based Screening and alert System for PASSengers and luggage
<https://cordis.europa.eu/project/id/787120/reporting>

High Court of Justice (2019): Bridges v. the South Wales Police
<https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf>

High-Level Expert Group on Artificial Intelligence (2020): Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment
<https://ec.europa.eu/digital-single-market/en/news/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment>

Hunton Andrew Kurth (2020): UK Court of Appeal Finds Automated Facial Recognition Technology Unlawful in Bridges v South Wales Police
<https://www.huntonprivacyblog.com/2020/08/12/uk-court-of-appeal-finds-automated-facial-recognition-technology-unlawful-in-bridges-v-south-wales-police/>

Kayalki, Laura (2019): French privacy watchdog says facial recognition trial in high schools is illegal
<https://www.politico.eu/article/french-privacy-watchdog-says-facial-recognition-trial-in-high-schools-is-illegal-privacy/>

Kayser-Bril, Nicolas (2020): EU Commission publishes white paper on AI regulation 20 days before schedule, forgets regulation
<https://algorithmwatch.org/en/story/ai-white-paper/>

Leyen, Ursula von der (2019): A Union that strives for more – My agenda for Europe
https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf

Leyen, Ursula von der (2020): Paving the road to a technologically sovereign Europe
<https://delano.lu/d/detail/news/paving-road-technologically-sovereign-europe/209497>

Leyen, Ursula von der (2020): Shaping Europe's digital future
https://twitter.com/eu_commission/status/1230216379002970112?s=11

Leyen, Ursula von der (2019): Opening Statement in the European Parliament Plenary Session by Ursula von der Leyen, Candidate for President of the European Commission
https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_19_4230

Nygård, (2019): The New Information Architecture as a Driver for Efficiency and Effectiveness in Internal Security
<https://www.eulisa.europa.eu/Publications/Reports/eu-LISA%20Annual%20Conference%20Report%202019.pdf>

Sabbagh, Dan (2020): This article is more than 1 month old South Wales police lose landmark facial recognition case
<https://www.theguardian.com/technology/2020/aug/11/south-wales-police-lose-landmark-facial-recognition-case>

South Wales Police (2020): Automated Facial Recognition
<https://afr.south-wales.police.uk/>

Valero, Jorge (2020): Vestager: Facial recognition tech breaches EU data protection rules
<https://www.euractiv.com/section/digital/news/vestager-facial-recognition-tech-breaches-eu-data-protection-rules/>



DEUTSCHLAND
STORY
SEITE 38
FORSCHUNG
SEITE 43



Erfahren Sie mehr auf Seite 44 unter „Verhinderung von Häftlingssuiziden“

Kann KI beim Kampf gegen den **Klima-** **wandel** helfen? Nicht wirklich.

Diverse Institutionen haben behauptet, KI könne einen Beitrag zur Bekämpfung des Klimawandels leisten, doch die Beweise hierfür sind dürftig. Im Gegenteil: KI trägt nachweislich dazu bei, dass mehr Treibhausgase ausgestoßen werden.

Von Nicolas Kayser-Bril

„Künstliche Intelligenz (KI) wird unser Leben verändern, indem sie ... zum Klimaschutz und zur Anpassung an den Klimawandel beiträgt.“ Mit diesen Worten **beginnt** das von der EU-Kommission am 19. Februar 2020 veröffentlichte Weißbuch Künstliche Intelligenz. Auch die Datenethikkommission **bescheinigt** Künstlicher Intelligenz in ihrem Gutachten von 23. Oktober 2019 „große Potenziale ... für Maßnahmen für einen wirksamen Klimaschutz.“

Von diesen kühnen Aussagen inspiriert, machte ich mich auf die Suche nach Beispielen, welchen Beitrag die KI zur Bekämpfung des Klimawandels leistet. Deutschland, wo AlgorithmWatch seinen Sitz hat, war das perfekte Testfeld.

Deutschland ist ein Land der Gegensätze. Einerseits ist es die Heimat der populärsten grünen Partei in Europa. Auf der anderen Seite unterstützt die politische Klasse die größten Autohersteller Europas, obwohl sich einige von ihnen schuldig bekannt haben, Abgastests zu betrügen und aktiv die Gesundheit von Millionen von Deutschen zu schädigen. In Deutschland gibt es mehr Windparks als in jedem anderen europäischen Land, und zugleich ist es für die höchsten CO₂-Emissionen auf dem Kontinent verantwortlich (der Pro-Kopf-Ausstoß liegt größtenteils ebenfalls über dem europäischen Durchschnitt).

Sollte KI in der Tat imstande sein, den Status quo in die eine oder andere Richtung zu beeinflussen, dann müsste sich das zuerst in Deutschland zeigen.

/ Intelligente Waschmaschinen


Zunächst las ich das 500 Seiten starke **Gutachten** des Wissenschaftlichen Beirats der Bundesregierung Globale Umweltveränderungen (WBGU). Dort wird betont, dass jede Strategie zur Bekämpfung des Klimawandels ein neues Nachdenken über die moralischen Werte erforderlich mache, auf denen die Marktwirtschaft beruhe. Welche genau und wie, wurde allerdings weit weniger klar benannt. Neue Technologien, so die Autor:innen, könnten sowohl CO₂-Sparer („low-carbon enabler“) als auch Energieschleudern („power drainer“) sein. Und tatsächlich: Datenzentren allein verbrauchten 2014 etwa 1 % der weltweit erzeugten Ener-

gie und produzierten 0,3 % aller CO₂-Emissionen. Beide Werte dürften heute sogar noch höher liegen. Automatisierte Heimanwendungen könnten den Energieverbrauch senken, doch die Herstellung der Geräte hinterlässt einen großen CO₂-Fußabdruck.

In der Hoffnung, konkretere Beispiele dafür zu finden, auf welche Weise KI zum Klimaschutz beitragen kann, traf ich mich mit zwei den Autor:innen des Berichts nahestehenden Akademiker:innen. Nach einem einstündigen Gespräch in ihren glanzvollen Büroräumen in der Mitte Berlins hatten sich meine Hoffnungen in Nichts aufgelöst. Zu den wenigen Beispielen, die sie mir nennen konnten, gehörte – halb im Scherz – die intelligente Waschmaschine. Ein KI-gesteuertes Gerät, das automatisch darauf programmiert werden

könnte, nachts zu laufen. Auf diese Weise ließe sich der Stromverbrauch gleichmäßiger über den Tag verteilen, was wiederum eine effizientere Energieproduktion ermöglichen würde.

Ich bin kein Wissenschaftler, aber meine Erfahrung mit Waschmaschinen hat gezeigt, dass die Geräte bereits mit einem Timer ausgestattet sind. Und der Grund, warum wir sie nachts nicht laufen lassen, liegt darin, dass sie sehr laut sind und die Nachbarn stören.



„KÜNSTLICHE INTELLIGENZ (KI) WIRD UNSER LEBEN VERÄNDERN, INDEM SIE ... ZUM KLIMASCHUTZ UND ZUR ANPASSUNG AN DEN KLIMAWANDEL BEITRÄGT.“
EUROPÄISCHE KOMMISSION

/ Schwammige Grundlagen

Wenn also die Wissenschaft nicht helfen konnte, dann vielleicht die Zivilgesellschaft. In Deutschland existiert die größte Regionalgruppe von Fridays For Future, einer weltweiten Bewegung von Schüler:innen und Student:innen, die die Umsetzung des Pariser Klimaschutzabkommens fordern. Ich fragte Annika Rittmann, eine der Sprecher:innen der Bewegung, ob KI ihres Wissens jemals zu einer Nettoerduktion von Treibhausgasen geführt hätte.

„Wir arbeiten auf der Grundlage wissenschaftlicher Erkenntnisse, die in diesem Bereich noch sehr schwammig sind“, schrieb sie in ihrer Antwort. „Dementsprechend können wir zu dem Thema nicht viel sagen.“

Ich kam einer Antwort nicht näher; klar schien allerdings zu sein, dass sowohl die EU-Kommission als auch die Datenethikkommission mit ihren Aussagen womöglich etwas zu enthusiastisch gewesen waren.

„KI IST EIN WERKZEUG, KEIN SELBSTZWECK.“

SOFTWAREENTWICKLER
BEI NEXT KRAFTWERKE

/ Windturbinen

Eine andere zivilgesellschaftliche Organisation, GermanWatch (keine Beziehung zu AlgorithmWatch), veröffentlichte 2019 ein [Hintergrundpapier](#) zu den Chancen und Risiken von KI im Zusammenhang mit der *Energiewende*. Dieses Konzept bildet die Grundlage für die Neuordnung des Energiesektors in Deutschland und bedeutet konkret, dass der Preis, den die deutschen Verbraucher:innen im Austausch für das Versprechen, die Energieerzeugung nachhaltig zu gestalten, für ihren Strom zahlen, ein Drittel über dem EU-Durchschnitt liegt.

Endlich fand ich konkrete Beispiele. BirdVision etwa ist ein Tochterunternehmen von Bürgerwindpark Hohenlohe. Das Unternehmen, das neun Windparks in Süddeutschland betreibt, setzt Bilderkennungsoftware ein, die Vögel entdeckt und automatisch die Geschwindigkeit der Turbinen reduziert, wenn sie zu nahe herankommen. Acht Turbinen sind mit diesem System ausgerüstet, und obwohl es sich noch in der Entwicklung befindet, haben laut Aussage des Unternehmens potenzielle Nutzer:innen bereits Interesse bekundet und wollen es sich anschauen, sobald es auf den Markt kommt. Dies, so GermanWatch, könnte zu einer größeren Akzeptanz der Windenergie führen und den Bau weiterer Windkraftanlagen befördern.

Am vielversprechendsten ist der Einsatz von KI dem Bericht zufolge im Bereich Energieeffizienz. Smarte Geräte könnten zum Beispiel so programmiert werden, dass sie nur dann laufen, wenn Strom aus erneuerbaren Quellen in großer Menge verfügbar ist. Und was die Verteilung betrifft, so könnten mit Hilfe von „Smart Grids“, intelligenten Stromnetzen, Produktion und Verbrauch besser aufeinander abgestimmt werden und helfen, den Verbrauch in Spitzenzeiten zu verringern.

/ Kein Selbstzweck

Inzwischen wusste ich, dass sich die Verfasser:innen von Berichten hinreißen lassen können. Also befragte ich ein Dutzend Unternehmen, die auf dem Gebiet der Erneuerbaren Energien tätig sind, wie sie KI einsetzen.

Der Stromhändler Next Kraftwerke nutzt maschinelles Lernen, um das Produktionsniveau von Solarfarmen und die Verbrauchsmengen seiner Kunden zu prognostizieren. Diese Informationen werden sowohl von der Verkaufsabteilung als auch von dem für die Netzstabilität verantwortlichen Personal genutzt. Andere Projekte seien allerdings weniger erfolgreich, so ein Softwareentwickler des Unternehmens (So hatten sie etwa versucht, ein Modell des Strommarktes zu erstellen). „KI ist eine Toolbox“, meinte er. „Next Kraftwerke sieht in KI nüchtern ein Werkzeug und keinen Selbstzweck.“

Andere, wie etwa Naturstrom AG, ein Stromerzeuger und -händler mit mehr als 250.000 Kunden, oder VSB, ein Hersteller schlüsselfertiger Windkraftanlagen, der bisher mehr als 650 Windturbinen und knapp 60 Solarfarmen errichtet hat, sagten mir, bei ihnen käme KI überhaupt nicht zum Einsatz.

/ Das Effizienzproblem

Um die Klimakrise zu bewältigen ist es erforderlich, die Anzahl von klimaschädlichen Treibhausgasen in der Atmosphäre zu reduzieren. Entweder, indem man sie einfängt – eine überaus teure und unausgereifte Technologie – oder, indem man ihren Ausstoß verhindert.

Klar ist, dass KI die Energieeffizienz steigern kann. Die Frage, ob Energieeffizienz irgendeinen Beitrag zur Bekämpfung des Klimawandels leistet, ist jedoch nach wie vor unbeantwortet. Nehmen Sie zum Beispiel Glühbirnen. Dank der LED-Technologie lässt sich Licht ohne Hitzeentwicklung erzeugen, so dass eine 15-Watt-LED-Glühbirne

einer 120-Watt-Wolfram-Glühbirne entspricht – also achtmal effizienter ist. Dennoch bleibt unklar, ob dies zu einem geringeren Energieverbrauch führt, denn billige LEDs eröffnen den Weg zu einer Fülle von Innovationen, darunter brillanteren, interaktiven Werbebildschirmen. Jeder einzelne von ihnen braucht mehr als 2.000 Watt, das ergibt in der Summe etwa das Sechsfache des jährlichen Energieverbrauchs eines Haushalts.

Ein grundlegendes Dogma in Marktwirtschaften lautet, dass ein geringerer Verbrauch für einige in geringeren Energiekosten führt, zugleich aber in einem höheren Verbrauch für die anderen bedeutet. Deshalb warnte der Wissenschaftliche Beirat der Bundesregierung Globale Umweltveränderungen (WBGU), dass sich zuerst die Regeln ändern müssten, bevor man über technische Lösungen diskutiere.

/ Feuermelder

In den Jahren 2018 und 2019 gingen in Deutschland über 2.000 Hektar Wald in Flammen auf. Wenig im Vergleich zu Südeuropa (in Portugal verbrannten 2018 allein 43.000 Hektar Wald), doch viel für Deutschland. Lediglich Schweden und Lettland verzeichneten im Vergleich zum Durchschnitt der Jahre 2008-2017 größere Steigerungen.

Angesichts der Tatsache, dass jeder Hektar verbranntes Land zwischen 15 und 35 Tonnen CO₂ freisetzt, waren Waldbrände im Jahr 2019 für mehr als 40.000 Tonnen CO₂-Emissionen in Deutschland verantwortlich. Das ist weniger als ein Hundertstel von einem Hundertstel von einem Prozent aller Treibhausgase, die in dem betreffenden Jahr in Deutschland freigesetzt wurden, dennoch eine erhebliche Menge. Allerdings eine ohne Rebound-Effekt: Die Verhinderung eines einzigen Waldbrandes ist ein echter Beitrag zum Klimaschutz, kein potenzieller.

KI kann in mehrerlei Hinsicht einen Beitrag zum Kampf gegen Waldbrände leisten: von einer Vorhersage möglicher Brandgebiete bis hin zur Entdeckung von Waldbränden mittels Satellitenaufnahmen. Diese Technologie ist so vielversprechend, dass das US-Magazin *Fast Company* vor zwei Jahren einen Artikel mit dem Titel „[Future wildfires will be fought with algorithms](#)“ (Waldbrände werden künftig von Algorithmen bekämpft) brachte.

Wieder war ich voller Hoffnung, dass es mir gelingen würde, ein Beispiel für einen positiven Nettoeffekt von KI auf das Klima zu finden. Ich schrieb E-Mails an die Notrufzentralen und Forstverwaltungen jener Regionen in Deutschland, die

am heftigsten von Waldbränden betroffen sind (Brandenburg, Sachsen und Sachsen-Anhalt). Ich war neugierig zu erfahren, wie sie KI bereits nutzten oder ob sie deren Nutzung planten. Und jetzt kommt's: Sie tun es nicht! Nicht, dass sie technologiescheu wären. Natürlich nutzen sie Sensoren und Software, um Waldbrände zu lokalisieren, wie zum Beispiel automatische Rauchmelder, Satellitenbilder, Drohnen und Infrarotkameras, doch eben nichts, bei dem Künstliche Intelligenz zum Einsatz kommt. (Das Bundesland Brandenburg schließt einen Einsatz von KI nicht generell aus, doch 2015 zeigte eine [Dissertation](#), dass das bisher verwendete System FireWatch alle Maschinenlernmodelle ausstach.)

/ KI statt Öl

Ich war gerade dabei, mir einen weiteren Bereich genauer anzusehen, in dem KI möglicherweise zum Einsatz kommen könnte, um den Klimawandel aufzuhalten. Doch Greenpeace funkte mir dazwischen. Die Umweltschutzorganisation veröffentlichte am 19. Mai 2020 einen Bericht mit dem Titel [Oil in the Cloud](#). Darin untersuchten sie, auf welche Weise Energiekonzerne KI nutzen, um die Fördermengen bei Öl und Gas zu steigern – ein Nettobeitrag zur Klimakatastrophe.

Obwohl sich der Bericht auf die USA konzentrierte, enthüllte er, dass Amazon, Microsoft und, in etwas geringerem Ausmaß, Google die fossile Brennstoffindustrie ermunterten, ihre Dienstleistungen zu nutzen. KI war nicht bloß geeignet, ihnen zu helfen, schneller und mehr fossile Brennstoffe zu finden, sondern bereits zu diesem Zweck im Einsatz.

Microsoft merkte in einem Fall an, seine Technologien könnten für eine Steigerung der Fördermenge um „bis zu 50.000 Barrel Öl-Äquivalente pro Tag“ sorgen. Verfeuert in einem Verbrennungsmotor oder einem Kraftwerk, entspricht das mehr als 20.000 Tonnen CO₂, die in die Atmosphäre gelangen – etwa der Hälfte dessen, was durch die Waldbrände in Deutschland freigesetzt wurde. *Täglich*, für ein einziges Projekt.

/ Zur Quelle vordringen

Jeder Effizienzgewinn, den KI den Betreibern im Bereich Erneuerbare Energien bietet, ist auch für diejenigen relevant, die fossile Energieträger verbrennen. Da fossile Energieunternehmen weitaus größer sind und sehr viel mehr Macht besitzen, kann man fairerweise sagen, dass KI zu einem vermehrten Ausstoß von Treibhausgasen führt und nicht zu einem verringerten. Solange KI nicht nachweislich

bei Projekten zum Einsatz kommt, die zu einer echten Netto-reduzierung der CO₂-Konzentrationen führen, wie etwa Carbon Capture and Storage (das Einfangen und Speichern von Kohlendioxid) bleibt die Behauptung, KI trüge zur Bekämpfung des Klimawandels bei, bestenfalls ein frommer Wunsch.

Am Ende drang ich zu den Quellen vor, auf die sich die EU-Kommission beruft, wenn sie zu dem Schluss kommt, dass KI unser Leben verändern wird, indem sie einen Beitrag zum Klimaschutz und zur Klimaanpassung leistet. Womöglich hatten sie ja Zugang zu Dokumenten und Expert:innen, die mir entgangen waren.

Eine Quelle, die sich in der Materie auskannte, schickte mir drei Links. Der erste führte zu einem Listicle mit dem Titel „8 ways AI can help save the planet“ (8 Möglichkeiten, wie KI helfen kann, unseren Planeten zu retten), verfasst von einer Mitarbeiterin des global agierenden Wirtschaftsprüfungunternehmens PricewaterhouseCoopers (PwC). Er enthielt keine Referenzen und war vom Weltwirtschaftsforum

veröffentlicht worden, einer Gruppe von Geschäftsleuten, die bekannt ist für ihre alljährlichen Treffen in Davos. Der zweite führte zu einem [Blog-Beitrag](#), der in seiner Gänze auf einem Bericht desselben Weltwirtschaftsforums fußte. Und der dritte zu einem [Artikel](#) der in einem Peer-Review-Verfahren geprüft und in *Nature* veröffentlicht wurde. Darin stand zum einen, KI hätte das Potenzial, einen Beitrag zum Klimaschutz zu leisten, aber zugleich auch, dass sie ebenfalls das Potenzial hätte, die Klimakrise weiter zu verschlimmern.

Die leitenden Forscher:innen der deutschen Datenethikkommission beantworteten meine Anfrage zur Offenlegung ihrer Quellen nicht.

Ich überlasse das letzte Wort Karsten Smid, der bei Greenpeace Deutschland für Veröffentlichungen zu klimabezogenen Themen verantwortlich ist. In einer E-Mail an AlgorithmWatch schrieb er: „Wir würden uns ja schon freuen, wenn statt ‚künstlicher Intelligenz‘ zumindest ‚menschliche Vernunft‘ im Klimaschutz Anwendung findet.“

**„WIR WÜRDEN UNS JA SCHON
FREUEN, WENN STATT KÜNSTLICHER
INTELLIGENZ ZUMINDEST
MENSCHLICHE VERNUNFT IM
KLIMASCHUTZ ANWENDUNG WINDET.“**

KARSTEN SMID

Deutschland

Von Louisa Well

Kontextualisierung

Themen im Zusammenhang mit ADM-Systemen tauchen in der öffentlichen Diskussion immer wieder auf, für gewöhnlich jedoch nicht auf den Titelseiten. Automatisierte Entscheidungsfindung oder ADM (automated decision-making) wird im Allgemeinen unter dem Label „Künstliche Intelligenz“ (KI) zusammengefasst, und die Trennlinien zwischen Digitalisierung, Automatisierung und Maschinellen Lernen sind in der Regel fließend. Eine Diskussion über ADM entzündete sich, als die Bundespolizei am Bahnhof Südkreuz in Berlin ein Pilotprojekt zur Gesichtserkennung startete. Akteur:innen der Zivilgesellschaft kritisierten das Programm nicht nur wegen seiner qualitativ schlechten Ergebnisse, sondern auch wegen seines Potenzials, Grundrechte zu verletzen. In Deutschland kommen viele ADM-Systeme auf den Markt, und den meisten von ihnen wird kein solches Maß an Aufmerksamkeit zuteil.

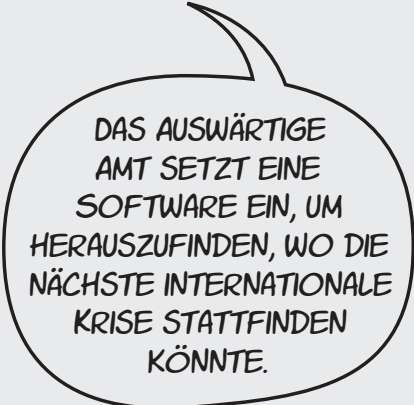
Anwendungen zur Risikoprognose sind auf dem Vormarsch. Polizeiintern werden sie genutzt, um das Bedrohungspotenzial unter Beobachtung stehender Islamisten:innen zu analysieren oder Hotspots für Einbrüche vorherzusagen. Das Auswärtige Amt setzt eine Software ein, um herauszufinden, wo die nächste internationale Krise stattfinden könnte. Schulen und Büros nutzen ein automatisiertes Analysetool, um gewalttätiges Verhalten aufzuspüren, ob nun durch Amokläufer:innen, Terrorist:innen oder gewalttätige Partner. Automatisierte Text-, Sprach-, und Bilderkennung hilft, Häftlingssuizide zu verhindern, Dialekte von Migrant:innen zu analysieren oder Kinderpornografie aufzuspüren. Automatisierung ist auch im Sozialsystem zu finden, wo es überwiegend genutzt wird, um Verwaltungsabläufe zu beschleunigen.

Alle diese Fälle zeigen, dass ADM-Systeme im Begriff sind, viele verschiedene Bereiche der Gesellschaft zu durchdringen – ein Trend, der sich in den nächsten Jahren noch verstärken wird. In Deutschland geht der Kurs auf Regierungsebene klar in Richtung Entwicklung und Implementierung von

KI-Anwendungen.

Die KI-Strategie der Regierung nimmt die Finanzierung entsprechender Forschung sowie die Unterstützung kleiner und mittlerer Unternehmen (KMU) und Startups in den Fokus.

Deutschland ist ein wichtiger Akteur in der europäischen KI-Landschaft.



DAS AUSWÄRTIGE AMT SETZT EINE SOFTWARE EIN, UM HERAUSZUFINDEN, WO DIE NÄCHSTE INTERNATIONALE KRISE STATTFINDEN KÖNNTE.

Katalog mit Anwendungsfällen: ADM im Einsatz

/ Vorhersagende Polizeiarbeit

/ Bundesebene

Das Bundeskriminalamt (BKA) nutzt seit 2017 das Risikoprognose-Tool [RADAR-iTE](#) (Bundeskriminalamt, 2017), um „militante Salafist:innen“ in drei Gefährdungsstufen einzuteilen (hoch, verdächtig und moderat). Das System wurde in Zusammenarbeit mit dem Fachbereich Forensische Psychologie der Universität Konstanz entwickelt. Um eine Person einzuschätzen, die den Behörden bereits bekannt ist, füllt ein:e Fallbearbeiter:in einen standardisierten Fragebogen aus, in dem das „beobachtbare Verhalten“ des oder der Betroffenen erfasst wird, und greift neben Daten, die die Polizei bereits über ihn oder sie gesammelt hat, auch auf alle Informationen zurück, die ihr legal zugänglich sind. Auf der Grundlage der Ergebnisse Gefährdungsstufe entscheidet sie (oder die betreffende Abteilung) über das weitere Vorgehen.

/ Landesebene

Auf Landesebene haben verschiedene Polizeistellen Versuche mit „Predictive Policing“-Software (Heitmüller, 2019) unternommen, von denen derzeit einige im Einsatz sind. In Bayern errechnet das Tool PRECOBS (Institut für musterbasierte Prognosetechnik, 2018), an welchen Orten mit der höchsten Wahrscheinlichkeit Einbruchsdiebstähle begangen werden, wohingegen die Arbeit mit derselben Software in Baden-Württemberg aufgrund von Problemen mit der Qualität der gelieferten Daten (Mayer, 2019) 2019 wieder eingestellt wurde. Auf dem Gebiet der Einbruchsprävention kommt im Bundesland Hessen das Prognose-Tool KLB-operativ zum Einsatz (Polizei Hessen, 2017), und Berlin nutzt KrimPro (Dinger, 2019). Beide Tools analysieren Daten, um herauszufinden, wo ein potenzieller Einbruch mit der höchsten Wahrscheinlichkeit stattfinden wird. Nordrhein-Westfalen hat auf der Basis von IBM-Produkten SKALA entwickelt (Polizei Nordrhein-Westfalen, 2020), während Niedersachsen PreMAP nutzt (Niedersächsisches Ministerium für Inneres und Sport, 2018); diese beiden Tools treffen Vorhersagen zu Hotspots für Einbruchsdiebstähle, die die Polizei in ihre Aktionspläne einbezieht.

/ Identifizierung von Kinderpornografie

Ein KI-Tool zur Identifizierung pornografischer Abbildungen von Kindern (Richter, 2019) wurde in Zusammenarbeit zwischen Microsoft und dem Justizministerium des Bundeslandes Nordrhein-Westfalen, und zwar der bei der Staatsanwaltschaft Köln angesiedelten Zentral- und Ansprechstelle Cybercrime (ZAC) entwickelt. Bevor das sensible Bildmaterial auf die Server von Microsoft hochgeladen wird, verwischt eine Software auf den Servern der Polizei die im Rahmen der polizeilichen Ermittlungen gespeicherten Bilder, so werden auch die strengen Bestimmungen zu Datenschutz und Kindeswohl eingehalten. In der Cloud analysieren Algorithmen die Bilder auf pornografische Inhalte, identifizieren die Gesichter von Opfern und Täter:innen und vergleichen sie mit bereits existierenden Profilen in einer Datenbank. Anschließend werden die Ergebnisse zur weiteren Analyse

wieder an die Polizei übergeben. Diese Software hat das Ziel, das Arbeitspensum und die mentale Belastung der Polizeibeamt:innen bei Ermittlungen zu Kindesmissbrauch zu reduzieren.

/ Verhinderung von Häftlings-suiziden

Das Justizministerium des Bundeslandes Nordrhein-Westfalen hat 2019 ein Programm gestartet, das Suizide von Häftlingen verhindern soll (Schmalen, 2019). Die Behörde beauftragte die Firma FusionSystems GmbH in Chemnitz (Systems, 2019) mit dem Aufbau eines Videoüberwachungssystems, das verdächtige Objekte wie etwa geknotete Stricke oder Feuerzeuge in einer Zelle erkennen kann und die diensthabenden Beamt:innen informiert. Das System ist für den Einsatz bei inhaftierten Personen gedacht, bei denen ein mittleres bis hohes Selbstmordrisiko vorliegt, und soll die derzeit alle 15 Minuten durchgeführten persönlichen Kontrollgänge ersetzen. Diese sind in die Kritik geraten (Schmalen, 2019), weil sie den emotionalen Stress der Insass:innen potenziell erhöhen.

/ Außenpolitisches Krisenmanagement

Das Auswärtige Amt nutzt das Datenanalysetool PREVIEW – Prediction, Visualization, Early Warning (Auswärtiges Amt, 2019), um sich anbahnende internationale Krisen zu erkennen. Das Tool analysiert öffentlich zugängliche Daten zu aktuellen politischen, ökonomischen und gesellschaftlichen Tendenzen und Konflikten, um sich entwickelnde Krisen zu identifizieren. Aussagen des Auswärtigen Amtes zufolge kommt KI zum Einsatz, um die Daten zu verarbeiten und anschließend Infografiken und Karten zu erstellen, die einen Beitrag zum Verständnis der aktuellen Situation in einem bestimmten Konflikt leisten. Darüber hinaus illustrieren Trendanalysen, welche Richtung politische und gesellschaftliche Entwicklungen nehmen könnten. PREVIEW kommt in der Abteilung S des Auswärtigen Amtes zum Einsatz, das für internationale Stabilisierungsmaßnahmen und Kriseneinsätze verantwortlich ist. Die Ergebnisse sollen die Beamt:innen bei der Entscheidung über die nächsten einleitenden Schritte unterstützen.

/ Identitätsprüfung bei Migrant:innen

Das Bundesamt für Migration und Flüchtlinge (BAMF) verwendet seit 2017 automatische Text- und Spracherkennungssysteme, um die Identität von Geflüchteten festzustellen (Thüer, Köver und Fanta, 2018). Mitarbeiter:innen können Asylsuchende bitten, ihnen Zugang zu Mobilite-

DIESE SOFTWARE HAT DAS ZIEL, DAS ARBEITSPENSUM UND DIE MENTALE BELASTUNG DER POLIZEIBEAMT:INNEN BEI ERMITTLUNGEN ZU KINDESMISSBRAUCH ZU REDUZIEREN.

lefonen, Tablets oder Laptops zu gewähren, um die von ihnen gemachten Angaben zu ihrer Herkunft auf ihren Wahrheitsgehalt zu überprüfen. Die Behörde ist in der Lage, alle Daten zu ermitteln, die sich auf diesen Geräten befinden, und mit einer Software zu untersuchen. Diese gibt den Bearbeiter:innen einen begrenzten Überblick über den Inhalt, inklusive einer Sprachanalyse der gefundenen Texte. Aussagen des BAMF zufolge wurden sowohl die Software als auch die Hardware von der [Firma Atos SE](#) (Biselli, 2017) zur Verfügung gestellt, wobei das [VICE Magazine Belege dafür gefunden hat](#) (Biselli, 2018b), dass auch [MSAB](#), ein Hersteller von Komplettlösungen für mobile Forensik, beteiligt war. Ein weiteres vom BAMF genutztes Tool [zielt auf die Identifizierung bewusst verschleierter Dialekte in der gesprochenen Sprache ab](#) (Biselli, 2018a). Verfügt eine Asylbewerber:in nicht über ein gültiges Ausweisdokument, wird die Person gebeten, in der eigenen Muttersprache eine etwa zwei Minuten lange Beschreibung eines Bildes zu geben. Diese Aufnahme wird einer Softwareanalyse unterzogen, die anschließend einen Prozentsatz errechnet, der angibt, wie nahe das Gesprochene einem bestimmten Dialekt kommt.

/ Sozialstaat

Die Behörde für Arbeit, Soziales, Familie und Integration in Hamburg nutzt eine [Software namens JUS-IT](#) (Behörde für Arbeit, Soziales, Familie und Integration, 2018) für den Allgemeinen Sozialen Dienst, den Pflegekinderdienst und die Wirtschaftliche Jugendhilfe. Sie kommt in der Fallverwaltung sowie für automatisierte Zahlungen zum Einsatz und ist mit Schnittstellen ausgerüstet, die eine Verbindung zu Polizeiberichten und Krankenkassen herstellen. Das System basiert auf [Cúram](#), einem modularen, handelsüblichen IBM-Produkt, das auf spezielle Bedürfnisse angepasst werden kann – und in der Kritik stand, [in Kanada nicht korrekt funktioniert zu haben](#) (Human Rights Watch, 2018). Im Jahr 2018 fand eine [Enquete-Kommission Kinderschutz und Kinderrechte](#) (Bürgerschaft der Freien Und Hansestadt Hamburg, 2018) heraus, dass JUS-IT die administrativen Prozesse verlängert und damit weniger Zeit für dringend notwendige Familienbesuche bleibt. Im Ergebnis empfiehlt die Kommission entweder eine umfassende Überarbeitung oder die [Abschaltung der Software](#) (Lasarzik, 2019).

Die Bundesagentur für Arbeit nutzt ein IT-System namens ALLEGRO für die [Verwaltung der Arbeitslosenhilfe](#). Die Mitarbeiter:innen der Behörde geben die Daten von Antragsteller:innen ein, und das System errechnet die entsprechenden finanziellen Hilfen. Das System ist in der

Lage, eine Verbindung zu Kranken- und Rentenkassen sowie dem Zoll und dem [Ausländerzentralregister](#) (Deutscher Bundestag, 2018) herzustellen. ALLEGRO ersetzte 2012 die Vorgängersoftware namens A2LL, ein [fehleranfälliges System von T-Systems](#). Die neue Software wurde [von der Bundesarbeitsagentur intern entwickelt](#) (Borchers, 2008).

/ Risiko-Scoring von Gewalt

Ähnlich dem oben beschriebenen, vom Bundeskriminalamt genutzten Tool haben Forschende des Instituts für Psychologie und Gefährdungsmanagement der Universität Darmstadt ein Tool mit Namen [Screener Islamismus](#) entwickelt ([DyRIAS, 2019c](#)). Dieses Tool kommt zum Einsatz, um das Risikopotenzial einer Person zu berechnen, aufgrund extremer islamistischer Ansichten Gewalt auszuüben. Es basiert auf dem IT-System [DyRIAS](#) (Kurzbezeichnung für Dynamic Risk Assessment Systems) ([DyRIAS, 2019a](#)). Hinter dem Tool steht die Philosophie, dass Gewalt nicht mit dem Charakter eines Menschen zusammenhängt, sondern am Ende eines Prozesses auftritt, bei dem Täter:innen, Opfer und situative Einflüsse miteinander interagieren. Nutzer:innen beantworten diverse Fragen zum Verhalten der Person, die sie überprüfen möchten. Das System organisiert diese Antworten zu einem Überblick und [fügt Handlungsempfehlungen](#) hinzu ([DyRIAS, 2019c](#)), wie mit der Situation umzugehen ist. Welche Institutionen das System implementieren wird, bleibt abzuwarten, da es offenbar nicht gezielt für die Polizei entwickelt wurde, sondern für andere Einrichtungen wie Schulen und Büros etc.

Andere [DyRIAS-Anwendungen sind bereits im Einsatz](#). So nutzen zum Beispiel ([DyRIAS, 2019b](#)) Frauenhäuser in den Städten Singen und Weimar die Software [DyRIAS-Intimpartner](#), um das Gefährdungspotenzial gewalttätiger männlicher Partner zu ermitteln. Ein weiteres Beispiel findet sich beim Landesschulamt Sachsen-Anhalt. Dort kommt [DyRIAS-Schule](#) zum Einsatz, um das Gefährdungspotenzial von Schüler:innen zu analysieren, die in dem Verdacht stehen, potentiell einen Amoklauf zu planen. In einem anderen Fall durchleuchten Schweizer Arbeitgeber:innen ihre Angestellten mit Hilfe von [DyRIAS-Arbeitsplatz](#), um sich deren Gewaltpotenzial anzusehen oder herauszufinden, ob sich jemand eventuell zu einem Stalker entwickeln könnte. Mitarbeiter:innen füllen einen Fragebogen über die Person aus, bei der sie ein Gewaltpotenzial vermuten, und die Tools „erstellen anschließend automatisch einen Bericht“, dem ein Threat-Score entnommen werden kann.

Politische Debatten rund um ADM – Regierung und Parlament

/ Deutsche KI-Strategie

Die [KI-Strategie Deutschland](#) (Die Bundesregierung, 2019), veröffentlicht im November 2018, definiert drei Kernziele:

- 1) Deutschland und Europa zu einem führenden Standort für die Entwicklung und Anwendung von KI-Technologien zu machen und die künftige Wettbewerbsfähigkeit Deutschlands zu sichern,
- 2) eine verantwortungsvolle und gemeinwohlorientierte Entwicklung und Nutzung von KI sicherzustellen.
- 3) KI im Rahmen eines breiten gesellschaftlichen Dialogs und einer aktiven politischen Gestaltung ethisch, rechtlich, kulturell und institutionell in die Gesellschaft einzubetten.

Im Rahmen der Strategie werden 3 Milliarden Euro für die Finanzierung von KI-Projekten zur Verfügung gestellt. Im Nachgang zu einer [schriftlichen Anfrage der Bundestagsfraktion Bündnis 90/ Die Grünen](#) (Deutscher Bundestag, 2019) hat die Bundesregierung außerdem offengelegt, wie die Mittel verteilt werden sollen.

Den Löwenanteil erhält das Bundesministerium für Bildung und Forschung (170 Millionen Euro). Er dient der Konsolidierung bestehender KI-Forschungsinstitute in Deutschland, der Finanzierung zusätzlicher Lehrstühle, soll internationale Experten-innen nach Deutschland locken, die akademische Lehre verbessern und in junge Wissenschaftler-innen investiert werden. Ein geringfügig kleinerer Anteil geht an das Bundesministerium für Wirtschaft und Energie (147 Millionen Euro), gefolgt vom Bundesministerium für Arbeit und Soziales (74 Millionen Euro), während der Rest zwischen sieben weiteren Ministerien sowie dem Bundeskanzleramt aufgeteilt wird.

Im Rahmen der KI-Strategie hat die Bundesregierung beschlossen, beim Bundesministerium für Arbeit und Soziales ein [KI-Observatorium](#) einzurichten. Es wird technologische Bewertungen im Hinblick auf die Verbreitung von KI und

deren Auswirkungen auf den Arbeitsmarkt und die Gesellschaft als Ganzes durchführen. Auf diese Weise sollen die interinstitutionelle Kooperation innerhalb Europas befördert, ein gemeinsames Verständnis von KI entwickelt und ein Rahmenwerk von Richtlinien und Prinzipien für die Entwicklung und den Einsatz von KI auf dem Arbeitsmarkt geschaffen werden.

Die [KI-Strategie wird laufend angepasst](#), und viele Projekte, die sich mit Fragestellungen rund um Forschung, Innovation, Infrastruktur, Administration, Data-Sharing-Modelle, internationale Standards und zivile Anwendungen beschäftigen, werden derzeit aufgesetzt.

/ Deutsche Datenstrategie

Im November 2019 veröffentlichte die Bundesregierung die [Kernpunkte](#) ihrer künftigen Datenstrategie. Dazu gehören der erleichterte Zugang zu Daten, die Forderung nach einem verantwortungsvollen Umgang mit Daten, die Verbesserung der Datenkompetenzen innerhalb der Gesellschaft und das Ziel, Deutschland zu einem Pionier der Datenkultur zu machen. Da Daten ein integraler Bestandteil von ADM-Systemen sind, wird diese Strategie einen erheblichen Einfluss darauf haben, welche Anwendungen in Zukunft auf den Markt gelangen. Der Entwicklungsprozess der Datenstrategie wurde begleitet von [Anhörungen](#) mit Vertreter-innen aus Zivilgesellschaft, Wirtschaft und Wissenschaft, an denen sich auch die Öffentlichkeit online beteiligen konnte.

/ Datenethikkommission

Im Jahr 2018 wurde von der Bundesregierung die [Datenethikkommission](#) (Datenethikkommission, 2019) ins Leben gerufen, um die ethischen Implikationen von Big Data und KI zu diskutieren. Ein Jahr später stellte das Gremium einen Bericht vor, der etliche Empfehlungen zum Umgang mit algorithmischen Systemen enthielt. Eine seiner generellen Forderungen bezog sich darauf, dass Algorithmen ein menschenzentriertes Design haben sollten. Dazu gehören: Einhaltung gesellschaftlicher Grundwerte, Nachhaltigkeit, Qualität und Leistungsfähigkeit, Robustheit und Sicherheit, Minimierung von Verzerrungen (bias) und Diskriminierung, Transparenz, Erklärbarkeit und Nachvollziehbarkeit sowie klare Rechenschaftsstrukturen. Interessanterweise führte die Kommission zudem das Konzept der „Systemkritikalität“ ein, insbesondere im Hinblick auf Transparenz und Kontrolle. Mit diesem Ansatz stellte sie die Forderung auf, die Wahrscheinlichkeit und

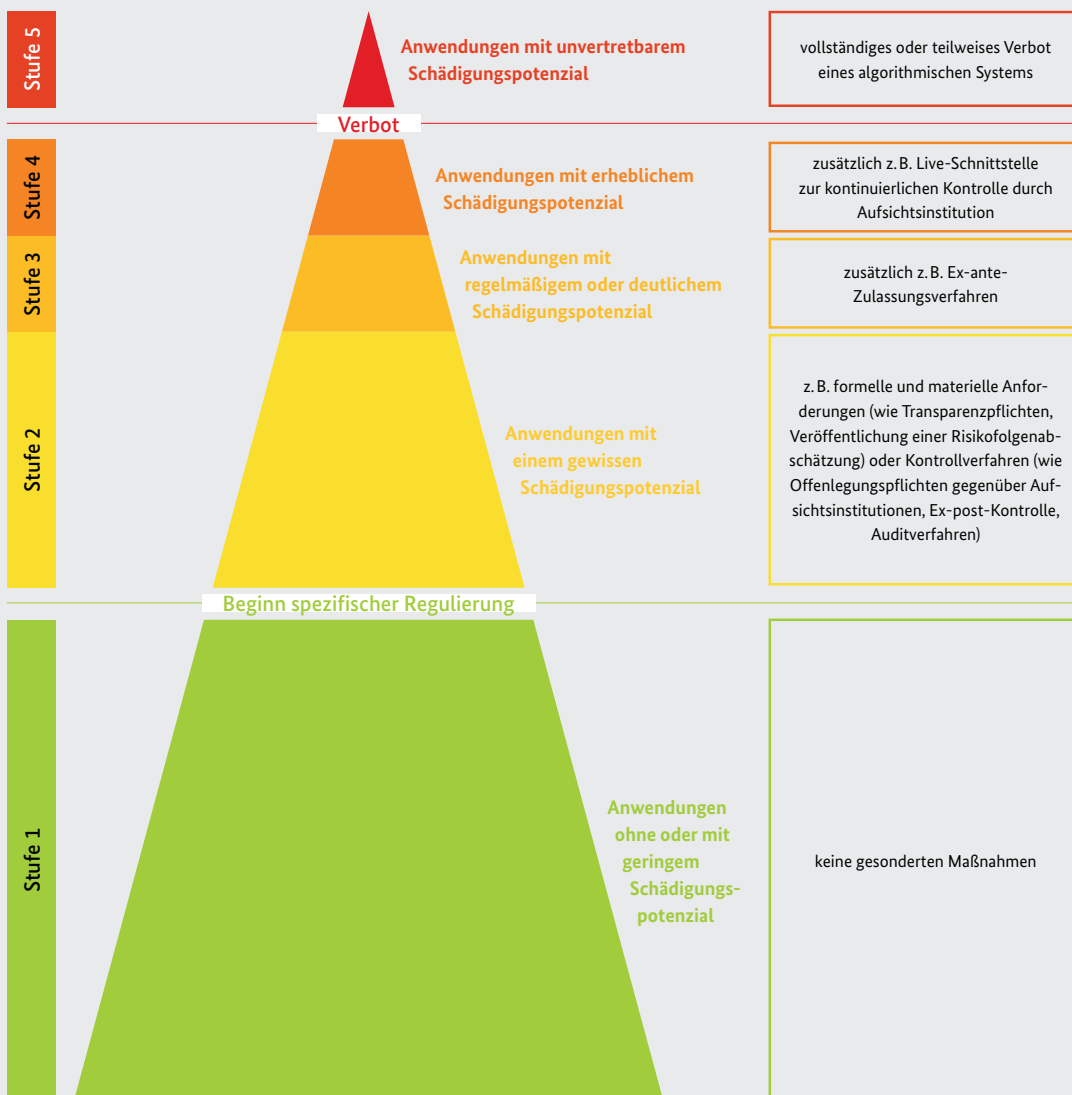
die Schwere potenzieller, von algorithmischen Systemen verursachter Schäden einzubeziehen. Darüber hinaus empfahl die Kommission, Algorithmen anhand einer „Kritikalitätspyramide“ zu prüfen. Die Kritikalität algorithmischer Systeme wird mit Hilfe einer Stufenskala von „1= nicht kritisch, keine Maßnahmen erforderlich“ bis „5 = hohes Schadenspotenzial, teilweises oder vollständiges Verbot des Systems“ bewertet. Zwischen diesen beiden Extremen schlagen sie abgestufte Transparenzverpflichtungen, Ex-

ante-Zulassungsverfahren sowie eine kontinuierliche Kontrolle durch die Aufsichtsbehörden vor.

/ Enquete-Kommission Künstliche Intelligenz

Ende 2019, ein Jahr nach Einrichtung der Kommission im Parlament, veröffentlichten die Projektgruppen die Zwischenergebnisse von Expertenanhörungen und Diskussio-

Kritikalitätspyramide und risikoadaptiertes Regulierungssystem für den Einsatz algorithmischer Systeme



Quelle: Bericht Datenethikkommission 2019

nen, die im Jahr zuvor stattgefunden hatten. Die [Projektgruppe KI und Wirtschaft](#) (Enquete-Kommission Künstliche Intelligenz, 2019c) stellte fest, dass KI das Potenzial besitzt, sowohl Produktivität zu steigern und Nachhaltigkeit zu verbessern, als auch soziale Ungerechtigkeiten und eine eingeschränkte Teilhabe am Arbeitsmarkt und am gesellschaftlichen Leben zu verstärken. Die [Projektgruppe KI und Staat](#) (Enquete-Kommission Künstliche Intelligenz, 2019b) betonte die Fürsorgepflicht des Staates bei der Einführung von KI-Systemen und schlug Mechanismen vor, die die Fähigkeit der Bürger:innen stärken, Entscheidungen zu treffen und Entscheidungen zurückzuverfolgen. Angesichts der Tatsache, dass Gesundheitsdaten sehr sensibel sind und eines hohen Schutzes bedürfen, sieht die [Projektgruppe KI und Gesundheit](#) (Enquete-Kommission Künstliche Intelligenz, 2019a) KI als einen potenziellen Treiber für verbesserte Diagnosen und Therapien sowie Kranken- und Altenfürsorge.

/ Deutsch-französische Arbeitsgruppe // Disruptive Innovationen und Künstliche Intelligenz

Um sich dem weltweiten Wettbewerb auf dem Gebiet der Technologie und Innovation zu stellen, insbesondere der Künstlichen Intelligenz, wurde 2019 von der Deutsch-französischen Parlamentarischen Versammlung eine [Arbeitsgruppe](#) (Deutsch-Französische Parlamentarische Versammlung, 2019) zu diesem Thema gegründet. Ziel der Gruppe ist es, die Arbeit der beiden Parlamente in Hinblick auf KI-Themen zu koordinieren und ein europäisches Rahmenwerk zu schaffen, das für ein verbessertes, von europäischen Werten geleitetes Innovationsvermögen sorgt.

ALS INNENMINISTER HORST SEEHOFER EINEN GESETZENTWURF VORSTELLTE, DER ES DER BUNDESPOLIZEI GESTATTEN SOLLTE, LANDESWEIT GESICHTSERKENNUNGSSOFTWARE EINZUSETZEN, STIESSEN DIESE PLÄNE BEI VIELEN PARTEIEN AUF WIDERSTAND UND SEEHOFER ZOG DEN GESETZENTWURF ZURÜCK.

/ Gesetzentwurf zu Gesichtserkennung gescheitert

In den Jahren 2017/18 führte die Polizei am Bahnhof Südkreuz in Berlin ein Pilotprojekt zur Gesichtserkennung durch. Es wurde nicht nur von Bürgerrechts- und Datenschutzaktivist:innen [kritisiert](#), sondern erbrachte auch eine [hohe Anzahl falsch positiver und falsch negativer Ergebnisse](#). Als Innenminister Horst Seehofer einen Gesetzentwurf vorstellte, der es der Bundespolizei gestatten sollte, landesweit Gesichtserkennungssoftware einzusetzen, stießen diese Pläne bei vielen Parteien auf Widerstand und [Seehofer zog den Gesetzentwurf zurück](#).

/ Zivilgesellschaft und Akademischer Bereich

/ Zivilgesellschaft

Die Zivilgesellschaft ist auf dem Gebiet der Digitalisierung im Allgemeinen und der ADM-Prozesse im Speziellen sehr aktiv. Ein wichtiger Diskussionsstrang in der Öffentlichkeit konzentriert sich auf das Thema KI und Nachhaltigkeit. Zu den Highlights dieser Debatte gehörte die [Konferenz Bits & Bäume](#) (Bits und Bäume, 2019) 2018, die zum Ziel hatte, die Communitys der Umweltschützer:innen und Technolog:innen zusammenzubringen, um die Wechselbeziehungen zwischen Fragen der Digitalisierung und des Umweltschutzes zu diskutieren. Im Jahr 2019 stellte der [Wissenschaftliche Beirat der Bundesregierung Globale Umweltveränderungen](#) (WBGU, 2019) nach einer Analyse der komplexen Wechselbeziehungen zwischen den beiden Themenfeldern ein detailliertes Gutachten mit dem Titel [Unsere gemeinsame digitale Zukunft](#) vor. Der Bericht machte auf eine Lücke in den Nachhaltigen Entwicklungszielen in Hinblick auf ADM-Prozesse aufmerksam und stellte Forderungen nach Rückverfolgbarkeit, juristisch einklagbaren Rechten und einer gesetzgeberischen Debatte zu den Haftungsrechten im Zusammenhang mit algorithmischen Systemen auf. Die Bertelsmann Stiftung entwickelte im Rahmen ihres Projektes [Ethik der Algorithmen](#) (Bertelsmann Stiftung, 2019b) die [Algo.Rules](#) (Bertelsmann Stiftung, 2019a). Sie enthalten neun Leitlinien für das Design algorithmischer Systeme. Darüber hinaus veröffentlichte die Stiftung Arbeitspapiere zu einer breiten Anzahl von Themen und führte außerdem Studien zum [Stand des Wissens und zur Akzeptanz für Algorithmen in Deutschland und Europa](#) (Grzymek, 2019) durch. Der Think Tank [Stiftung Neue Verantwortung](#) (Stiftung Neue Verantwortung, 2019)

führte ein zweijähriges Projekt zum Thema Algorithmen und Gemeinwohl durch und konzentrierte sich dabei auf die Bedeutung einer starken Zivilgesellschaft, Gesundheitsthemen, Personalpolitik, Predictive Policing, Corporate Digital Responsibility und den Wohlfahrtssektor. Die Investigativplattform [Netzpolitik.org](https://netzpolitik.org) (Netzpolitik.org, 2019) und die Watchdog-Organisation [AlgorithmWatch](https://algorithmwatch.org) (AlgorithmWatch, 2020a) berichten kontinuierlich über [neue ADM-Anwendungsfälle](https://algorithmwatch.org) (AlgorithmWatch, 2020b). AlgorithmWatch führt außerdem Forschungsprojekte zu den Themen Plattform-Governance, Human Resource Management und Kredit-Scoring durch und erstellt Mappings zum aktuellen Stand von ADM in Deutschland und Europa.

/ Akademischer Bereich

Der akademische Diskurs zu ADM findet auf verschiedenen Ebenen statt: Manches davon fokussiert sich auf die Entwicklung von ADM-Systemen, anderes auf die Forschung zu den ethischen Implikationen. Ein Beispiel für Forschung über ADM-Systeme, die bereits im Einsatz ist, findet sich im Gesundheitssektor. Das Bundesministerium für Wirtschaft und Energie finanziert im Rahmen des [Wettbewerbs Smarte Datenwirtschaft](https://www.wettbewerbs-smarte-datenwirtschaft.de) (Hans Böckler Stiftung, 2018) zwei medizinische Projekte, bei denen KI zur Verarbeitung von Patient-innendaten zum Einsatz kommen soll. 2019 startete das Projekt [Smart Medical Doctor – From Data to Decision](https://www.smart-medical-doctor.de) (Bundesministerium für Wirtschaft und Energie, 2019). Ziel dieses Projektes ist die Entwicklung einer medizinischen Datenplattform, auf der KI zur Anonymisierung und Aufbereitung von Daten zu Krankenhausaufenthalten von Patient:innen zum Einsatz kommt, um sie für die Nutzung durch ein unterstützendes Diagnosetool für Ärzt:innen mit Namen Ada DX verfügbar zu machen. An dem Projekt sind die Ada Health GmbH, die Helios Kliniken GmbH, die Universitätsklinik Charité und die Beuth-Fachhochschule Berlin beteiligt. Ein zweites Projekt namens [Telemed5000](https://www.telemed5000.de) (Telemed5000, 2019) hat zum Ziel, ein ferngesteuertes Patient:innen-Managementsystem zu entwickeln, das es Ärzt:innen erlaubt, eine große Zahl von Herzpatient:innen zu verwalten. Das deutsch-österreichische Gemeinschaftsprojekt soll Systeme des Internets der Dinge (IoT) mit Deep-Learning-Technologien verbinden, um Herzpatient:innen zu überwachen.

Darüber hinaus existieren diverse Forschungsprojekte zu den gesellschaftlichen Implikationen von Algorithmen in Deutschland. Zu den erwähnenswerten Beispielen gehören unter anderem das [Weizenbaum Institut – Forschung für die vernetzte Gesellschaft](https://www.weizenbaum-institut.de) (Weizenbaum Institut, 2019); das

Algorithm Accountability Lab (TU Kaiserslautern, 2019) und der [Studiengang Sozioinformatik](https://www.studiengang-sozioinformatik.de) an der Technischen Universität Kaiserslautern; daneben die neugegründete [Hochschule für Politik an der Technischen Universität München](https://www.hochschule-politik.de) (TUM, 2019).

Erkenntnisse

Die Narrative rund um ADM bewegen sich zwischen den Polen des Technologischen Solutionismus, dessen Vertreter:innen KI und ADM als Lösung all unserer gesellschaftlichen Probleme preisen, und der technophobisch geprägten Ablehnung alles Neuen. In vielen Fällen liegt die Wahrheit irgendwo dazwischen, und daher ist es wichtig, mehr über die Feinheiten von ADM-Systemen in Erfahrung zu bringen, um Gut und Böse voneinander unterscheiden zu können. Vielleicht wird mit Hilfe von maschinellem Lernen ein Heilmittel für Krebs gefunden, und ADM kann potenziell bei der Bewältigung der Herausforderungen im Zusammenhang mit der Reduzierung der CO₂-Emissionen helfen. Gleichzeitig können eine allgegenwärtige Datifizierung zahlreicher Lebensbereiche und deren Einbindung in ADM-Systeme Grundrechte bedrohen und müssen daher sorgfältigst auf Ungerechtigkeiten abgeklopft werden. Eines ist gewiss: Die Welt wird sich verändern. Die Verschiebungen, die künftig durch ADM-Systeme in Wirtschaft und Wissenschaft, im Labor, im Gesundheitswesen, im Finanzwesen, in der Sozialfürsorge und auf vielen anderen Gebieten stattfinden werden, müssen von einer Gesellschaft begleitet werden, die bereit ist, Mängel zu kritisieren und gleichzeitig dort Veränderungen zuzulassen und anzunehmen, wo sie nützlich und erwünscht sind. Zu lange waren Diskussionen über ADM allein mit dem autonomen Fahren oder dem Bild von Robotern verbunden, die Menschen ersetzen. Es ist notwendig und an der Zeit, dass die Gesellschaft mehr über die Herausforderungen und Chancen lernt, die ADM-Systeme mit sich bringen. Auf Regierungsebene ist die Enquete-Kommission Künstliche Intelligenz ein guter Ausgangspunkt für politische Entscheidungsträger:innen, sich eine bessere Perspektive auf das Thema zu erarbeiten. Akteure der Zivilgesellschaft wie AlgorithmWatch arbeiten ununterbrochen daran, die Funktionsweise verborgener Algorithmen zu durchleuchten. Wir brauchen eine breite gesellschaftliche Diskussion darüber, wie ADM das Leben besser machen kann und wo wir die roten Linien sehen, die nicht überschritten werden dürfen.

Quellen:

AlgorithmWatch (2020a)
AlgorithmWatch. Verfügbar auf: <https://algorithmwatch.org/en/>
(Abgerufen: 17. Januar 2020).

AlgorithmWatch (2020b)
Stories. Verfügbar auf: <https://algorithmwatch.org/en/stories-en/>
(Abgerufen: 17. Januar 2020).

Auswärtiges Amt (2019)
PREVIEW. Verfügbar auf: https://www.auswaertiges-amt.de/de/aussenpolitik/themen/krisenpraevention/8_Krisenfrueherkennung (Abgerufen: 30. November 2010).

Behörde für Arbeit Soziales Familie und Integration (2018) JUS-IT Hamburg. Verfügbar auf: <https://www.hamburg.de/jus-it/> (Abgerufen: 17. Januar 2020).

Bertelsmann Stiftung (2019a) Algo. Rules. Verfügbar auf: <https://algorules.org/en/home/> (Abgerufen: 17. Januar 2020).

Bertelsmann Stiftung (2019b) Ethics of Algorithms. Verfügbar auf: <https://www.bertelsmann-stiftung.de/en/our-projects/ethics-of-algorithms/project-description/> (Abgerufen: 17. Januar 2020).

Bertelsmann Stiftung (2019c): Predictive Policing – Mit Algorithmen vor die Lage kommen. Verfügbar auf: <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/predictive-policing-mit-algorithmen-vor-die-lage-kommen/>

Biselli, A. (2017) Syrien oder Ägypten? Software zur Dialektanalyse ist fehleranfällig und intransparent, Netzpolitik.org. Verfügbar auf: <https://netzpolitik.org/2017/syrien-oder-aegypten-software-zur-dialektanalyse-ist-fehleranfaellig-und-intransparent/> (Abgerufen: 17. Januar 2020).

Biselli, A. (2018a) Die IT-Tools des BAMF: Fehler vorprogrammiert, Netzpolitik.org. Verfügbar auf: <https://netzpolitik.org/2018/die-it-tools-des-bamf-fehler-vorprogrammiert/> (Abgerufen: 17. Januar 2020).

Biselli, A. (2018b) Handys von Asylbewerbern zu analysieren, kostet viel mehr als geplant, Vice Magazine, April 2018. Verfügbar auf: <https://www.vice.com/de/article/kzxy8n/handys-von-asylbewerbern-zu-analysieren-kostet-viel-mehr-als-bisher-geplant>

Biselli, A. (2019) Asylbehörde sucht mit Künstlicher Intelligenz nach auffälligen Geflüchteten. Verfügbar auf: <https://netzpolitik.org/2019/asylbehoerde-sucht-mit-kuenstlicher-intelligenz-nach-auffaelligen-gefluechteten/#vorschaltbanner>

Bits und Bäume (2019) Bits und Bäume. Verfügbar auf: <https://bits-und-baeume.org/en> (Abgerufen: 17. Januar 2020).

Borchers, D. (2008) Hartz-IV-Software: Allegro soll A2LL ablösen, Heise, 10. März. Verfügbar auf: <https://www.heise.de/newsticker/meldung/Hartz-IV-Software-Allegro-soll-A2LL-abloesen-189165.html>

Bundeskriminalamt (2017) Presseinformation: Neues Instrument zur Risikobewertung von potentiellen Gewaltstraftätern. Verfügbar auf: https://www.bka.de/DE/Presse/Listenseite_Pressemitteilungen/2017/Presse2017/170202_Radar.html (Abgerufen: 17. Januar 2020).

Bundesministerium für Wirtschaft und Energie (2019) Telemed5000. Verfügbar auf: https://www.digitale-technologien.de/DT/Redaktion/DE/Standardartikel/Smarte-Datenwirtschaft-Projekte/smarte_datenwirtschaft_telemed5000.html (Abgerufen: 17. Januar 2020).

Bürgerschaft der Freien und Hansestadt Hamburg (2018) Bericht der Enquete-Kommission „Kinderschutz und Kinderrechte weiter stärken: Verfügbar auf: https://www.buergerschaft-hh.de/parldok/dokument/65251/bericht_der_enquete_kommission_kinderschutz_und_kinderrechte_weiter_staerken_ueberpruefung_weiterentwicklung_umsetzung_und_einhaltung_gesetzlicher_gru.pdf

Datenethikkommission (2019) Opinion of the Data Ethics Commission. Verfügbar auf: https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.pdf?__blob=publicationFile&v=1 (Abgerufen: 17. Januar 2020).

Deutsch-Französische Parlamentarische Versammlung (2019) Deutsch-Französische Offensive für disruptive Innovationen und Künstliche Intelligenz. Verfügbar auf: https://www.bundestag.de/resource/blob/659556/6050efd45a80ac7a7214e2c49d03427a/20190924_beschluss_arbeitsgruppe_riki-data.pdf

Deutscher Bundestag (2018) Drucksache19/5014. Verfügbar auf: <https://dip21.bundestag.de/dip21/btd/19/050/1905014.pdf>

Deutscher Bundestag (2019) Drucksache19/10101. Verfügbar auf: <https://dip21.bundestag.de/dip21/btd/19/101/1910101.pdf>

Die Bundesregierung (2019) KI-Strategie. Verfügbar auf: <https://www.ki-strategie-deutschland.de/home.html> (Abgerufen: 17. Januar 2020)

Dinger, A. (2019) Vorhersage-Software: So genau ist das Programm der Polizei, Berliner Morgenpost, 11 September. Verfügbar auf: <https://www.morgenpost.de/berlin/article216410297/Vorhersage-Software-So-genau-ist-das-Programm-der-Polizei.html>

DyRIAS (2019a) DyRIAS. Verfügbar auf: <https://www.dyrias.com/en/> (Abgerufen: 17. Januar 2020)

DyRIAS (2019b) Referenzen. Verfügbar auf: <https://www.dyrias.com/de/referenzen.html> (Abgerufen: 17. Januar 2020)

DyRIAS (2019c) Screener Islamismus. Verfügbar auf: https://www.dyrias.com/images/Screener_Islamismus/DyRIAS_Broschuere_Screener_Radikal.pdf

Enquete-Kommission Künstliche Intelligenz (2019a) Projektgruppe „KI und Gesundheit“. Verfügbar auf: <https://www.bundestag.de/resource/blob/672950/fed938366dcf1b3f79c2ff177e0f86f5/PG-3-Projektgruppenbericht-data.pdf>

Enquete-Kommission Künstliche Intelligenz (2019b) Projektgruppe „KI und Staat“. Verfügbar auf: <https://www.bundestag.de/resource/blob/672932/8dafccf73bce9d9560223d1683d82cb9/PG-2-Projektgruppenbericht-data.pdf>

Enquete-Kommission Künstliche Intelligenz (2019c) Projektgruppe „KI und Wirtschaft“. Verfügbar auf: <https://www.bundestag.de/resource/blob/672894/e82fdd5cb437de0aaf0da3d7033f84a3/PG-1-Projektgruppenbericht-data.pdf>

German Advisory Council on Global Change (WBGU) (2019) Towards our Common Digital Future. Verfügbar auf: <https://www.wbgu.de/en/> (Abgerufen: 17. Januar 2020)

Grzymek, V. (2019) Was Europa über Algorithmen weiß und denkt. Verfügbar auf: <https://algorithmenethik.de/2019/02/06/was-europa-ueber-algorithmen-weiss-und-denkt/> (Abgerufen: 17. Januar 2019).

Hans Böckler Stiftung (2018) Smarte Datenwirtschaft. Verfügbar auf: https://www.boeckler.de/pdf/p_mbf_praxis_2018_010.pdf

https://www.digitale-technologien.de/DT/Redaktion/DE/Kurzmeldungen/Aktuelles/2018/smarte_datenwirtschaft_20181217.html (Abgerufen: 20 November 2019).

Heitmüller, U. (2019) Missing Link: Predictive Policing – die Kunst, Verbrechen vorherzusagen, Heise, 19 May. Verfügbar auf: <https://www.heise.de/newsticker/meldung/Missing-Link-Predictive-Policing-die-Kunst-Verbrechen-vorherzusagen-4425204.html>

Human Rights Watch (2018) May 2019 Submission to the UN Special Rapporteur on Extreme Poverty & Human Rights Regarding His Thematic Report on Digital Technology, Social Protection & Human Rights. Verfügbar auf: <https://www.ohchr.org/Documents/Issues/Poverty/DigitalTechnology/HumanRightsWatch.pdf>

Institut für musterbasierte Prognosetechnik (2018) PRECOBS in Bayern. Verfügbar auf: https://www.ifmpt.de/project_bayern.html (Abgerufen: 17. Januar 2020)

Lasarzik, A. (2019) "Empathie ist wichtiger als strenge Kontrolle", Die Zeit, 16. Januar. Verfügbar auf: <https://www.zeit.de/hamburg/2019-01/kinderschutz-christian-prasser-jugendamt-hamburg-enquete-kommission-yagmur>

Mayer, N. (2019) Strobl entscheidet sich gegen Precobs, Stuttgarter Nachrichten, 3. September. Verfügbar auf: <https://www.stuttgarter-nachrichten.de/inhalt.aus-fuer-die-einbruchvorhersage-software-strobl-entscheidet-sich-gegen-precobs.19a18735-9c8f-4f1a-bf1b-80b6a3ad0142.html>

Netzpolitik.org (2019) Netzpolitik.org. Verfügbar auf: <https://netzpolitik.org/> (Abgerufen: 17. Januar 2010).

Netzpolitik.org (2020) Predictive Policing <https://netzpolitik.org/?s=predictive+policing>

Niedersächsisches Ministerium für Inneres und Sport (2018) Zweite Pilotphase startet: Einsatz der Prognosesoftware PreMAP wird zum 1. November weiter intensiviert. Verfügbar auf: https://www.mi.niedersachsen.de/aktuelles/presse_informationen/zweite-pilotphase-startet-einsatz-der-prognosesoftware-premap-wird-zum-1-november-weiter-intensiviert-170622.html (Abgerufen: 17. Januar 2020).

Polizei Hessen (2017) Prognose-Software gegen Wohnungseinbruchdiebstahl. Verfügbar auf: <https://www.polizei.hessen.de/icc/internetzentral/nav/f7b/broker.jsp?UCon=50460710-abca-6f51-a3ac-106409c082ee&uTem=bff71055-bb1d-50f1-2860-72700266cb59&uMen=f7b602efe369-b951-df5a-47e60ef798e7> (Abgerufen: 17. Januar 2020).

Polizei Nordrhein-Westfalen (2020) Projekt SKALA (Predictive Policing in NRW) – Ergebnisse. Verfügbar auf: <https://polizei.nrw/artikel/projekt-skala-predictive-policing-in-nrw-ergebnisse> (Abgerufen: 17. Januar 2020).

Richter, I. (2019) Automatische Bilderkennung hilft im Einsatz gegen Kinderpornografie, Microsoft. Verfügbar auf: <https://news.microsoft.com/de-de/ki-im-einsatz-gegen-kinderpornografie/> (Abgerufen: 17. Januar 2020).

Schmalen, L. (2019) https://www.nw.de/nachrichten/zwischen_weser_und_rhein/22594348_Wie-Technik-hilft-Suizide-im-Gefaengnis-zu-vermeiden.html, Neue Westfälische, 11. Oktober. Verfügbar auf: https://www.nw.de/nachrichten/zwischen_weser_und_rhein/22594348_Wie-Technik-hilft-Suizide-im-Gefaengnis-zu-vermeiden.html

Spiegel (2019): KI soll Kinderpornografie schneller erkennen Verfügbar auf: <https://www.spiegel.de/netzwelt/netzpolitik/nrw-kuenstliche-intelligenz-soll-kinderpornografie-schneller-erkennen-a-1280480.html>

Stappenbeck, Jens (2017): Frühwarnung weiterdenken: Errichtung einer digitalen Vorhersageplattform <https://peacelab.blog/2017/03/fruehwarnung-weiter-denken-errichtung-einer-digitalen-vorhersageplattform>

Stiftung Neue Verantwortung (2019) Stiftung Neue Verantwortung. Verfügbar auf: <https://www.stiftung-nv.de/en> (Abgerufen: 17. Januar 2020).

Systems, F. (2019) Fusion Systems. Verfügbar auf: <https://www.fusionsystems.de/home.html> (Abgerufen: 17. Januar 2020).

Telemed5000 (2019) Telemed5000. Verfügbar auf: <https://www.telemed5000.de/> (Abgerufen: 17. Januar 2020).

Thüer, L., Köver, C. and Fanta, A. (2018) Asylverfahren: Handy-Durchsuchung bringt keine Vorteile, Netzpolitik.org. Verfügbar auf: <https://netzpolitik.org/2018/asylverfahren-handy-durchsuchung-bringt-keine-vorteile/> (Abgerufen: 17. Januar 2020).

TU Aachen (2019) Algorithmic Accountability. Verfügbar auf: <http://aalab.informatik.uni-kl.de/en/projekte/> (Abgerufen: 17. Januar 2020).

TUM (2019) A Unique Profile – Bringing Politics & Technology Together. Verfügbar auf: <https://www.hfp.tum.de/en/home/> (Abgerufen: 17. Januar 2020).

Weizenbaum Institut (2019) Research for the Networked Society. Verfügbar auf: <https://www.weizenbaum-institut.de/index.php?id=1&L=5> (Abgerufen: 17. Januar 2020).

Team

/ Beate Autering

Gestaltung und Layout



Beate Autering ist freiberufliche Diplomdesignerin und ist Mitbetreiberin der Ateliergemeinschaft beworx. Neben der Erstellung von Designs, Grafiken und Illustrationen bietet sie Bildbearbeitung und Postproduktion an.

/ Fabio Chiusi

Projektmanager und Mitherausgeber des Reports sowie **Autor der Einleitung und des EU-Kapitels**



Fabio ist Projektmanager des Automating Society Reports 2020. Nach zehn Jahren im Bereich Tech Reporting arbeitete er als Berater und wissenschaftlicher Mitarbeiter in den Bereichen Daten und Politik (Tactical Tech) und KI im Journalismus (Polis LSE). Er koordinierte den Bericht

„Persuasori Social“ über die Regulierung politischer Kampagnen in den sozialen Medien für das Projekt PuntoZero und arbeitete während der laufenden Legislaturperiode als technikpolitischer Mitarbeiter in der Abgeordnetenkammer des italienischen Parlaments. Als Fellow am Nexa Center for Internet & Society in Turin ist er Lehrbeauftragter an der Universität San Marino und unterrichtet die Fächer „Journalismus und Neue Medien“ sowie „Verlagswesen und Digitale Medien“. Er ist Autor mehrerer Essays über Technologie und Gesellschaft, zuletzt 'lo non sono qui. Visioni e inquietudini da un futuro presente' (DeA Planeta, 2018), das derzeit ins Polnische und Chinesische übersetzt wird. Er schreibt als Tech-Policy-Reporter beim Kollektiv-Blog ValigiaBlu.

/ Samuel Daveti

Co-Autor des **Comics**



Samuel Daveti ist ein Gründungsmitglied der Kulturvereinigung „Double Shot“. Er ist der Autor des französischsprachigen Graphic Novels Akron Le guerrier (Soleil, 2009) und er ist Kurator der Anthologie Fascia Protetta (Double Shot, 2009). Im Jahr 2011 wurde er Gründungsmitglied des Comic-Kollektivs Mammaiuto. Samuel schrieb auch Un Lungo Cammino (Mammaiuto, 2014; Shockdom, 2017), als Film für das Medienunternehmen Brandon Box veröffentlicht wird. 2018 schrieb er The Three Dogs, mit Zeichnungen von Laura Camelli, das den Micheluzzi-Preis auf der Napoli Comicon 2018 und den Boscarato Preis für den besten Webcomic auf dem Treviso Comic Book Festival.

/ Sarah Fischer

Mitherausgeberin



Sarah Fischer ist als Projektmanagerin für das Projekt „Ethik der Algorithmen“ der Bertelsmann Stiftung tätig, wo sie vor allem Linie für die wissenschaftlichen Studien verantwortlich zeichnet. Zuvor arbeitete sie Postdoctoral Fellow am Graduiertenkolleg „Vertrauen und Kommunikation in einer digitalisierten Welt“ an der Universität Münster, wo sie sich insbesondere mit dem Thema „Vertrauen in Suchmaschinen“ beschäftigte. Ebenfalls im Rahmen dieses Graduiertenkollegs erwarb sie mit einer Promotion über Vertrauen in Gesundheitsdienstleistungen im Internet ihren Dokortitel. Sie studierte Kommunikationswissenschaften an der Friedrich-Schiller-Universität Jena und ist Co-Autorin der Aufsätze „Wo Maschinen irren können. Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung“ und „Was Deutschland über Algorithmen weiß und denkt“.

/ Leonard Haas

Redaktionsassistent



Leonard Haas arbeitet als wissenschaftlicher Mitarbeiter bei AlgorithmWatch und war unter anderem die Konzeption, Weiterentwicklung und Pflege des Projekts AI Ethics Guidelines Global Inventory mitverantwortlich. Er studiert im Master Sozialwissenschaften an der Humboldt Universität Berlin und hält zwei Bachelorabschlüsse von der Universität Leipzig in Digital Humanities und Politikwissenschaften. Sein Forschungsschwerpunkt bildet die Automatisierung der Arbeit und des Regierens. Außerdem interessiert er sich für Allgemeinwohl-orientierte Datenpolitik und Arbeitskämpfe in der Tech-Branche.

/ Graham Holliday

Lektor der englischsprachigen Gesamtausgabe



Foto: Josh White

Graham Holliday ist freiberuflicher Redakteur, Autor und Journalismus-Trainer. Er war fast zwanzig Jahre lang in diversen Positionen für die BBC tätig und außerdem Korrespondent für Reuters in Ruanda. Derzeit ist er Redakteur für die CNN-Sendungen „Parts Unknown“ und „Roads & Kingdoms“. Die ersten beiden von ihm verfassten Bücher wurden von dem inzwischen verstorbenen Anthony Bourdain herausgebracht und erhielten unter anderem Rezensionen in der *New York Times*, der *Los Angeles Times*, dem *Wall Street Journal*, *Publisher's Weekly*, dem *Library Journal* und auf National Public Radio.

/ Nikolas Kayser-Bril

Mitherausgeber und Autor der Story



Foto: Julia Bornkessel

Nicolas ist Datenjournalist und arbeitet für AlgorithmWatch als Reporter. Er war Wegbereiter für neue Formen des Journalismus in Frankreich und Europa und ist einer der führenden Experten für Datenjournalismus. Er hält regelmäßig Vorträge auf internationalen Konferenzen, unterrichtet Journalismus an französischen Journalismusschulen und gibt Schulungen in Redaktionen. Als autodidaktischer Journalist und Entwickler (und Absolvent der Wirt-

schaftswissenschaften) entwickelte er 2009 zunächst kleine interaktive, datengesteuerte Anwendungen für Le Monde in Paris. Anschließend baute er 2010 das Datenjournalismus-Team bei OWNI auf, bevor er von 2011 bis 2017 Journalism++ mitbegründete und leitete. Nicolas ist auch einer der Hauptverfasser des *Datajournalism Handbook*, dem Nachschlagewerk für Datenjournalismus.

/ Lorenzo Palloni

Co-Autor des Comics



Lorenzo Palloni, mehrfach ausgezeichneter Cartoonist, ist Autor mehrerer Graphic Novels und Webcomics und Mitbegründer des Comic-Künstlerkollektivs Mammaiuto. Derzeit arbeitet er an diversen Titeln für den französischen und italienischen Buchmarkt. Er ist außerdem

Lehrer für Scriptwriting und Storytelling an der Scuola Internazionale di Comics di Reggio Emilia (Internationale Comic-Schule in Reggio Emiliana).

/ Kristina Penner

Co-Autorin des EU-Kapitels



Foto: Julia Bornkessel

Kristina Penner ist Referentin der Geschäftsführung bei AlgorithmWatch. Ihre Forschungsinteressen umfassen ADM in Sozialsystemen, Social Scoring und die gesellschaftlichen Auswirkungen von ADM, sowie die Nachhaltigkeit neuer Technologien aus einer ganzheitlichen

Perspektive. Ihre Analyse des EU-Grenzverwaltungssystems baut auf ihrer bisherigen Erfahrung in der Forschung und Beratung zum Asylrecht auf. Ihre bisherigen Erfahrungen umfassen Projekte zu konfliktsensitivem Journalismus und zur Nutzung von Medien in der Zivilgesellschaft und sowie zur Beteiligung von Interessengruppen an Friedensprozessen auf den Philippinen. Sie hat einen Master-Abschluss in International Studies / Peace and Conflict Research von der Goethe-Universität in Frankfurt.

/ Alessio Ravazzani

Co-Autor des **Comics**



Alessio Ravazzani ist redaktioneller Grafikdesigner, Cartoonist und Illustrator. Er arbeitet für die renommiertesten Comic- und Graphic-Novel-Verlage Italiens und gehört zu den Autoren und Gründungsmitgliedern des Kollektivs Mammaiuto.

/ Friederike Reinhold

Mitarbeit an der **Einleitung**, einschließlich der **Handlungsempfehlungen**



Foto: Julia Bornkessel

Friederike Reinhold ist als Senior Policy Advisor von AlgorithmWatch verantwortlich für die Weiterentwicklung des Policy- und Advocacy-Bereichs. Vor ihrer Tätigkeit bei AlgorithmWatch arbeitete Friederike als Referentin im Auswärtigen Amt, für den Norwegian Refugee Council (NRC) im Iran, für die Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) in Afghanistan sowie am Wissenschaftszentrum Berlin für Sozialforschung (WZB).

/ Matthias Spielkamp

Mitherausgeber



Foto: Julia Bornkessel

Matthias Spielkamp ist Mitgründer und Geschäftsführer der Organisation AlgorithmWatch, die mit der Theodor-Heuss-Medaille ausgezeichnet und für einen Grimme Online Award nominiert wurde. Er war Sachverständiger in Anhörungen des Europarats, der EU-Parlaments, des

Bundestags und ist Mitglied der Global Partnership on Artificial Intelligence (GPAI). Matthias ist Vorstandsmitglied bei Reporter ohne Grenzen, Mitglied des Kuratoriums der Stiftung Warentest und im Beirat des Whistleblower Netzwerks. Er ist Autor und Herausgeber von Büchern zu Algorithmen, KI und Automatisierung, Internet Governance, der Zukunft des Journalismus und des Urheberrechts. Seine journalistischen Beiträge sind in MIT Technology Review, Die Zeit, brand eins und vielen anderen Publikationen erschienen. MA Philosophie, FU Berlin, MA Journalismus, University of Colorado at Boulder.

/ Marc Thümmler

Koordination



Foto: Julia Bornkessel

Marc Thümmler ist bei AlgorithmWatch für die Bereiche Öffentlichkeitsarbeit und Outreach verantwortlich. Er hat einen M.A. in Medienwissenschaft, arbeitete als Produzent und Redakteur für die Produktionsfirma Filmgestalten und betreute Projekte für die Deutsche Kinemathek und die zivilgesellschaftliche Organisation Gesicht Zeigen. Zusätzlich zu seinen Kernaufgaben bei AlgorithmWatch war Marc an der Crowdfunding- und Crowdsourcing-Kampagne OpenSCHUFA beteiligt und koordinierte die erste Ausgabe des Automating Society Reports, die 2019 veröffentlicht wurde.

/ Louisa Well

Autorin des **Forschungskapitels Deutschland**



Louisa Well ist wissenschaftliche Mitarbeiterin von Dr. Anna Christmann, MdB, der Sprecherin für Innovations- und Technologiepolitik der Parlamentarischen Gruppe Bündnis 90/Die Grünen im Deutschen Bundestag. In dieser Position erarbeitet Louisa Politikstrategien in den Bereichen KI und Digitalisierung. Sie hat einen Master-Abschluss in Science and Technology in Society der Universität Edinburgh sowie einen Abschluss als Bachelor of Arts in Politikwissenschaft und englischer Literatur der Universität Heidelberg. Für ihre Masterarbeit – in deren Mittelpunkt die Frage stand, ob Automatisierung in Deutschland zu einer verstärkten Diskriminierung bei der automatisierten Zuweisung von Arbeitslosengeld II führt – absolvierte sie ein Forschungspraktikum bei AlgorithmWatch.

ORGANISATIONEN

/ AlgorithmWatch

AlgorithmWatch ist eine gemeinnützige Forschungs- und Advocacy-Organisation mit dem Ziel, Systeme algorithmischer / automatisierter Entscheidungsfindung (ADM) und deren Auswirkungen auf die Gesellschaft zu beobachten und zu analysieren. Um menschliche Selbstbestimmung und Grundrechte zu schützen sowie das Gemeinwohl zu maximieren, halten wir es für entscheidend, ADM-Systeme zur Rechenschaft zu ziehen und sie einer demokratischen Kontrolle zu unterwerfen. Der Einsatz von ADM-Systemen, die wesentlich individuelle und kollektive Rechte beeinträchtigen, muss nicht nur auf klare und zugängliche Weise öffentlich gemacht werden, Einzelpersonen müssen auch in der Lage sein zu verstehen, wie Entscheidungen zu zustandekommen und sie gegebenenfalls angefochten werden können. Deshalb befähigen wir die Bürgert:innen, ADM-Systeme besser zu verstehen und Wege zu entwickeln diese Prozesse demokratische zu kontrollieren - mit einer Mischung aus Technologien, Regulierung und geeigneten Aufsichtsinstitutionen. Damit streben wir an zu einer fairen und inklusiven Gesellschaft beizutragen und den Nutzen von ADM-Systemen für die Gesellschaft zu maximieren für die gesamte Gesellschaft zu maximieren. Damit wollen wir zu einer fairen und inklusiven Gesellschaft beitragen und den Nutzen von ADM-Systemen für die Gesellschaft insgesamt zu maximieren.

<https://algorithmwatch.org>



/ Bertelsmann Stiftung

Die Bertelsmann Stiftung setzt sich für eine gerechte Teilhabe aller am gesellschaftlichen Leben ein. Sie engagiert sich in den Bereichen Bildung, Demokratie, Gesellschaft, Gesundheit, Kultur und Wirtschaft. Durch ihr Engagement will sie alle Bürgerinnen und Bürger ermutigen, sich für das Gemeinwohl einzusetzen. Die 1977 von Reinhard Mohn gegründete, gemeinnützige Einrichtung hält die Mehrheit der Kapitalanteile der Bertelsmann SE & Co. KGaA. Die Bertelsmann Stiftung arbeitet operativ und ist unabhängig vom Unternehmen sowie parteipolitisch neutral. Die Bertelsmann Stiftung setzt sich im Projekt „Ethik der Algorithmen“ mit den gesellschaftlichen Folgen algorithmischer Entscheidungsfindung auseinander. Das Ziel des Projekts ist es, zu einer Gestaltung algorithmischer Systeme beizutragen, die zu mehr Teilhabe für alle führt. Nicht das technisch Mögliche, sondern das gesellschaftlich Sinnvolle muss Leitbild sein – damit maschinelle Entscheidungen den Menschen dienen.

<https://www.bertelsmann-stiftung.de>

| BertelsmannStiftung