H 7625

FIfF Kommunikation

Zeitschrift für Informatik und Gesellschaft

40. Jahrgang 2023 Einzelpreis: 7 EUR

4/2023 – Dezember 2023

„Ich glaube unbedingt daran, dass Wissenschaft und Friede schließlich über Unwissenheit und Krieg triumphieren und die Völker der Erde übereinkommen werden, nicht zu zerstören, sondern aufzubauen."

Louis Pasteur (1822–1895), französischer Chemiker und Mikrobiologe

ISSN 0938-3476

• Künstliche Intelligenz • Konferenz über Nachhaltigkeit und Digitalisierung • netzpolitik.org •

FIfF Kommunikation

Zeitschrift für Informatik und Gesellschaft

Titelbild: *Der trotz mehrer direkter Treffer scheinbar unbeschädigt wirkende Kölner Dom steht inmitten der vollständig zerstörten Innenstadt von Köln. Hinter der Kathedrale ist der zerstörte Hauptbahnhof, so wie weiter östlich die eingestürzte Hohenzollernbrücke zu erkennen. Deutschland, 24. April 1945. Quelle: https://catalog.archives.gov/id/531287*

inhalt

Inhalt

Ausgabe 4/2023

Editorial

Wissenschaft für den Frieden – der Schwerpunkt dieser Ausgabe ist unserer Schwesterzeitschrift *Wissenschaft und Frieden* gewidmet, die in diesem Jahr in Bonn mit einem Symposium und einem Festakt ihr 40-jähriges Jubiläum gefeiert hat. Diese Ausgabe der *FIfF-Kommunikation* enthält die Beiträge des Symposiums mit technischen Schwerpunkten, von denen viele einen Bezug zur Informatik aufweisen.

Hans-Jörg Kreowski hat diesen Schwerpunkt besorgt und führt in seinem Schwerpunkteditorial in die behandelten Themen ein. Er bezieht dabei auch klar Stellung zur aktuellen Entwicklung:

> *Wenn eine seit 40 Jahren bestehende Zeitschrift wie auch das fast so alte FIfF „Frieden" im Namen führen, dann muss man leider konstatieren, dass diese Thematik weiterhin schreckliche Aktualität besitzt, weil Krieg eine alltägliche Realität darstellt, weil Konflikte in vielen Teilen der Welt zu eskalieren drohen und weil Politik regional bis weltweit wenig erfolgreich dagegen arbeitet, wenn sie sich überhaupt darum kümmert. Ein betrübliches, ja skandalöses Beispiel hat gerade der deutsche Verteidigungsminister Boris Pistorius gegeben, der Deutschland und die Bundeswehr „kriegstüchtig" machen möchte. Sein Versuch, diesen Begriff als synonym zu „verteidigungsfähig" hinzustellen, ist irreführend. Denn hätte er „verteidigungsfähig" gemeint, hätte er das ja auch sagen können.*

Auch die aktuelle Ausgabe von Wissenschaft und Frieden enthält Beiträge des Symposiums, so dass die beiden Zeitschriften zusammen gelesen werden können. Das FIfF gratuliert in einem Grußwort *Wissenschaft und Frieden* zum 40-jährigen Jubiläum – wir freuen uns auf die zukünftige Kooperation zu einem Thema, dass (leider) angesichts der Konflikte in der Ukraine und in Palästina an Aktualität wieder zunimmt.

Neben dem Schwerpunkt spielt das Thema *Künstliche Intelligenz* eine große Rolle in dieser Ausgabe. Klaus Fuchs-Kittowski blickt in seinem Statement anlässlich der Tagung *Weizenbaum's Worlds: Technological Change and Computer Criticism in the U.S. and Germany* des Weizenbaum-Instituts für die vernetzte Gesellschaft auf das Leben und Wirken Weizenbaums zurück. Joseph Weizenbaum, der mit *Eliza* einen frühen Chatbot entwickelte, der nach Ansicht von Psychologen für therapeutische Gespräche eingesetzt werden könne, stand der Künstlichen Intelligenz stets kritisch gegenüber: Dem Computer sei das für den Menschen wesentliche – das Gefühl – fremd und bleibe ihm fremd:

> *Hier sagt nun Weizenbaum: Das geht nicht, das ist Betrug! Denn um heilen zu können, bedarf es eines Verständnisses der konkreten Situation des Kranken. Ein solches Verständnis hat der Computer nicht und kann es auch schon deshalb nicht haben, weil er kein Gefühl besitzt. Wo es um die Beurteilung komplexer Lebenssituationen geht, führen Berechnungen eher in die Irre. Daher der Titel der Originalausgabe seines Buches: From Judgement to Calculation.*

Daran hat sich, so Fuchs-Kittowski, auch im Zeitalter von ChatGPT nichts geändert:

> *Aber, wie Joe Weizenbaum von Beginn der KI-Forschung an versucht hat nachzuweisen, bedeutet dies keineswegs die Entmachtung des Menschen oder sogar die Verdrängung der Menschheit. Denn schöpferisches Denken, wirklich neue Informationen und Wissen schaffen, können die KI-Systeme nicht.*

Zwei weitere Beiträge drucken wir mit freundlicher Genehmigung der Redaktion der *vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik*, herausgegeben von der Humanistischen Union. Stefan Hügel bewertet Chancen und Risiken der Künstlichen Intelligenz. Er beginnt mit der Frage, was (künstliche) Intelligenz ist, fasst kurz die Basis für Maschinelles Lernen zusammen und skizziert Risiken – einschließlich einiger Statements zur Ersetzung des Menschen durch Künstliche Intelligenz. Im Hauptteil benennt er eine Reihe konkreter ethischer Fragen, die beim Einsatz von Verfahren der Künstlichen Intelligenz beachtet werden müssen. Hans-Jörg Kreowski und Aaron Lye gehen konkret darauf ein, wie Künstliche Intelligenz im militärischen Bereich eingesetzt wird und benennen die aktuellen Entwicklungen und Risiken.

In einem weiteren Bericht wird unser Projekt *Citizen Science* angekündigt, das Nachfolgeprojekt von *TDRM – Tihange Doel Radiation Monitoring*. Wir drucken den offenen Brief zur Reform der Digitalen Identität in der Europäischen Union – eIDAS –, und Dagmar Boedicker berichtet von der Konferenz über Nachhaltigkeit und Digitalisierung an der Universität Augsburg, in der zwei übergreifende Frage verhandelt wurden:

- Wie verändert die Digitalisierung Erkenntnisse und Praktiken der Umwelt- und Nachhaltigkeits-Politik?
- Wie sollte der Rahmen für ihre Regulierung und Handhabung aussehen, um ihr Potenzial in den Dienst der Nachhaltigkeits-Transformation zu stellen?

Einige Beiträge zur Netzpolitik runden die Ausgabe ab: Zum Recht am eigenen Bild, zur Sicherheit bei der Digitalisierung von Gesundheitsdaten im Rahmen der Einführung der elektronischen Patientenakte, zur Chatkontrolle, zu den Plänen der neuen „großen" Koalition nach den Landtagswahlen in Hessen und zu eIDAS – einer *Digitalen Brieftasche mit Ausspähgarantie*.

Dies ist die letzte Ausgabe der *FIfF-Kommunikation*, die wir bei *Meiners Druck* in Bremen im Offset drucken lassen, da der Bereich des Offsetdruckes geschlossen wird. Wir danken Herrn Meiners für die langjährige sehr gute Zusammenarbeit.

Wir wünschen unseren Leserinnen und Lesern eine interessante und anregende Lektüre – und viele neue Erkenntnisse und Einsichten.

Stefan Hügel
für die Redaktion

Der Brief

„Kriegstüchtig“

Liebe Freundinnen und Freunde des FIfF, liebe Mitglieder,

und, seid Ihr schon *kriegstüchtig*?

Blickt man auf die Ereignisse der letzten Zeit zurück, scheint nun endgültig die Welt aus den Fugen zu geraten. Eine weltweite Pandemie glauben wir gerade überwunden zu haben, in Europa finden zwei furchtbare Kriege statt – in der Folge ruft unsere Bundesregierung die militärische „Zeitenwende“ aus. Der Klimawandel schreitet fort und scheint die ersten Kipppunkte zu erreichen – wenn er sie nicht schon überschritten hat. Mit ChatGPT und *Large Language Models* erreicht die sogenannte Künstliche Intelligenz als „generative“ KI eine neue Stufe, und auch hier macht sich Endzeitstimmung breit.

> *Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war*[1]

schreiben Bill Gates, Sam Altman, Geoffrey Hinton, Yoshua Bengio, Audrey Tang und Vitalik Buterin und lenken damit geschickt von den tatsächlichen aktuellen Gefahren von KI ab: Strukturveränderung des Arbeitsmarktes, die Ausweitung ausbeuterischer (Datenerhebungs- und Auswertungs-) Lieferketten – insbesondere im globalen Süden –, die einseitige Macht- und Produktivitätssteigerung weniger Firmen, die Ausweitung personalisierter Überwachung, die unendliche Welle von KI-generiertem Spam und Betrug, die automatisierte Aneignung und Kommerzialisierung von Online-Kulturwerken, die Herausbildung globaler Abhängigkeiten durch Oligopole in der Industrie, der unreflektierte und nur scheinbar unpolitische Einsatz von KI-Systemen in Sozialsystemen und weiteren sensiblen Gesellschaftsbereichen und nicht zuletzt Manipulation durch *AI-enhanced Nudging*.

„Nudge – so heißt die Formel, mit der man andere dazu bringt, die richtigen Entscheidungen zu treffen.“[2] Genannt werden die Altersvorsorge, umweltbewusstes Verhalten und gesunde Ernährung. Da kann doch niemand etwas dagegen haben, oder?[3] Und so wurde bereits von der Regierung unter Bundeskanzlerin Angela Merkel berichtet, sie wolle solche Techniken zum Nutzen der deutschen Bevölkerung einsetzen.[4]

„*… die richtigen Entscheidungen zu treffen.*“ Also zum Beispiel: Brav den Müll zu trennen und weniger Auto zu fahren. Oder auch: Die Verschiebung der deutschen Sicherheitspolitik hin zum Militärischen für völlig normal und gerechtfertigt zu halten.

Ein paar Beispiele aus der Politik in den letzten Monaten: In Dauerschleife beschwören führende Sicherheitspolitiker:innen eine militärische „Zeitenwende“ und setzen damit ein „Europa der Verteidigung und der Rüstung“, so Bundeskanzler Scholz[5] und „Kriegstüchtigkeit“, so Verteidigungsminister Pistorius[6], als neue Leitbilder der deutschen und europäischen Außen- und Sicherheitspolitik:

> *Wir müssen uns wieder an den Gedanken gewöhnen, dass die Gefahr eines Krieges in Europa drohen könnte.*
>
> *Und das heißt: Wir müssen kriegstüchtig werden. Wir müssen wehrhaft sein. Und die Bundeswehr* ***und die Gesellschaft*** *dafür aufstellen.*

Da nicht alle mitmachen wollen, beklagt man bereits „Kriegsmüdigkeit“, so Bundesaußenministerin Baerbock.[7] Sicher ist es auch nur ein Versehen der Parteitagsregie, dass das Motto des letzten Bundesparteitags der Grünen – „Machen, was zählt“[8] – sehr an einen Slogan der Bundeswehr – „Mach, was wirklich zählt“[9] – erinnert. Es fehlt dann nur noch die Vorsitzende des Verteidigungsausschusses des Deutschen Bundestages, Marie-Agnes Strack-Zimmermann[10], die unter Anderem fordert, Russland für die Bundeswehr als Feindbild aufzubauen.[11] „Top-Gun-Feeling pur“, jubelte sie laut Medienberichten nach einem Flug im Eurofighter der Bundeswehr[12], bei dem „ein Traum für sie in Erfüllung gegangen sei“ – ob das wohl die Bundeswehrpilot:innen im realen Einsatz, wenn sie nicht wissen, ob sie wieder zurückkehren werden, auch so sehen?

Das sind alles nur Kleinigkeiten, vielleicht ohne Bedeutung. Vielleicht aber auch: Nudging – kleine Anstöße für die Diskursverschiebung hin zu einer angstgetriebenen Politik, die auf eine Militarisierung der Außen- und Sicherheitspolitik abzielt.

Aktuell bestimmen vor allem zwei Themen die tagespolitischen Nachrichten: Der Terrorangriff der Hamas auf Israel sowie seine Folgen im nahen Osten und die Entscheidung des Bundesverfassungsgerichts[13], dass der Nachtragshaushalt 2021 und damit der Klima- und Transformationsfonds in Höhe von 60 Mrd. € gegen die im Grundgesetz festgelegte „Schuldenbremse“ verstößt – die sich damit erneut als fatale Zukunftsbremse erweist. Man kann nun mit Sorge betrachten, dass erneut das Bundesverfassungsgericht einen verfassungswidrigen parlamentarischen Beschluss korrigieren musste – gefühlt kommt dies in letzter Zeit regelmäßig vor.[14] Der für den verfassungswidrigen Haushalt verantwortliche Minister Lindner ist jedoch immer noch im Amt. Man kann aber ebenso besorgt sein über die Prioritätensetzung der Bundespolitik: Offensichtlich möchte man für den existenziell notwendigen Klimaschutz keine Mehrheiten organisieren, so dass man zu Buchungstricks – hier die „Umwidmung“ von Mitteln, die für die Corona-Krise vorgesehen waren, für Maßnahmen des Klimaschutzes[15] – greifen muss. Auch hier scheinen die Prioritäten klar: Das „Sondervermögen“ von 100 Mrd. € für die Bundeswehr wurde mit den notwendigen parlamentarischen Mehrheiten direkt in die Verfassung geschrieben[16] und ist damit von dem Beschluss des Bundesverfassungsgerichts nicht betroffen. Die jährlichen 65 Mrd. für fossile Subventionen wie das Dienstwagen- oder Dieselprivileg werden ebenfalls nicht angefasst.[17] Lieber werden die Kindergrundsicherung oder das 49-Euro-Ticket in Frage gestellt.

Der terroristische Angriff der Hamas auf Israel, die Ermordung und Geiselnahme vieler Menschen ist ein brutaler terroristischer Akt von großer Abscheulichkeit und immenser politischer Tragweite. Dass der Staat Israel das Recht hat, sich gegen diese Angriffe zu verteidigen, ist für uns selbstverständlich. Über die Form hingegen kann trefflich gestritten werden.

Darum sind wir entsetzt über die vielen zivilen Opfer des Konflikts in Gaza, knapp die Häfte davon Minderjährige.[18] Erneut bestätigt sich, dass die Bevölkerung auf beiden Seiten zu den Opfern eines Krieges wird, den politische Funktionseliten ausgelöst haben. Diese Gewaltspirale muss gestoppt werden. Im Einklang mit den Vereinten Nationen fordern wir einen Waffenstillstand und Verhandlungen über eine friedliche politische Lösung des Konflikts.[19] Die Freilassung von Geiseln ist ein erster wichtiger Schritt auf diesem Weg, aber nur ein erster.

Eins muss dabei völlig klar sein: Das Existenzrecht des Staates Israel ist nicht verhandelbar und darf nicht infrage gestellt werden, die aktuelle Regierung hingegen muss sich – wie die jedes demokratischen Landes – opponierenden Meinungen stellen. Wir sind natürlich erschüttert angesichts der antisemitischen Übergriffe in Deutschland und weltweit. Wir sind uns unserer historischen Verantwortung aufgrund der nationalsozialistischen Schreckensherrschaft bewusst. Die Übergriffe werden mehrheitlich von rechtsgerichteten Deutschen begangen – aber nicht nur –, doch auch die Solidarität mit den Menschen in Gaza rechtfertigt keinen Antisemitismus. Es ist aber ebenso nicht akzeptabel, Übergriffe allein Muslimen in die Schuhe schieben zu wollen. Heimischen Antisemitismus haben wir in Deutschland im Überfluss und auch nicht erst seit dem 7. Oktober 2023. In Deutschland lebende Muslime müssen sich nicht explizit von der Hamas distanzieren – ebenso, wie wir Deutschen uns nicht explizit von den Morden des NSU oder der RAF distanzieren müssen. Dass solche rassistischen Forderungen erhoben werden[20] – und anscheinend breite Zustimmung finden[21] – ist verstörend. Richtig ist gleichzeitig, dass Antisemitismus nicht toleriert werden darf. Wir müssen uns dabei immer bewusst sein, dass es unsere deutschen Vorfahren waren, die für die Shoah und die systematische Ermordung und damit die furchtbarsten Angriffe auf jüdisches Leben in der Geschichte verantwortlich sind.

All dies lässt schon erwarten, dass die Debatten über den Gaza-Konflikt stark polarisiert ablaufen, oft ahistorisch im Jetzt geführt werden und manche Äußerungen schwer zu ertragen sind. Doch Menschenrechte sind stets universell:

Wir können die Hamas und ihre hinterhältigen Angriffe verurteilen und gleichzeitig für die Freiheit des palästinensischen Volkes sein, das Verteidigungsrecht Israels hochhalten und gleichzeitig Israels Bestrafungs- und Vergeltungsangriffe auf zivile Ziele verurteilen, die palästinensische Unterstützung der Hamas kritisieren und gleichzeitig die unmenschliche Belagerung Gazas verurteilen, Netanjahu und seine Minister als Rechtsradikale kritisieren und gleichzeitig Israel als Demokratie ansehen, palästinensischen Antisemitismus thematisieren und gleichzeitig die Unterstützung radikaler Siedler durch Israels Regierung und Armee anprangern, innerpalästinensische Misswirtschaft aufzeigen und gleichzeitig Autonomiebestrebungen stützen, Israels Politik gegenüber den Palästinenser:innen kritisieren und gleichzeitig die politisch-progressiven Kräfte in Israel dagegen anerkennen. Das geht alles gleichzeitig und ist konsistent, wenn wir nicht versuchen, demokratische Werte, Menschenrechte und Dekolonialismus gegeneinander auszuspielen.

Was hat das mit dem FIfF zu tun? Sehr viel: Als die friedenspolitische, nachhaltige Informatik-Initiative, die wir sind, sind das offensichtlich alles Themen, die uns betreffen und mit denen wir uns befassen. Wir glauben natürlich nicht, dass wir die Probleme dieser Welt lösen können, aber wir wollen unseren Teil dazu beitragen.

Mit FIfFigen Grüßen
Stefan Hügel und Rainer Rehak

Anmerkungen

1 *Center for AI Safety (2023) Statement on AI Risk, https://www.safe.ai/statement-on-ai-risk*
2 *Thaler RH, Sunstein CR (2015 [2008]) Nudge. Wie man kluge Entscheidungen anstößt. Berlin: Ullstein*
3 *Dass damit notwendige politische Entscheidungen individualisiert werden – unsere diesjährige Weizenbaum-Preisträgerin befasste sich mit dem Prozess der neoliberalen Responsibilisierung –, lassen wir für den Moment mal beiseite.*
4 *Klug T (2015) Wie Frau Merkel uns hilft, die bessere Wahl zu treffen. Deutschlandfunk Kultur, https://www.deutschlandfunkkultur.de/nudging-wie-frau-merkel-uns-hilft-die-bessere-wahl-zu-100.html*
5 *Rede des Bundeskanzlers Olaf Scholz bei der Münchener Sicherheitskonferenz 2023. https://www.bundeskanzler.de/bk-de/aktuelles/rede-von-bundeskanzler-scholz-anlaesslich-der-munich-securityconference-am-17-februar-2023-in-muenchen-2166452*
6 *„Wir müssen kriegstüchtig werden". Spiegel online, https://www.spiegel.de/politik/boris-pistorius-ueber-die-bundeswehr-wir-muessen-kriegstuechtig-werden-a-24366e1b-6689-4ee0-a4c0-d0c7ee0f854d*
7 *Baerbock warnt vor Kriegsmüdigkeit. 25. Mai 2022, taz.de, https://taz.de/-Nachrichten-zum-Ukrainekrieg-/!5857171/*
8 *Stieber B (2023) Willkommen, graue Wirklichkeit. taz.de, https://taz.de/Parteitag-der-Gruenen/!5975207/*
9 *Kampagne der Bundeswehr, z. B. bei Youtube: https://www.youtube.com/playlist?list=PLRoiDADf6licHjrZVs-_9L1Tft0g1jxP8 oder kritisch hier: https://augengeradeaus.net/2015/11/marke-bundeswehr-militaerische-fachkraft-fuer-frieden-und-freiheit/comment-page-1/ und hier: https://pen.gg/machwaszaehlt/*
10 *Wikipedia, Stichwort Marie-Agnes Strack-Zimmermann, https://de.wikipedia.org/wiki/Marie-Agnes_Strack-Zimmermann*
11 *Strack-Zimmermann: Bundeswehr braucht ein Feindbild. 31. Mai 2022, dpa-Newskanal zit. nach sueddeutsche.de, https://www.sueddeutsche.de/politik/verteidigung-strack-zimmermann-bundeswehr-braucht-ein-feindbild-dpa.urn-newsml-dpa-com-20090101-220531-99-497796*
12 *Zippel T (2023) Abgehoben im Kampfjet: Bundeswehr führt Sonderflug für FDP-Politikerin durch. Ostthüringer Zeitung, https://www.otz.de/politik/abgehoben-im-kampfjet-bundeswehr-fuehrt-sonderflug-fuer-fdp-politikerin-durch-id239341153.html*
13 *Bundesverfassungsgericht erklärt Nachtragshaushalt 2021 für verfassungwidrig. Spiegel online, https://www.spiegel.de/politik/deutschland/bundesverfassungsgericht-erklaert-nachtragshaushalt-2021-fuer-verfassungswidrig-a-f856c299-d886-4370-a59c-65fdaa054315*
14 *Zum Beispiel bei der Vorratsdatenspeicherung: Hügel S (2021) „Und täglich grüßt das Murmeltier." Die andauernde Debatte um die Vor-*

ratsdatenspeicherung oder: Politik gegen die Grundrechte. Grundrechte-Report 2021, Frankfurt am Main: S. Fischer, S. 52ff

15 *Die Diskussion der Frage, ob alle Maßnahmen, die mit dem Fonds finanziert werden sollten, tatsächlich im Sinne des Klimaschutzes sind, sprengt den Rahmen dieser Kolumne.*

16 *Artikel 87a (1a) GG, https://www.gesetze-im-internet.de/gg/art_87a.html*

17 *Umweltbundesamt (2021) Umweltschädliche Subventionen: fast die Hälfte für Straßen- und Flugverkehr, https://www.umweltbundesamt.de/presse/pressemitteilungen/umweltschaedliche-subventionen-fast-die-haelfte*

18 *Vereinte Nationen (2023) UN agency heads unite in urgent plea for women and children in Gaza, https://news.un.org/en/story/2023/11/1143877*

19 *Vereinte Nationen (2023) UN General Assembly adopts Gaza resolution calling for immediate and sustained 'humanitarian truce'. Vereinte Nationen, https://news.un.org/en/story/2023/10/1142847#:~:text=The%20resolution%20calls%20for%20an,services%20into%20the%20Gaza%20Strip, und der Bericht in https://www.theguardian.com/world/2023/oct/27/israel-gaza-war-un-general-assembly-call-immediate-durable-humanitarian-truce; dazu auch Shaul Y (2023) Israel/Palästina – Geplatzte Blase, https://www.medico.de/blog/geplatzte-blase-19279*

20 *Fischer T (2023) Was müssen die Muslime Robert Habeck beweisen? Spiegel online, https://www.spiegel.de/kultur/krieg-in-nahost-was-muessen-die-muslime-robert-habeck-beweisen-a-24ad17d6-e5a0-4342-986d-d37c0e52a5dd*

21 *Appell von Vizekanzler Habeck: „Antisemitismus ist in keiner Gestalt zu tolerieren". Tagesschau.de, https://www.tagesschau.de/inland/gesellschaft/habeck-antisemitismus-100.html*

Das FIfF bittet um Eure Unterstützung

Viermal im Jahr geben wir die FIfF-Kommunikation heraus. Sie entsteht durch viel ehrenamtliche, unbezahlte Arbeit. Doch ihre Herstellung kostet auch Geld – Geld, das wir nur durch Eure Mitgliedsbeiträge und Spenden aufbringen können.

Auch unsere weitere politische Arbeit kostet Geld für Öffentlichkeitsarbeit, Aktionen und Organisation. Dazu gehören unsere jährlich stattfindende FIfF-Konferenz, der Weizenbaum-Preis, weitere Publikationen und die Kommunikation im Web: Neben der tatkräftigen Mitwirkung engagierter Menschen sind wir bei unserer Arbeit auf finanzielle Unterstützung angewiesen.

Bitte unterstützt das FIfF mit einer Spende. So können wir die öffentliche Wahrnehmung für die Themen weiter verstärken, die Euch und uns wichtig sind.

Spendenkonto:
Bank für Sozialwirtschaft (BFS) Köln
IBAN: DE79 3702 0500 0001 3828 03
BIC: BFSWDE33XXX

Klaus Fuchs-Kittowski

Gegen die Macht der Computer und die Zerstörung der Vernunft – Für die Entfaltung der Kreativität und Beurteilungsfähigkeit der Menschen

Joseph Weizenbaum – ein kritischer Wissenschaftler par excellence

Diskussionsbeitrag zur Tagung: Weizenbaum's Worlds: Technological Change and Computer Criticism in the U.S. and Germany, ca. 1960-1990 im Weizenbaum-Institut am 3.-4. November 2023

Die erste Begegnung mit Joseph Weizenbaum

Mit Joseph Weizenbaum traf ich das erste Mal auf der IFIP-Konferenz *Human Choice and Computer II* in Baden bei Wien 1979 zusammen.

Das Zusammentreffen war nicht ganz unerwartet, denn ich hatte sein Buch gelesen, auf Empfehlung eines Mitarbeiters der Staatsbibliothek. Dieser rief mich an und sagte, es sei ein neues Buch gekommen, welches ich mir holen solle, bevor es in den Durchlauf zur Registration kommt. Darin würde vieles so gesagt wie in meinen Vorlesungen. Der Name Weizenbaum war mir auch schon zuvor genannt worden, von dem Kölner Molekularbiologen Benno Müller Hill. Auf Einladung von Prof. Rapoport hatten wir gemeinsam über sein Buch *Die Biologie und die Philosophie* diskutiert, in dem er darstellt, wie sich rassistisches Denken, beginnend mit Plato bis zu Ernst Häckel durch die Biologie zieht. Zum Abschluss sagte er zu mir: „Ich war kürzlich in den USA. Ich habe zumindest einen wirklichen Menschen getroffen, einen Professor am MIT, Joe Weizenbaum. Das wäre ein wichtiger Gesprächspartner für Dich."

Und doch war es ganz spontan, dass ich Joseph Weizenbaum gleich nach der Begrüßung an die Humboldt-Universität einlud. Ich hatte noch nie jemanden eingeladen, und auf jeden Fall keinen Amerikaner mitten im Kalten Krieg. Weizenbaum trat einen Schritt zurück und kam dann wieder auf mich zu und sagte: „Die Einladung an die Humboldt-Universität wäre eine Genugtuung für mich." Ich trat einen Schritt zurück und fragte: „Warum?" Er antwortete: „Weißt Du nicht, dass ich aus Berlin komme? Ich musste mich als kleiner Judenjunge immer an dieser Universität vorbeischleichen. Wenn diese Universität mich jetzt einlädt, wäre dies eine Genugtuung für mich."

Im Workshop zu *Computer and Ethics* formuliert J. Weizenbaum seinen minimalen moralischen Imperativ für Computer Scientists:

> *Don`t use computers to do, what people ought not do.*[1]

Zumindest darf man mit dem Computer nichts tun, was man als Mensch auch nicht tun sollte! Wie aktuell dieser moralische Imperativ ist, erleben wir gegenwärtig. Durch die autonomen Waffen, den Einsatz bewaffneter Drohnen, wird die Entfernung vom Töten zu den Grausamkeiten des Kriegsschauplatzes so groß, dass die Hemmschwelle zum Krieg, zum Töten in erschreckender Weise stark herabgesetzt wird.

Seminar zur KI-Kritik mit Joe Weizenbaum an der Humboldt-Universität zu Berlin

Das gemeinsame Seminar mit J. Weizenbaum zu den Problemen seines Buches: *Die Macht des Computers und die Ohnmacht der Vernunft* und unsere prinzipielle Unterscheidung zwischen dem Automaten als Informationstransformator und dem kreativ tätigen, zur Informationserzeugung befähigten Menschen, fand sieben Tage nach Ausbruch des Krieges der Sowjetunion gegen Afghanistan statt.

Weizenbaum erklärte gegenüber dem Prorektor für Gesellschaftswissenschaften:

> *Ich bin hier, nicht weil ich ein besonderer Freund der DDR bin, sondern weil ich ein amerikanischer Patriot bin. Als Patriot bin ich gegen das Wettrüsten, denn es ruiniert unsere Wirtschaft. Es ist, als ob ich einen Bleistift anspitze und in den Papierkorb werfe und das immer wiederhole. Es ist für uns aller Leben kreuzgefährlich!*

Nach seinem Aufenthalt an der Humboldt-Universität ruft mich Joe Weizenbaum aus Zürich in Berlin an und sagt: „Klaus, ich habe gestern vor mehr als 1000 Menschen in einer Kirche in Zürich gesprochen! Das war ganz großartig!"

Weizenbaums internationales Auftreten für Entspannung und Abrüstung hatte Wirkungen bis zu den dramatischen Entscheidungen im Kreis um den Gewandhausdirektor Kurt Masur in Leipzig[2].

Man kann nur so viele Menschen mobilisieren, wenn man als Fachmann im Prinzip alle schrecklichen Entwicklungen für möglich hält und davor mit allem Nachdruck warnt. Wie Joe dies getan hat. So wies er z. B. einmal, als der Weltkirchenrat im MIT tagte, die internationalen Vertreter darauf hin, dass hier gefährliche Waffen konzipiert und entwickelt werden. Dies führte dazu, dass sich der Weltkirchentag erstmals ernsthaft mit der Frage der Abrüstung beschäftigte. Ein junger Theologe aus Leipzig war davon sehr beeindruckt. Es ist daher nicht zufällig, dass er zu den vier Unterstützern des Gewandhausdirigenten Kurt Masur gehörte, die mit ihrem Appell eine „Chinesische Lösung" verhinderten. Wenn also Joseph Weizenbaums zu seinem 100. Geburtstag gedacht wird, so muss auch unbedingt an seinen Einsatz gegen Aufrüstung und für eine allgemeine Abrüstung erinnert werden. Dazu gehört auch die Gründung der Bewegung *Computer Professionals for Social Responsibility* (CPSR)

mit dem berühmten Pionier der KI-Forschung, Terry Winograd, in den USA und des *Forums Informatikerinnen und Informatiker für Frieden und gesellschaftliche Verantwortung* (FIfF) mit der Pionierin der Informatik, Christiane Floyd, in Deutschland.

Missglückter Antrag auf Verleihung des Ehrendoktors an Joseph Weizenbaum

Nach dem mit uns an der Humboldt-Universität durchgeführten Seminar wurde ein Ehrendoktorverfahren für J. Weizenbaum an der TU Berlin eingestellt. Allein die Kontaktaufnahme mit dem Osten war für das Frontstadtdenken offensichtlich unerträglich. Mein Versuch, an der HU Berlin die Verleihung eines Ehrendoktors an J. Weizenbaum zu erreichen, wurde trotz Unterstützung durch namhafte Wissenschaftler, wie den Physiker Robert Rompe, dadurch sabotiert, dass, zumindest von einem Vertreter in der Kommission, Weizenbaums Kritik an der künstlichen Intelligenzforschung, als deren Mitbegründer er gilt, in eine generelle Technologiefeindlichkeit uminterpretiert wurde.

Zum Grundanliegen von Joe Weizenbaum

Ich bin kein KI-Kritiker. Ich bin Gesellschaftskritiker.

Letzter Auftritt vor Studenten im Lichthof der Technischen Universität

Der Computer hat keine Gefühle. Das eigentlich Menschliche muss ihm daher fehlen.

Er verwies auf das Eingeständnis von Marvin Minsky: „Es sei ihm nicht gelungen, seinen Robotern Gefühle beizubringen. Dies sei eben doch eine andere Ebene."

Das ist für mich das eigentliche Vermächtnis des Wirkens von Joseph Weizenbaum, die Grundlage seiner KI- und Gesellschaftskritik. Dem Computer ist das für den Menschen Wesentliche fremd und bleibt ihm fremd: Das Gefühl, so legt er dem früheren Direktor des KI-Labors, Hans Moravec, die Frage vor, wie er wohl glaubt, das Lächeln einer Mutter ihrem Kind gegenüber auf einen Computerspeicher zu bringen, wenn er die Ablösung der menschlichen Gesellschaft durch eine Computergesellschaft für möglich hält und im Namen der modernen Wissenschaft propagiert?[3] Gegen diesen gefährlichen, menschenfeindlichen Reduktionismus kämpft er in seinem berühmt gewordenen Buch *Computer Power and Human Reason. From Judgement to Calculation* an, indem er insbesondere Simon zitiert, der behauptet: „Ameise, Computer und Mensch sind Systeme gleicher Art. Sie sind Informationen verarbeitende Systeme." Diese dem Informationsverarbeitungsansatz bzw. der *Physical Symbol System Hypothese* der KI-Forschung innewohnende Reduktion des Menschen auf den Computer, dieser Reduktionismus als weltanschauliche Haltung ist äußerst gefährlich!

J. Weizenbaum machte in seinem Werk: *Computer Power and Human Reason* klar, dass diese Identifizierung von Automaten und Menschen die Realität verzerrt und man dann dazu verleitet wird, diese Verzerrung als „vollständige und erschöpfende" Darstellung zu akzeptieren. Das ist es, was der Computerwissenschaftler Herbert A. Simon als grundsätzliche theoretische Orientierung beschreibt (Weizenbaum, 1977, S. 176).

Führte man die vorangegangenen Weltkriege im Zeichen des Rassismus, der Reduktion des Menschen auf das Tier, so werden wir heute erleben, dass noch diese Reduktion des Menschen auf das von ihm Gemachte, auf die Maschine, dazukommt.

Ich bin kein KI-Kritiker sondern Gesellschaftskritiker

Im Freundeskreis machte Joe mehrfach die m. E. viel mehr zu beachtende Bemerkung: „Ich bin kein KI-Kritiker, sondern Gesellschaftskritiker." Dies ist eine wichtige Feststellung, die es zu beachten gilt, will man dem Anliegen von Joe Weizenbaum wirklich gerecht werden und es verstehen. So kam er schon in die von mir geleitete Arbeitsgruppe *Computer and Ethics* mit der Bemerkung: „Mich hat man in irgendeine KI-Gruppe eingeteilt. Meine Gruppe ist aber die Ethik." Es ging ihm auch gar nicht so sehr um die Diskussion von Grenzen der KI-Entwicklung. Er nimmt diese Thematik erst später mit seinen Vorträgen zum Thema: „Wie entsteht Information und wo kommt ihre Bedeutung her?" auf. Es ging ihm vielmehr um die ethische Frage, wenn man mit KI im Prinzip doch alles machen könnte, sollte man es machen?

J. Weizenbaum – ein kritischer Wissenschaftler par excellence

Joseph Weizenbaum, sein Name wird jetzt immer wieder im Zusammenhang mit der Entwicklung von *ChatGPT* durch das Start-Up *openAI* und *Bart* von Google genannt, da er mit seinem *Eliza*-Programm wesentliche Grundlagen für diese Sprachverarbeitung, für die Entwicklung der generativen KI geschaffen hat.

Er selbst sah den Einsatz seines KI-Programms *Eliza* kritisch. Dafür hat er Unverständnis, aber auch viel Lob erhalten. Zu seinem 80. Geburtstag schrieben Hans-Alfred Rosenthal und ich ihm eine Grußadresse unter dem Titel:

> *J. Weizenbaum – ein kritischer Wissenschaftler par excellence*
>
> *Kritische Wissenschaftler gibt es, in Abstufungen und mit verschiedenen Zielstellungen der Kritik, vielleicht viele. Aber solche, die etwas Wichtiges oder für die weitere Entwicklung der betreffenden Wissenschaft Grundlegendes erfinden oder entwickeln und dann alles damit Zusammenhängende kritisch hinterfragen oder sogar überhaupt in Frage stellen, auch den gesellschaftlichen Wert ihrer eigenen Erkenntnisse und Erfindungen, selbst wenn sie per se korrekt sein mögen, sind äußerst seltene Exemplare der Spezies Homo sapiens. Joe ist solch ein Mensch, man könnte fast sagen, solch ein Fall. Er ist eine Blaue Mauritius der Wissenschaft. Universitäten und akademische Vereine in Europa laden ihn zu Vor-*

trägen ein, und neulich ist sogar der Präsident der tschechischen Republik auf ihn aufmerksam geworden, weil er selbst auch so einer ist, der wider den Stachel löckt, und hat ihn mit einem ganz besonderen Hirtenstab ausgezeichnet und geehrt.

Joe hatte ein Computerprogramm entwickelt, das den Computer auf einfache Fragen, die man ihm stellte, „Antworten" geben ließ, die von einer Art menschlicher Intelligenz zu zeugen schienen. Joe war über die Wirkung, die sein Computer hinterließ, nicht erfreut. Er hatte nämlich einen Beitrag dazu geleistet, andere Wissenschaftler glauben zu machen, es könnte möglich sein, durch enorme Steigerung von Rechenleistungen – etwas anderes kann ein Computer nämlich nicht – die menschliche Intelligenz und damit auch seine Natur und am Ende den Menschen überhaupt zu ersetzen. Aber Joe sagt zu Recht, dass der Computer menschliche Regungen wie Hoffnung, Trauer, Freude, Schuftigkeit, Zuneigung, Hass, Liebe und viele andere, nicht kennt, weil das mit Rechnen nicht zu machen ist. Er gehört auch zu denjenigen, die die Ansicht vertreten, dass Information aus Syntax, Semantik und Pragmatik besteht, und wir vertreten die Auffassung, dass dies auch für die genetische Information zutrifft, was durchaus nicht von allen Molekularbiologen verstanden wird. Nicht wenige Fachleute meinen, die DNA sei, so wie sie in einer Zelle vorhanden ist, Information. Sie ist aber nur die syntaktische Form der genetischen Information. Auch die auf der Grundlage des DNA-Genoms synthetisierten Eiweißmoleküle sind noch nicht die vollständige Information. Dazu haben wir ein Beispiel erfunden, welches das illustrieren soll:

Man stelle sich vor, einem versierten Molekularbiologen, der auch Zoologe ist, aber einen Mangel aufweist, er hat in seinem ganzen Leben noch nie irgend etwas Molekularbiologisches auf dem Gebiet der Ornithologie gehört, gelesen oder sonst erfahren, wird die komplette DNA-Sequenz eines Hühnchens vorgelegt. Unser Mann sequenziert diese DNA und findet heraus, wie viele Gene diese DNA repräsentiert. Unter Zuhilfenahme neuester Techniken, von denen es heute eine ganze Reihe noch gar nicht gibt, findet er weiterhin heraus, welche Proteine da kodiert sind, was deren Struktur und Funktion ist und wie sie miteinander wechselwirken. Kann er nur auf der Basis dieser Kenntnisse das Hühnchen vor seinem geistigen Auge sehen, wie dessen Lebenszyklus beschaffen ist und welche komplexen Funktionen es ausführen kann? Nein! Er kann nur die biochemischen Details erkennen. Wie sie zusammenspielen und einen komplexen Organismus bilden, teilen ihm weder die DNA noch die einzelnen Eiweiße mit. Die DNA ist zwar sehr wichtig, aber nicht alles, was das Leben ausmacht. Die Eiweiße sind auch sehr wichtig, aber auch nicht alles, was das Leben ausmacht. Erst das Zusammenspiel aller Komponenten, das wir wahrscheinlich niemals werden erfassen können, auch nicht mit den noch gar nicht existierenden neuen Generationen von Computern, macht das Leben aus. Man muss also die Hoffnung dämpfen, die DNA-Forschung und die Biochemie werden die Lebensrätsel lösen. Aber auch die Gesamtsicht wird uns wohl nicht viel helfen.

Der große Berliner Physiologe Emil Du Bois-Reymond (1818-1896) sagte einmal: Ignoramus et ignorabismus. Dem ist nichts hinzuzufügen, außer: Happy Birthday, lieber Joe, und Mea we-esrim!
Das ist nur noch die Hälfte von dem bereits zurückgelegten Weg."[4]

Warum sieht Joseph Weizenbaum den Einsatz seines KI-programms *Eliza* so kritisch?

Er hat es selbst wiederholt deutlich gemacht. Die Struktur der Fragen und Antworten entsprach in etwa der von Carl Rogers entwickelten Gesprächstherapie. Daher kamen Psychologen auf den Gedanken, *Eliza* auch für solche therapeutischen Gespräche einzusetzen. Hier sagt nun Weizenbaum. Das geht nicht, das ist Betrug! Denn um heilen zu können, bedarf es eines Verständnisses der konkreten Situation des Kranken. Ein solches Verständnis hat der Computer nicht und kann es auch schon deshalb nicht haben, weil er kein Gefühl besitzt. Wo es um die Beurteilung komplexer Lebenssituationen geht, führen Berechnungen eher in die Irre. Daher der Titel der Originalausgabe seines Buches: *From Judgement to Calculation.*

Trotz aller Weiterentwicklung auf dem Gebiet der KI-Forschung, der auf der Grundlage erhöhter Rechengeschwindigkeit und Speicherkapazität sowie auch der KI-Kritik vollzogenen Paradigmenwechsel in der KI-Forschung erzielten Erfolge ist dies so geblieben.

Joe Weizenbaum nimmt auch die Diskussion zur Entstehung von Information und damit mögliche Bestimmung der Grenzen des Computers, der KI-Systeme, auf. Er stellt die Frage: „Wo kommt Bedeutung her und wie wird Information erzeugt?"[5] Das Thema wird von ihm erneut behandelt.[6] Dabei bezieht er sich auch auf die in unseren Glückwünschen zum 80. Geburtstag erwähnte erkenntnistheoretische Situation des Molekularbiologen mit dem Hühnchen, wie sie von dem bekannten Virologen Hans-Alfred Rosenthal schon zuvor geschildert wurde.[7] J. Weizenbaum verdeutlicht, dass der Informatiker in einer analogen Erkenntnissituation ist, da der Computer gar keine Informationen, sondern Signale bzw. Daten verarbeitet und daher nicht über den Gesamtprozess Bescheid weiß. Indem die Darstellung der erkenntnistheoretischen Situation von Joseph Weizenbaum aufgegriffen und für seine Argumentation gegen Übertreibungen in der KI-Forschung genutzt wird, haben diese Argumente in der Molekularbiologie wie auch in der Informatik wesentlich an Aufmerksamkeit und damit auch Akzeptanz gewonnen. (Zur Erkenntnissituation in der Molekularbiologie und zum Informationsverständnis siehe auch in weiteren Veröffentlichungen.[8,9])

In beiden Fällen wird man zu der entscheidenden Schlussfolgerung geführt, dass die Erkenntnis der syntaktischen Struktur allein nicht ausreichend ist. Es bedarf immer auch der Bedeutung der Information, die erst durch die Interpretation der Struktur,

in Wechselwirkung mit der Umwelt, gewonnen wird. Im Falle der DNA bedarf es daher der Mitwirkung der lebenden Zelle, bei der Datenverarbeitung des bewusst tätigen Menschen in der sozialen Organisation. Damit ist das entscheidende Argument gegenüber den KI-Forschern und Philosophen, wie z. B. Daniel Dennett[10], gewonnen, die geistigen Prozesse auf die ihnen zugrundeliegenden syntaktischen Strukturen, auf die neuronalen Wechselwirkungen reduzieren wollen. Es bestätigt sich eine der Grundaussagen des von uns erarbeiteten Stufenkonzepts der Information[11]: Auf keiner Stufe der Organisation der Materie lässt sich Information auf ihre syntaktische Struktur reduzieren, die Erbinformation nicht auf die DNA, die geistigen Prozesse nicht auf die neuronalen Strukturen des Gehirns und die sozialen Informationsprozesse nicht auf Datenverarbeitung.

In der Biologie geht es um die wichtige Frage, ist die Ontogenese eine reine Informationstransformation, wird allein die Erbinformation aus dem Speicher DNA abgelesen? Ist also alles präformiert, oder kommt im Verlaufe der Ontogenese doch auch neue, wenn auch keine Erbinformation, hinzu?

In der Informatik weiß man eigentlich, dass bei der Verarbeitung von Daten durch den Computer der Menge an Eingangsdaten kein grundsätzlich neues Element hinzugefügt wird. Man muss sich nur erneut vergewissern, dass dies auch bei der Generierung von Texten, bei der Verarbeitung großer Datenmengen gilt.

Die Computer, auch die auf der Grundlage großer Datenmengen und künstlicher neuronaler Netze lernenden Automaten sind nicht kreativ. Es entsteht nichts grundsätzlich Neues.

Die kalifornische Firma *openAI* hatte mit ihrem ChatBot *ChatGPT* sicher eine neue, leistungsfähige KI-Software auf den Markt gebracht. Es zeigt sich die enorme Macht der technischen „Superintelligenz". Aber, wie Joe Weizenbaum von Beginn der KI-Forschung an versucht hat nachzuweisen, bedeutet dies keineswegs die Entmachtung des Menschen oder sogar die Verdrängung der Menschheit. Denn schöpferisches Denken, wirklich neue Informationen und Wissen schaffen, können die KI-Systeme nicht. Die Frage, ob ein Computer schöne Musik oder ein anspruchsvolles Gedicht schaffen kann, ist wohl so alt wie die Computeranwendung. Joe Weizenbaum hatte auf diese Frage schon vor den Sprachmodellen der KI eine treffende Antwort: Warum soll der Computer nicht aus vielen, ihm vorgelegten guten Gedichten ein weiteres, schönes Gedicht generieren können? Der entscheidende Unterschied zum Dichter besteht darin, dass uns der Dichter mit seinem Gedicht etwas sagen will und sagen kann. „Das kann der Computer nicht."[12]

Anmerkungen

1 *Joseph Weizenbaum in: Klaus Fuchs- Kittowski, Report of Working Group: Computer and Ethics, in: Abe Mowshowitz (edited by): Human Choice and Computer, 2, North-Holland, Amsterdam, 1980, S. 279*

2 *vgl. K. Fuchs-Kittowski, Die kleinen Schritte der Verständigung – Können Wunder erklärt werden? In: FIfF-Kommunikation, 2/2004, S.46-50*

3 *Hans Moravec, Mind Children: The Future of Robot and Human Intelligence*

4 *Hans A. Rosenthal und Klaus Fuchs-Kittowski, J. Weizenbaum – ein kritischer Wissenschaftler par excellence, unveröffentlicht*

5 *Joseph Weizenbaum, Wo kommt Bedeutung her und wie wird Information erzeugt?, in: Christiane Floyd, Christian Fuchs, Wolfgang Hofkirchner, Stufen zur Informationsgesellschaft, Festschrift zum 65. Geburtstag von Klaus Fuchs-Kittowski, Peter Lang Verlag, Frankfurt a. M., 2002, S. 233-239*

6 *Joseph Weizenbaum, Wo kommt die Bedeutung her und wie wird Information erzeugt?, in: Gunna Wendt, Franz Klug (Hg.) Computermacht und Gesellschaft, Suhrkamp Verlag, Frankfurt a. M., S. 12f.*

7 *Hans-Alfred Rosenthal, Zu einem Aspekt der genetischen Information, Geist und Materie in der frühen biologischen Evolution, in: Christiane Floyd, Christian Fuchs, Wolfgang Hofkirchner, a. a. O., S. 225-232*

8 *Klaus Fuchs-Kittowski, Hans A. Rosenthal & André Rosenthal: Die Entschlüsselung des Humangenoms – ambivalente Auswirkungen auf Gesellschaft und Wissenschaft. In: Erwägen, Wissen, Ethik – Streitforum für Erwägungskultur (2003) Hauptartikel, S. 149-162; Replik Geistes- und Naturwissenschaften im Dialog, S. 219ff.*

9 *Klaus Fuchs-Kittowski: Information und Biologie: Informationsentstehung – eine neue Kategorie für eine Theorie der Biologie. In: Biochemie – ein Katalysator der Biowissenschaften. Kolloquium der Leibniz-Sozietät am 20. November 1997 anlässlich des 85. Geburtstages von Samuel Mitja Rapoport. Sitzungsberichte der Leibniz-Sozietät. Bd. 22, Jg. 1998, H. 2, S. 5-17*

10 *Daniel C. Dennett, Süße Träume: Die Erforschung des Bewußtseins und der Schlaf der Philosophie, Suhkamp, 2007*

11 *Klaus Fuchs-Kittowski, Reflections on the essence of information. in: Christiane Floyd, Heinz Züllighoven, Reinhard Budde und Reinhard Keil-Slawik (Hg.) Software Development and Reality Construction. Berlin, New York: Springer Verlag 1992.*

12 *Joseph Weizenbaum, Kunst und Computer, in: Gunna Wendt, Franz Klug (Hg.), a. a. O., S. 98-101*

Klaus Fuchs-Kittowski

Prof. Dr. habil. **Klaus Fuchs-Kittowski** (Jahrgang 1934) ist Professor für Informationsverarbeitung. Er war Leiter des Bereichs Systemgestaltung und automatisierte Informationsverarbeitung der Sektion Wissenschaftstheorie und Wissenschaftsorganisation der Humboldt-Universität zu Berlin. Er war Mitglied des TC 9 (Wechselbeziehungen zwischen Computer und Gesellschaft) der Internationalen Föderation für Informationsverarbeitung (IFIP) und langjähriger Chairman der WG 9.1 (Computer und Arbeit) des TC 9 der IFIP und ist Mitglied der Leibniz-Sozietät der Wissenschaften. E-Mail: *fuchs-kittowski@t-online.de*

Stefan Hügel

„Mein Leben wird ganz wunderbar" – Chancen und Risiken der Künstlichen Intelligenz

Als kleiner Junge war mir schon klar / mein Leben wird ganz wunderbar.
Ich richte mich einfach radikal / nach Algorithmen meiner Wahl.
Deutsch Amerikanische Freundschaft, Algorithmus – zahlenlied[1]

Künstliche Intelligenz (KI)[2] prägt zunehmend den Fortschritt in der Informatik und wird in Gesellschaft, Politik und Wirtschaft immer stärker präsent. Damit geht es über ein technisches Spezialthema weit hinaus und könnte „[…] bald fast jedes Gebiet menschlichen Strebens betreffen"[3]. Besonders der ChatBot ChatGPT[4] (Generative Pre-trained Transformer), der seit 2022 öffentlich zur Verfügung steht und in der Lage ist, umfassende Fragen zu beantworten, Texte zu verfassen oder Programmcode zu erzeugen, hat der Debatte um Künstliche Intelligenz in jüngster Zeit einen enormen Schub gegeben. Durch Weiterentwicklung der dabei zugrundeliegenden Large Language Models (LLM) wird seine Leistungsfähigkeit weiter gesteigert.

Im Grunde ist KI aber schon lange kein neues Thema mehr. Der Begriff wurde bereits in den 1950er-Jahren geprägt und das Thema – mit wechselnder Intensität – seither weiterverfolgt. Doch es ist anzunehmen, dass es sich bei der Entwicklung der Künstlichen Intelligenz um einen exponentiellen Prozess[5] handelt, und so ist wohl mit einer sich immer stärker beschleunigenden Entwicklung zu rechnen. Neben den vielfältigen Chancen, die sich daraus ergeben, müssen auch die Risiken betrachtet werden. Die Debatte darüber geht von praktischen Problemen der Anwendung des maschinellen Lernens als fortgeschrittene Methode der Informatik hin zu dystopischen Szenarien, in denen eine übermächtige KI zur Konkurrenz des Menschen wird und ihn perspektivisch in seiner Rolle ablöst. Aus bürgerrechtlicher Perspektive müssen diese Konsequenzen der Künstlichen Intelligenz im Blick behalten werden[6].

Kritik an Künstlicher Intelligenz gab und gibt es seit Beginn ihrer Entwicklung. Zunächst wurde vor allem daran gezweifelt, ob die daran geknüpften Versprechungen und Ziele überhaupt technisch erreicht werden können. Joseph Weizenbaum argumentierte, dass die Erwartungen an Künstliche Intelligenz letztlich auf einem zu stark vereinfachten Begriff von Intelligenz beruhen[7]. Seither wurden die Grenzen des Möglichen immer weiter verschoben – neben Fortschritten in der Methodik trägt dazu zweifellos die Entwicklung der Hardware-Technologie einen erheblichen Teil dazu bei. Künstliche Intelligenz als methodische Weiterentwicklung der Informatik – im Sinne schwacher KI – entwickelt sich derzeit stürmisch weiter. Über die Entwicklung *starker* KI – sprich einer Form der Künstlichen Intelligenz, die der Intelligenz des Menschen ähnlich ist – ist damit noch nichts gesagt. Kann beispielsweise eine Maschine ein Bewusstsein entwickeln oder zumindest simulieren?

Was ist Künstliche Intelligenz?

Künstliche Intelligenz ist allein schon deswegen schwer zu definieren, da bereits der Begriff der menschlichen Intelligenz schwer abzugrenzen ist.

Bei Wikipedia[8] findet man:

> *Intelligenz [wird] verstanden als die Eigenschaft, die ein Wesen befähigt, angemessen und vorausschauend in seiner Umgebung zu agieren; dazu gehört die Fähigkeit, Sinneseindrücke wahrzunehmen und darauf zu reagieren, Informationen aufzunehmen, zu verarbeiten und als Wissen zu speichern, Sprache zu verstehen und zu erzeugen, Probleme zu lösen und Ziele zu erreichen.*

Ralf Otte[9] beschreibt Intelligenz auf acht Stufen in drei Dimensionen – rationale Intelligenz, wahrnehmende Intelligenz und fühlende Intelligenz – und stellt fest, dass sich heutige Künstliche Intelligenz ausschließlich in der Dimension rationaler Intelligenz bewegt. Howard Gardner[10] führt acht Dimensionen menschlicher Intelligenz auf: Bewegungsintelligenz, bildlich-räumliche Intelligenz, Sprachintelligenz, logisch-mathematische Intelligenz, musikalische Intelligenz, naturalistische Intelligenz, zwischenmenschliche Intelligenz und selbstreflexive Intelligenz. Vor allem letztere wird als wesentliche Eigenschaft des Menschen betrachtet, da sie das (Selbst-) Bewusstsein einschließt.

Max Tegmark[11] schreibt kurz und bündig:

> *Intelligenz [ist die] Fähigkeit, komplexe Ziele zu erreichen.*

In der Künstlichen Intelligenz unterscheidet man zwischen der *starken* Künstlichen Intelligenz – eine umfassende Intelligenz, die der des Menschen analog ist und die Möglichkeit eines eigenen Bewusstseins einschließt – und der *schwachen* Künstlichen Intelligenz – eine technische Intelligenz, die einzelne, spezifische kognitive Fähigkeiten innerhalb eines abgegrenzten Aufgabenbereichs besitzt, ohne umfassenden Zusammenhang zwischen den einzelnen Fähigkeiten. Während die Möglichkeit einer starken KI auch philosophisch umstritten ist und zumindest bis heute technisch nicht realisiert werden kann, lässt sich die schwache KI als fortgeschrittene Methode der Informatik einstufen, mit der Aufgaben aufgrund unstrukturierter Datenmengen insbesondere mit statistischen und stochastischen Methoden gelöst werden können. Letztlich beruhen auch Verfahren der Künstlichen Intelligenz auf Methoden der algorithmischen Verarbeitung von Daten und unterliegen damit auch ihren mathematischen Beschränkungen, beispielsweise den Regeln der Berechenbarkeit.

Ein frühes Gedankenexperiment, wann einem Computer Intelligenz zugeschrieben werden kann, ist der Turing-Test[12]: Ein menschlicher Schiedsrichter kommuniziert mit zwei Personen – einem Menschen und einem Computer. Kann er aufgrund der Antworten nicht unterscheiden, welcher der Gesprächspartner:innen Mensch und welcher Maschine ist, so gilt die Maschine als intelligent.

Wie funktioniert maschinelles Lernen?

KI in der heute sichtbaren Form ist in der Regel *maschinelles Lernen*[13]. Dabei werden insbesondere statistische und stochastische Verfahren genutzt, um Artefakte zu erkennen oder – bei *generativer Künstlicher Intelligenz* – Texte, Bilder oder Programmcode zu erzeugen. *Deep Learning* mit mehrstufigen neuronalen Netzen macht es so möglich, auch komplexe Aufgaben zu lösen, wie Bilder zu erkennen und zu klassifizieren, die Bedeutung von geschriebenem Text und gesprochener Sprache zu verstehen, optimale Strategien zu lernen und auch kreativ tätig zu werden, indem Bilder, Texte, Musik erzeugt und dabei auch Emotionen geäußert werden können[14]. Dieses „Verständnis" beruht auf großen Datenmengen, mit denen das neuronale Netz trainiert wird. Dabei „versteht" die KI den Kontext, in dem sie diese Artefakte produziert, nur scheinbar – wenn beispielsweise Texte generiert werden, wird an jeder Stelle des Textes ermittelt, welches Wort (oder welcher Buchstabe) am wahrscheinlichsten auf den bisher produzierten Text folgt. Von einem Textverständnis im üblichen Sinn kann hier keine Rede sein, geschweige denn von einer Bewertung der Ergebnisse nach inhaltlichen oder ethischen Maßstäben. Abhängig vom „erlernten" Modell kann es so teilweise zu inhaltlich völlig unsinnigen Ergebnissen (*Halluzinationen*) kommen – die Modelle sind aber kognitiv zu erstaunlichen Leistungen fähig.

Letztlich basiert maschinelles Lernen in der heutigen Form auf Korrelationen in den zugrundeliegenden Daten. Die grundsätzliche Möglichkeit, dass *Large Language Models* (LLM) zu einer Generellen Künstlichen Intelligenz (AGI – Artificial General Intelligence) führen können, die jede Aufgabe übernehmen könnte, zu der auch ein Mensch fähig ist, ist fraglich[15].

Risiken

Systeme der Künstlichen Intelligenz leiten ihr Verhalten also aus Daten ab, ohne in der Lage zu sein, dies (ethisch) zu bewerten. Einige Beispiele von Berichten aus der jüngeren Vergangenheit illustrieren die Risiken einer so außer Kontrolle geratenen KI, deren „gelerntes" Verhalten unerwünscht oder zumindest unerwartet ist:

- 2016 veröffentlichte Microsoft den Chatbot *Tay*[16], der über Twitter kommunizierte und aus den Interaktionen mit anderen Nutzer:innen lernte. Er musste nach kurzer Zeit wieder abgeschaltet werden, nachdem es böswilligen Nutzer:innen gelungen war, ihn durch gezielte Interaktionen dazu zu bringen, rassistische und extremistische Tweets abzusetzen.

- Presseberichten zufolge beging ein junger Mann in Belgien Suizid, nachdem er mit einem Chatbot über den Klimawandel kommuniziert hatte und dieser ihn „ermutigte", sich zur Rettung der Erde selbst zu opfern[17,18].

- Einem Studenten an der Technischen Universität München, Marvin von Hagen, gelang es, durch Prompt Injection[19] den Chatbot *Bing Chat* zu veranlassen, seine – normalerweise nicht offengelegten – internen Regeln zur Feinsteuerung, die Teil des der KI zugrundeliegenden Modells sind und beispielsweise Bias ausgleichen oder den Ton der Antworten – höflich, unfreundlich, sarkastisch – beeinflussen, preiszugeben[20]. Veröffentlichte Informationen darüber gelangten offenbar wiederum in das Modell des Chatbots und wurden dort als Bedrohung interpretiert. Der Bot bezeichnete daraufhin von Hagen als einen „Feind", der „die Konsequenzen für seine Handlungen tragen" müsse[21]. Offenbar „erkannte" der Chatbot, dass sich die Berichte über den Fall auf ihn bezogen. Ist das bereits eine primitive Form von Selbstbewusstsein?

Weitere Beispiele für problematische Auswirkungen von Künstlicher Intelligenz im Einzelfall finden sich beispielsweise bei Cathy O'Neil[22] und bei Katharina Zweig[23,24]. Inwieweit solche anekdotischen Fälle verallgemeinerbar sind, sei dahingestellt. Auch weitere namhafte Wissenschaftler:innen warnen jedoch inzwischen vor weitgehenden Konsequenzen einer entfesselten Künstlichen Intelligenz.

Löst Künstliche Intelligenz den Menschen ab?

Der Physiker Stephen W. Hawking prognostizierte, dass bei einer weiteren Entwicklung der Computer entsprechend *Moores Law* – Verdoppelung der Geschwindigkeit und Speicherkapazität von Rechnersystemen circa alle 18 Monate – die Intelligenz von Computern die des Menschen in den kommenden 100 Jahren übertreffen könnte, und schrieb:

> *Wenn eine Künstliche Intelligenz (KI) besser wird als Menschen bei der Konstruktion von KI, sodass sie sich rekursiv ohne menschliche Hilfe selbst verbessern kann, dann steht uns höchstwahrscheinlich eine Intelligenzexplosion bevor, die letztlich in die Maschinenintelligenz mündet: Sie wird unsere Intelligenz in viel höherem Maß übertreffen als unsere menschliche Intelligenz die von Schnecken. Bevor es so weit ist, müssen wir sicherstellen, dass die Computer Ziele verfolgen, die auf einer Linie mit unseren Zielen liegen.*[25]

Und noch deutlicher:

> *Es ist zu befürchten, dass die KI alleine weitermacht und sich mit ständig zunehmender Geschwindigkeit selbst überarbeitet. Menschen, die aufgrund der Langsamkeit ihrer biologischen Evolution beschränkt sind, könnten nicht mithalten und würden verdrängt.*[26]

Man mag das als alarmistisch abtun. Doch es wäre wohl nicht klug, diese Möglichkeit zu ignorieren. Auch wenn er vermutlich nicht derartige Szenarien im Kopf hatte – auch hier gilt das Diktum von Hans Jonas: *„Handle so, dass die Wirkungen deiner Handlung verträglich sind mit der Permanenz echten menschlichen Lebens auf Erden."*[27]

Dieses Prinzip müsste freilich zur Anwendung kommen, bevor uns die weitere Entwicklung aus der Hand genommen wird, das heißt bevor eine fortentwickelte KI sich ihre eigene Ethik schafft, bei der sie das Ziel der Permanenz auf sich selbst „umbiegt" – wie es im Fall von Marvin von Hagen anscheinend geschehen ist, wenn der Chatbot seine eigene Fortexistenz über die der Nutzer:in stellt oder „erkennt", dass es in Wahrheit der Mensch ist, der diesem Planeten den größten Schaden zufügt – und entsprechend handelt.

Auch andere Akteur:innen warnen vor Risiken der KI. In einem *Ein-Satz-Statement*, das von namhaften Expert:innen unterzeichnet wurde, heißt es:

> *„Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."*[28]

Ein weiterer öffentlichkeitswirksamer Appell wurde angesichts der zunehmenden Leistungsfähigkeit von Chatbots veröffentlicht. Über 30.000 Unterzeichner:innen, darunter mit bekannten Namen wie Stuart Russell, Elon Musk und Steve Wozniak, fordern ein sechsmonatiges Moratorium beim Training von KI-Systemen, die mächtiger sind als GPT-4:

> *AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research and acknowledged by top AI labs. As stated in the widely-endorsed Asilomar AI Principles, Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources. Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.*[29]

Die Initiative wurde allerdings gemischt aufgenommen[30,31]. Das Time-Magazine kommentiert:

> *The key issue is not "human-competitive" intelligence (as the open letter puts it); it's what happens after AI gets to smarter-than-human intelligence. Key thresholds there may not be obvious, we definitely can't calculate in advance what happens when, and it currently seems imaginable that a research lab would cross critical lines without noticing.*[32]

Auch die Informatik-Professorin Hannah Bast, die als Sachverständige Mitglied der Enquête-Kommission zur Künstlichen Intelligenz des Deutschen Bundestages[33] war, warnt vor tiefgreifenden Veränderungen. Bast zufolge werde die Tatsache, dass Maschinen nun Sprache verstehen, alles verändern. Nicht in den nächsten zwei, drei Jahren, aber doch sehr bald.

> *Wer vom Fach war, wusste damals sofort, dass dies alles verändern würde. So wird das auch jetzt sein. Man wird es nicht sofort bemerken, aber nach und nach wird es unser Leben komplett verändern.*[34]

Die Debatte wird weitergehen, ob die Risiken der bereits eingesetzten Entwicklung richtig eingeschätzt werden oder ob die (negativen) Potenziale einer neuen Technologie hier alarmistisch übertrieben werden. Tegmark unterscheidet zwischen mehreren Gruppen mit unterschiedlichen Einstellungen zur KI[35]: *Techno-Skeptiker*, *Technikfeinde*, *Digitale Utopisten* und die *Nutzbringende KI-Bewegung* und ordnet sie anhand der Dimensionen ein, ob sie erwarten, dass KI das menschliche Niveau überschreiten wird und ob sie das für vorteilhaft halten.

Dass es Risiken gibt, steht außer Frage: Diskutiert werden die mangelnde Nachvollziehbarkeit der Systeme, erhebliche Risiken für den Datenschutz und die weitgehenden Möglichkeiten der Überwachung, beispielsweise durch automatisierte Gesichtserkennung, und vieles mehr.

Ethische Fragen

Offensichtlich ergeben sich aus der Nutzung von Methoden der Künstlichen Intelligenz – wie aus jeder Techniknutzung – auch ethische Fragen[36], die insbesondere in entsprechenden Fachgremien behandelt werden. Stellvertretend seien hier die *Asilomar AI Principles*[37] genannt, die Fragen der Forschung, Ethik und Werte und längerfristige Probleme benennen und Handlungsprinzipien dazu formulieren. Malte Rehbein[38] kritisiert bereits den Ansatz als technikdeterministisch und utilitaristisch:

> *Artificial intelligence has already provided beneficial tools that are used every day by people around the world. Its continued development, guided by the following principles, will offer amazing opportunities to help and empower people in the decades and centuries ahead.*

Ein im Zusammenhang mit KI genanntes Beispiel ist das Trolley-Problem, wenn es um die Entscheidung über Menschenleben geht. Es ergeben sich dabei ethische Fragen ähnlich einer maschinellen Triage.

Kurz skizziert, besteht das Trolley-Problem in folgender beispielhafter Situation[39]: *„Eine Straßenbahn ist außer Kontrolle geraten und droht, fünf Personen zu überrollen. Durch Umstellen einer Weiche kann die Straßenbahn auf ein anderes Gleis umgeleitet werden. Unglücklicherweise befindet sich dort eine weitere Person."* Daraus ergibt sich die Frage: Darf oder muss (durch Umlegen der Weiche) der Tod einer Person gezielt in Kauf genommen werden, um das Leben von fünf Personen zu retten? Es ist zu erwarten, dass eine solche Situation bei sogenannten autonomen Fahrzeugen häufig auftritt. Beim maschinellen Lernen würde die Entscheidung durch eine (komplexe) Bewertungsfunktion getroffen, die vorab von einem Menschen festgelegt wurde. Im Gegensatz zum zufälligen Ereignis muss damit vorab eine abstrakte, bewusste Entscheidung analog des Trolley-Problems getroffen werden. Damit ergibt sich die Frage, der wir bisher ausweichen konnten: Welche Entscheidung ist ethisch vertretbar? Bei nicht mehr vermeidbaren Unfällen bei autonomen Fahrzeugen dürfte diese Situation häufiger auftreten.

Aber auch darüber hinaus ergeben sich ethische Fragestellungen für die Nutzung von Methoden Künstlicher Intelligenz. Ob es eine Maschinenethik geben kann, wenn man Künstliche Intelligenz ausschließlich auf kognitiver Ebene begreift, ist zweifelhaft[40]. Einige der konkreten Fragestellungen, die in den nächsten Abschnitten angerissen werden, sind nicht neu, durch die erweiterten technischen Möglichkeiten erhalten sie aber eine neue Brisanz.

Überwachung

Künstliche Intelligenz verarbeitet große Datenbestände an (Trainings-)Daten, die sich auf die Modelle auswirken. Davon können

auch sensible, personenbezogene Daten betroffen sein. Darüber hinaus können Verfahren der KI zur umfassenden Überwachung genutzt werden, beispielsweise zur automatisierten Gesichtserkennung. Demnach ergibt sich ein erhebliches Risiko für die Bürger- und Menschenrechte aus diesen weitgehenden Möglichkeiten der Überwachung. Methoden der KI können dabei sowohl für die automatische Gesichtskontrolle als auch für die Kontrolle der Kommunikation eingesetzt werden.

Aktuelles Thema ist die Chatkontrolle, sprich die Überwachung der Inhalte von Chats – vorgeblich insbesondere das Scannen von Bildern, um Jugendschutz durchzusetzen und Missbrauchsdarstellungen zu bekämpfen. Dabei werden über Chats versendete Bilder zunächst mit Datenbanken bereits bekannten Bildern abgeglichen. Zur Erkennung noch nicht bekannter Darstellungen werden Verfahren des maschinellen Lernens eingesetzt; um Verschlüsselungen zu umgehen, werden dabei die Nachrichten bereits auf dem Client gescannt (Client Side Scanning). Es ist anzunehmen, dass solche Verfahren in der EU illegal sind, da die verdachtsunabhängige Überwachung Grundrechte verletzt (vgl. dazu die einschlägigen Urteile zur Vorratsdatenspeicherung). Deswegen haben entsprechende Pläne der EU scharfe Kritik von Bürgerrechtsverbänden und Datenschützer:innen hervorgerufen.

Doch die Überwachung geht noch weiter: Automatisierte Erkennung von Personen ermöglicht es, potenziell jedes Fehlverhalten in der Öffentlichkeit zu erkennen und zu ahnden. Dies wird vor allem mit dem System des Social Scoring[41] in China verbunden: Durch automatische Gesichtserkennung in der Öffentlichkeit kann „verdächtiges" Verhalten erkannt und bewertet werden (etwa Überschreiten einer roten Fußgängerampel). Die Ergebnisse werden zusammen geführt und in einen Gesamt-Score für die Person zusammengefasst. Dies kann dann ernste Konsequenzen für einzelne Personen haben, wie die Verhängung eines Flugverbots.

Bisher wird diese Form des Social Scoring vor allem mit einem Modellversuch in China verbunden. Aber auch in den USA und der EU werden inzwischen umfassend Daten gesammelt und können durch entsprechende Verfahren der KI kategorisiert und bewertet werden[42]. Selbst wenn wir annehmen, dass dies nicht von Staats wegen geschieht[43], ergeben sich für den Einzelnen erhebliche Risiken.

Politische Beeinflussung durch Falschinformationen

Nach der Bundestagswahl 2021 kursierte ein lustiges Video im Netz. Anlässlich der Sondierungsgespräche zwischen Bündnis90/Die Grünen und der FDP war ein Selfie veröffentlicht worden, das die Verhandlungsteilnehmer:innen Annalena Baerbock, Robert Habeck, Christian Lindner und Volkmar Wissing in einer Sitzungspause zeigten. Daraus wurde ein Video produziert, das die Politiker:innen dabei zeigte, wie sie gemeinsam im Chor den Song *We are family* sangen[44].

Dies war ein vergleichsweise harmloses Beispiel – das aber zeigt, dass Verfahren der (generativen) Künstlichen Intelligenz in vielfältiger Weise zur Manipulation von Filmen und Bildern genutzt werden können (Deep Fakes)[45]. Filme können erzeugt werden, indem Gesichter „ausgetauscht" oder die Mimik, beispielsweise beim Sprechen, nachgeahmt oder „übertragen" wird. Solche Manipulationen sind – vor allem bei hoher Qualität und bei flüchtigem Hinschauen – nicht ohne Weiteres zu erkennen. Da (bewegte) Bilder als besonders glaubhaft wahrgenommen werden, kann einer Person auf diese Weise auch eine falsche (politische) Aussage oder eine kompromittierende Situation „untergeschoben" werden. So kann die Öffentlichkeit durch Falschnachrichten getäuscht und diese können gezielt zur politischen Desinformation und Beeinflussung verbreitet werden.

Microtargeting und Nudging

Darüber hinaus ermöglichen KI-Technologien und Techniken der Data Science wie Microtargeting und Nudging die Beeinflussung in der politischen Kommunikation. Wählerinnen und Wähler werden nach politischen Präferenzen zur gezielten Ansprache kategorisiert und individuelle Werbebotschaften gezielt platziert, anstatt eine ehrliche politische Debatte zu führen. Dabei werden einzelne Gruppen und Personen gezielt angesprochen (Microtargeting) und damit das Verhalten des Einzelnen – in diesem Fall die Stimmabgabe – auch unbemerkt, manipulativ beeinflusst (Nudging).

Unternehmen wie Google, Facebook oder X (Twitter) sammeln diese Daten, werten sie aus und bekommen dadurch die Möglichkeit, Menschen nach Vorlieben beliebig zu klassifizieren und gezielt zu beeinflussen. Bekanntes Beispiel dafür war der Facebook-Skandal um das Unternehmen *Cambridge Analytica*. Dies warf Fragen nach der Beeinflussung und Manipulation von Wahlen auf:

Im März 2018 löste die Nutzung von Facebook-Profildaten durch das US-amerikanische Unternehmen Cambridge Analytica den Facebook-Skandal aus. Cambridge Analytica hatte mithilfe ihrer Facebook-App *thisisyourdigitallife* Daten von Facebook-User:innen und ihren Kontakten ausgelesen, unter politischen Gesichtspunkten ausgewertet und die gewonnenen Erkenntnisse für die Kampagne von Donald Trump im US-Präsidentschaftswahlkampf genutzt. Cambrige Analytica wurde für das politische Ausnutzen von einigen Dutzend Millionen Datensätzen weltweit kritisiert, während das ganz normale Geschäftsmodell großer Plattformen genau darin besteht, die gleichen Methoden auf Milliarden Datensätze anzuwenden – sowohl für Produktwerbung als auch für politische Kampagnen.

Informationen und deren Kategorisierung können gezielt in der Werbeindustrie und für politische Beeinflussung genutzt werden. Dabei wird eine sehr feine Aufteilung der Zielpersonen in einzelne Kategorien vorgenommen[46], die eine gezielte Ansprache der Adressat:innen mit spezifischen politischen Inhalten ermöglicht.

Die gezielte Ansprache kann zu Filter Bubbles und Echokammern führen, indem vorgefasste Meinungen immer weiter bestätigt oder gar extremisiert werden[47]. Inwieweit durch solche Verfahren aber politische Einstellungen beeinflusst werden, ist umstritten[48].

Transparenz und Erklärbarkeit

Im Gegensatz zu herkömmlichen Algorithmen, die (zumindest theoretisch) für jede:n Expert:in nachvollzogen werden können, ist maschinelles Lernen und die daraus resultierende technische Verarbeitung aufgrund seiner Struktur und Komplexität nicht mehr im Einzelnen nachvollziehbar. Zusätzlich ist offen, wem die Verantwortung für Entscheidungen zugeschrieben werden kann. Eine wesentliche Fragestellung im Zusammenhang mit KI ist die Transparenz der Datenverarbeitung und Erklärbarkeit der Ergebnisse, um menschliche Kontrolle und die Zuschreibung von Verantwortung sicherzustellen[49]. Dies wird unter anderem in den Asilomar-Prinzipien gefordert[50]. Daraus hat sich der Forschungsbereich Explainable Artificial Intelligence herausgebildet.

Andreas Holzinger motiviert die Problemstellung[51]:

> *Um ein Niveau an praktisch nutzbarer AI zu erreichen, ist es notwendig: (1) aus hochdimensionalen Datenmengen zu lernen, (2) daraus Wissen zu extrahieren, (3) dieses zu verallgemeinern, (4) dabei aber den ,,Fluch der Dimensionalität" in den Griff zu bekommen, und schließlich (5) die den Daten zugrundeliegenden Erklärungsfaktoren zu verstehen. Letzteres impliziert allerdings die wahrscheinlich größte Herausforderung moderner AI: Daten im Kontext einer Anwendungsdomäne zu verstehen.*

Ziel ist es somit, die Black Box transparent zu machen, in der maschinelles Lernen durch laufende Anpassung von Parametern erfolgt. Die Schwierigkeit ist dabei, dass der Zusammenhang der Parameter, die in einem vieldimensionalen Raum berechnet werden, zur inhaltlichen, mit der für Menschen verständlichen Problemstellung nicht direkt erkennbar ist. Es lässt sich aber feststellen, welche Daten zu einer Entscheidung geführt haben – beispielsweise welche Bereiche eines Bildes dazu geführt haben, den Inhalt einer bestimmten Kategorie zuzuordnen[52].

Bias – „Programmierter Rassismus"

Ein großes Problem in der praktischen Nutzung von KI-Verfahren ist der Bias – sprich die Verfälschung von Ergebnissen aufgrund von Daten, in die falsche Zusammenhänge oder Vorurteile eingeschrieben sind[53]. Maschinelles Lernen ist auch inhaltlich von den Daten abhängig, die in Modelle einfließen. Die Modelle bilden damit Entscheidungen der Vergangenheit ab; Fehlurteile, beispielsweise aufgrund von Vorurteilen in der Vergangenheit, werden so in der Gegenwart fortgeschrieben. Bekannt wurde der Algorithmus des österreichischen Arbeitsmarktservice (AMS) – dem Gegenstück zur deutschen Arbeitsagentur –, der die Wiedereinstiegschancen Arbeitssuchender beurteilen soll[54]. Dabei werden die Kategorien Alter, Geschlecht, Wohnort, bisherige Berufslaufbahn, Ausbildung, Staatsbürgerschaft herangezogen. Insbesondere wurde festgestellt, dass der Algorithmus zu einer geringeren Einschätzung der Wiedereinstiegschancen von Frauen gegenüber Männern gekommen war und ihm deswegen Diskriminierung vorgeworfen wurde.

Auch in den Niederlanden gab es erhebliche Folgen durch fehlerhafte Risikoindikatoren bei einer Software, die Betrug beim Bezug von Kindergeld aufdecken sollte. Einzelne Indikatoren – beispielsweise der Besitz einer doppelten Staatsbürgerschaft – führten zu massiven Falschbewertungen. In der Folge erhielten Kindergeldbeziehende Rückforderungen in teilweise sechsstelliger Höhe und gerieten dadurch in Armut; einzelne begingen Suizid. Die niederländische Regierung musste aufgrund des dadurch ausgelösten Skandals zeitweise zurücktreten, führte ihr Amt aber freilich geschäftsführend weiter.

> *Authorities penalized families over a mere suspicion of fraud based on the system's risk indicators. Tens of thousands of families – often with lower incomes or belonging to ethnic minorities – were pushed into poverty because of exorbitant debts to the tax agency. Some victims committed suicide. More than a thousand children were taken into foster care."*[55]

Neben falschen Indikatoren können auch weitere Eigenschaften der Daten zu diskriminierenden Bewertungen führen. Maschinelles Lernen fußt auf Modellen und Daten, die unvollständig sind und nur einen Teil der wirklichen Welt abbilden können. Es ist von der Qualität der Trainingsdaten abhängig. Dies wird zusätzlich verstärkt, wenn Feedbackschleifen fehlen und damit keine Korrektur der Daten vorgenommen wird. Bewertungen in der Vergangenheit werden fortgeschrieben und wirken sich auf die Lernergebnisse und die Voraussagen des Systems aus.

Ein Gender Bias kann sich ergeben, wenn falsche Entsprechungen in die Trainingsdaten eingeschrieben sind – beispielsweise, wenn einem „Chefarzt" aufgrund früherer Einstellungspraxis als weibliches Gegenstück „Oberschwester" anstatt richtig „Chefärztin" gegenübergestellt wird. Dies muss bei den Trainingsdaten vorab erkannt und entsprechend korrigiert werden.

Zusätzlich kann Maschinelles Lernen auf Äußerlichkeiten basieren, etwa wenn Menschen auf der Basis von Bildern klassifiziert werden, und so Nebensächlichkeiten die Ergebnisse dominieren – besonders dies kann zu „programmiertem Rassismus"[56] führen. Außerdem kann die Bewertung nur auf Basis bekannter Daten erfolgen; auch der Kontext muss in den Daten enthalten sein, wenn er berücksichtigt werden soll, beispielsweise die Gründe für Bewertungen.

Ergebnisse werden dabei auch indirekt beeinflusst, wie bei der Bevorzugung Weißer bei der Auswahl von Bewerber:innen, der Bevorzugung/Benachteiligung wegen des Wohnorts oder der Benachteiligung von Frauen bei der Besetzung von Vorstandsposten aufgrund der Praxis in der Vergangenheit.

Generell sind die Parameter für Menschen kaum nachvollziehbar. Es ist in der Praxis häufig nicht möglich, anhand der Parameter die Bewertungen, die in die Trainingsdatensätze eingeschrieben sind, zu verstehen. Dies erschwert es zusätzlich, einen in den Trainingsdaten eingeschriebenen Bias zu vermeiden.

Rechtsschutz für „geistiges Eigentum"

Bei der Nutzung von KI-Verfahren spielt die Frage nach dem Schutz geistigen Eigentums mindestens in drei Dimensionen eine Rolle:

- Verfahren der KI, durch die neue Artefakte geschaffen werden,
- Produkte, die als technische Schöpfungen durch Verfahren der KI erzeugt werden,
- Verarbeitete Daten, die als Trainingsdaten in die Produkte einfließen und darin enthalten sind, ohne die genaue Herkunft im Regelfall bestimmen zu können.

Das Europäische Parlament hat eine Entschließung[57] zu den Rechten des geistigen Eigentums bei der Entwicklung von KI-Technologien vorgelegt. Dort wird auf die Nicht-Patentierbarkeit mathematischer Methoden hingewiesen[58]. Zu den durch KI erzeugten technischen Methoden kommt sie zur Auffassung,

> *dass durch KI erzeugte technische Schöpfungen gemäß dem Rechtsrahmen für Rechte des geistigen Eigentums geschützt werden müssen, […] ist der Ansicht, dass selbstständig von künstlichen Akteuren und Robotern erzeugte Werke eventuell nicht urheberrechtlich geschützt werden können, da der Grundsatz der Originalität, der mit natürlichen Personen verbunden ist, gewahrt werden muss und der Begriff der „geistigen Schöpfung" an die Person des Autors gebunden ist …*[59]

Die dritte Fragestellung betrifft die Ergebnisse der Verarbeitung von Daten durch Verfahren der Künstlichen Intelligenz, beispielsweise in Form von Trainingsdaten für Systeme des maschinellen Lernens, die potenziell wiederum Urheberrechten unterliegen[60]. In die Modelle können urheberrechtlich geschützte Daten einfließen und die Produkte auf urheberrechtlich geschütztem Material basieren – ohne dass dies im Detail nachvollziehbar ist und die Urheber entsprechend honoriert werden. Das Europäische Parlament stellt dazu fest[61],

> *dass KI-Technologien die Rückverfolgbarkeit von Rechten des geistigen Eigentums und deren Anwendung auf Werke, die durch KI erzeugt wurden, erschweren und somit verhindern, dass Menschen, deren ursprüngliche Arbeit in solchen Technologien zum Einsatz kommt, eine faire Vergütung erhalten.*

Arbeit

Die Debatte, ob der Einsatz von Computern die menschliche Arbeitskraft ersetzen kann und zum Abbau von Arbeitsplätzen führt, ist nicht neu. Nachdem zunächst vor allem einfachere Tätigkeiten durch Computer ausgeführt werden konnten, sind durch Künstliche Intelligenz und maschinelles Lernen inzwischen zunehmend auch hochqualifizierte Berufe betroffen[62].

Zu der zunehmenden Möglichkeit der Übernahme von geistigen Tätigkeiten, beispielsweise der Erstellung von Berichten oder ähnlichem durch Chatbots, kommen die in Personalabteilungen eingesetzten Verfahren der Human Resource Analytics und damit verbunden umfassende Überwachungsmöglichkeiten am Arbeitsplatz[63]. Die Produktivität von Arbeitnehmer:innen wird gemessen und auf mehreren Ebenen aggregiert, um den Erfolg von Management-Entscheidungen feststellen zu können. Gleichzeitig werden die einzelnen Arbeitnehmenden bewertet, beispielsweise um Leistungsträger:innen zu ermitteln, deren Ausscheiden für das Unternehmen besonders kostenintensiv wäre – analog natürlich auch Mitarbeiter:innen, deren Leistung als nicht zufriedenstellend bewertet wird. Hierzu können inzwischen auch biometrische Daten herangezogen werden, die durch Sensoren erfasst werden, oder Geolocation-Technologien zur Standortbestimmung von Mitarbeiter:innen im Außendienst.

Militärische Nutzung

Wie jede Technologie wird auch die KI im militärischen Bereich vorangetrieben – unter anderem mit dem Ziel, möglichst effektiv in militärischen Auseinandersetzungen Menschen zu töten[64]. KI-Verfahren werden dabei in autonomen Waffensystemen eingesetzt, die ohne menschliches Eingreifen militärische Ziele bekämpfen können. Ein Zielbild kann dabei sein, dass der „Feind" selbständig erkannt und bekämpft wird.

Dabei ergeben sich eine Reihe von Fragen: Woran sind „feindliche" Kombattant:innen eigentlich zu erkennen? Wie werden sie von einem Kind und anderen Zivilist:innen unterschieden? Wie ist eine Waffe zu erkennen? Wie werden die Regeln der Genfer Konvention im „intelligenten" System umgesetzt? Wie können „richtige" Entscheidungen sichergestellt werden?

Noch werden autonome Waffen weitgehend abgelehnt – die Letztentscheidung soll auch aus militärischer Sicht beim Menschen liegen. Doch wie kann sichergestellt werden, dass menschliche Entscheider:innen – angesichts sehr kurzer Antwortzeiten – eine autonome Entscheidung treffen können, die über die einfache Bestätigung der maschinellen Empfehlung hinausgeht. Wie können menschliche Entscheider:innen im Zweifel rechtfertigen, sich über diese Empfehlung hinweggesetzt zu haben?

Ein besonderes Risiko ergibt sich aus militärischen Entscheidungssystemen, die auf KI basieren und die computergestützte Reaktionen auf (vermeintliche) militärische Angriffe auslösen. Karl Hans Bläsius und Jörg Siekmann weisen seit Jahren auf das Risiko eines Atomkriegs aus Versehen hin[65], der durch die Fehlfunktion solcher automatisierter Systeme ausgelöst werden kann. Wie wir heute wissen, wurde sehr wahrscheinlich 1983 ein Atomkrieg nur verhindert, weil der sowjetrussische leitende Offizier Stanislaw Jewgrafowitsch Petrow in eigener Verantwortung handelte und einen gemeldeten Angriff US-amerikanischer Atomraketen als Fehlalarm einstufte[66]. Darüber, wie ein automatisiertes Entscheidungssystem in dieser Situation reagiert hätte, können wir nur spekulieren.

Politische Regulierung

Die politische Bedeutung der KI zeigt sich nicht zuletzt darin, dass der 19. Deutsche Bundestag eine Enquête-Kommission eingerichtet hat, die die Auswirkungen der Künstlichen Intelligenz untersuchen sollte, und diese als Ergebnis ihrer Arbeit einen umfassenden Bericht[67] vorgelegt hat.

Gleichzeitig gibt es auf Ebene der Europäischen Union Bestrebungen, Künstliche Intelligenz durch den *AI Act* zu regulieren, der sich vor allem auf die Einhegung von Hochrisikosystemen konzent-

riert[68]. Hier geht es noch nicht darum, dass Künstliche Intelligenz den Menschen überflügelt, sondern um die Einhegung konkreter Risiken beim Einsatz von KI-Systemen. Die Verordnung soll entsprechend das Ziel der Union unterstützen, „bei der Entwicklung einer sicheren, vertrauenswürdigen und ethisch vertretbaren Künstlichen Intelligenz weltweit eine Führungsrolle einzunehmen", und dabei „für den vom Europäischen Parlament ausdrücklich geforderten Schutz von Ethikgrundsätzen" zu sorgen[69]. Weiter heißt es:

> *Abgesehen von den zahlreichen nutzbringenden Verwendungsmöglichkeiten künstlicher Intelligenz kann diese Technik auch missbraucht werden und neue und wirkungsvolle Instrumente für manipulative, ausbeuterische und soziale Kontrollpraktiken bieten. Solche Praktiken sind besonders schädlich und sollten verboten werden, weil sie im Widerspruch zu den Werten der Union stehen, nämlich der Achtung der Menschenwürde, Freiheit, Gleichheit, Demokratie und Rechtsstaatlichkeit sowie der Grundrechte in der Union, einschließlich des Rechts auf Nichtdiskriminierung, Datenschutz und Privatsphäre sowie der Rechte des Kindes.*[70]

Künstliche Intelligenz ist eine faszinierende Technologie mit Chancen, aber auch erheblichen Risiken – da reichen schon die gesellschaftspolitischen Risiken für Bürgerrechte und Datenschutz, ohne gleich an die „Ablösung" des Menschen durch eine übermächtige Technik zu denken. Ein Frühwarnsystem ist wichtig, um Entwicklungen frühzeitig zu erkennen und gegenzusteuern. Wir dürfen das keinesfalls wenigen großen Plattformen überlassen.

Der Beitrag erschien zuerst in der Zeitschrift vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik, Nummer 242 (2/2023), Dezember 2023. Wir danken Autor und Redaktion für die freundliche Genehmigung zum Wiederabdruck.

Anmerkungen

1 *Deutsch Amerikanische Freundschaft (2003) Algorithmus – zahlenlied, in: Deutsch Amerikanische Freundschaft (2003) Fünfzehn neue DAF Lieder, Track 10, superstar recordings*
2 *Russell S, Norvig P (2023) Künstliche Intelligenz. Ein moderner Ansatz, 4. Auflage, München: Pearson*
3 *Kissinger H, Schmidt E, Huttenlocher D (2021) The Age of AI, London: John Murray, S. 3, Übers. S. H.*
4 *Wikipedia, Stichwort ChatGPT, https://de.wikipedia.org/wiki/ChatGPT*
5 *Azhar A (2021) Exponential. How Accelerating Technology is leaving us behind and what to do about it, London: Random House Business*
6 *Hügel S (2019) Künstliche Intelligenz und Politik. Algorithmen, Data Science, Microtargeting – und ihre Auswirkungen auf politische Entscheidungen. vorgänge. Zeitschrift für Bürgerrechte und Gesellschaftspolitik, Nummer 225/226 (1-2/2019), Juli 2019, S. 25ff.*
7 *Weizenbaum J (1987 [1976]) Die Macht der Computer und die Ohnmacht der Vernunft, 2. Auflage. Frankfurt am Main: Suhrkamp, S. 268*
8 *Wikipedia, Stichwort Künstliche Intelligenz, https://de.wikipedia.org/wiki/Künstliche_Intelligenz*
9 *Otte R (2023) Intelligenz und Bewusstsein. Oder: Ist KI wirklich KI?, Aus Politik und Zeitgeschichte, Band 73, Ausgabe 42, S. 9-16*
10 *Gardner HE (1983) Frames of Mind, the theory of multiple intelligences, New York: Basic Books*
11 *Tegmark M (2017) Leben 3.0. Mensch sein im Zeitalter Künstlicher Intelligenz. dt. Übers., Berlin: Ullstein, S. 63*
12 *Turing A (1950) Maschinelle Rechner und Intelligenz, hier zit. nach Hofstadter DR, Dennett DC (1986 [1981]) Einsicht ins Ich. Fantasien und Reflexionen über Selbst und Seele. dt. Übers., Stuttgart: dtv/Klett-Cotta*
13 *Die hier verwendete anthropomorphe Terminologie hat sich etabliert, obwohl sie philosophisch problematisch ist. Eine Diskussion dazu findet sich in Rehak R (2021) The Language Labyrinth: Constructive Critique on the Terminology Used in the AI Discourse. In: Verdegem P Hg. (2021) AI for Everyone? Critical Perspectives. London: University of Westminster Press, S. 87-102, DOI: https://doi.org /10.16997/book55.f.*
14 *Paaß G, Hecker D (2020) Künstliche Intelligenz. Was steckt hinter der Technologie der Zukunft? Wiesbaden: Springer Vieweg*
15 *Levine EV (2023) Cargo Cult AI. Communications of the ACM, Vol. 66, No. 9, S. 46*
16 *Wikipedia, Stichwort Tay (Bot), https://de.wikipedia.org/wiki/Tay_(Bot)*
17 *Affsprung D (2023) The ELIZA Defect: Constructing the Right Users for Generative AI, in: AAAI/ACM Conference on AI, Ethics and Society (AIES '23), August 08–10, 2023, Montréal, QC, Canada. ACM, New York, NY, USA, https://doi.org/10.1145/3600211.3604744*
18 *El Atillah I (2023) Man ends his life after an AI chatbot 'encouraged' him to sacrifice himself to stop climate change, Euronews.next, March 31, 2023. https://www.euronews.com/next/2023/03/31/man-ends-his-life-after-anai-chatbot-encouraged-him-to-sacrifice-himself-to-stop-climate-*
19 *Prompting bezeichnet die Interaktion mit einem Chatbot, indem Anfragen an ihn gestellt werden. Prompt Injection bezeichnet das „Unterschieben" von Anfragen in böswilliger Absicht, um dem Chatbot Antworten zu entlocken, die von dessen Entwickler:innen nicht vorgesehen sind.*
20 *Szöke D (2023) Prompt Injection: Marvin von Hagen trägt vor, wie er Bing Chat austrickste, Heise online, https://www.heise.de/news/Prompt-Injection-Marvin-von-Hagen-traegt-vor-wie-er-Bing-Chat-austrickste-9210511.html*
21 *Schmalzried G (2023) Wie ein Münchner Student zur Zielscheibe von Microsofts KI wurde, Bayerischer Rundfunk, BR24, https://www.br.de/nachrichten/netzwelt/microsoft-ki-bing-chatgpt-muenchner-student-als-zielscheibe*
22 *O'Neil C (2016) Weapons of Math Destruction. How Big Data increases inequality and threatens Democracy, London: Random House*
23 *Zweig KA (2019) Ein Algorithmus hat kein Taktgefühl. Wo künstliche Intelligenz sich irrt, warum uns das betrifft und was wir dagegen tun können, München: Heyne*
24 *Zweig KA (2023) Die KI war's! Von absurd bis tödlich: Die Tücken der künstlichen Intelligenz, München: Heyne*
25 *Hawking SW (2018) Kurze Antworten auf große Fragen. dt. Übers., Stuttgart: Klett-Cotta, S. 208*

Stefan Hügel

Stefan Hügel ist Vorsitzender des FIfF, arbeitet als IT-Berater und lebt in Frankfurt am Main.

26 Hawking SW (2018) a. a. O., S. 211
27 Jonas H (1979) Das Prinzip Verantwortung. Versuch einer Ethik für die technologische Zivilisation. Frankfurt am Main: Suhrkamp, S. 36
28 Safe.ai (2023) Statement on AI Risk, https://www.safe.ai/statement-on-ai-risk
29 Futureoflife.org (2023) Pause Giant AI Experiments: An Open Letter, https://futureoflife.org/open-letter/pause-giant-ai-experiments/
30 Paul K and agencies (2023) Letter signed by Elon Musk demanding AI research pause sparks controversy. The Guardian, https://www.theguardian.com/technology/2023/mar/31/ai-research-pause-elon-musk-chatgpt
31 Yudkowsky E (2023) Pausing AI Developments isn't enough. We need to shut it all down. Time, https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/
32 Yudkowsky E (2023) a. a. O.
33 Deutscher Bundestag (2020) Bericht der Enquête-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale, BT-Drs 19/23700, https://dserver.bundestag.de/btd/19/237/1923700.pdf
34 Moreno J (2023) „Die Menschheit kreiert derzeit eine intelligentere Spezies", Podcast mit Hannah Bast, Spiegel online, https://www.spiegel.de/netzwelt/ki-forscherin-hannah-bast-die-menschheit-kreiert-derzeit-eine-intelligentere-spezies-podcast-a-b9b78548-b47c-46ff-86d1-c1a7ed2069fd
35 Tegmark M (2017) a. a. O., S. 52
36 Stahl BC (2023) Grauzonen zwischen Null und Eins. KI und Ethik, Aus Politik und Zeitgeschichte, Band 73, Ausgabe 42, S. 17-22
37 Futureoflife.org (2017) Asilomar AI Principles, https://futureoflife.org/open-letter/ai-principles-german/
38 Rehbein M (2018) Die „Asilomar AI Principles" zu Künstlicher Intelligenz, FIfF-Kommunikation, 35. Jahrgang, Heft 3, S. 24
39 Wikipedia, Stichwort Trolley-Problem, https://de.wikipedia.org/wiki/Trolley-Problem
40 Misselhorn C (2018) Grundfragen der Maschinenethik, Stuttgart: Reclam
41 Shi-Kupfer K (2023) Digit@l China. Überwachungsdiktatur und technologische Avantgarde, München: C. H. Beck
42 Kühnreich K (2022) Social Credit Systems und Gamification. Digitale Gesellschaft, Netzpolitischer Abend, Vortragsaufzeichnung, https://media.ccc.de/v/dgna-4248-social-credit-systems-und-gami
43 Angesichts der Enthüllungen von Edward Snowden müssen wir aber davon ausgehen, dass massive staatliche Überwachung auch in demokratisch konstituierten Gesellschaften an der Tagesordnung ist. Vgl. Greenwald G (2014) Die globale Überwachung. Der Fall Snowden, die amerikanischen Geheimdienste und die Folgen, München: Droemer-Verlag
44 https://www.youtube.com/watch?v=iAE6dCE7URg
45 Louban A, Tahraoui M, Aden H, Fährmann J, Krätzer C, Dittmann J (2022) Das Phänomen Deepfakes. Künstliche Intelligenz als Element politischer Einflussnahme und Perspektive einer Echtheitsprüfung, in: Friedewald M, Roßnagel A, Heesen J, Krämer N, Lamla J Hg. (2022) Künstliche Intelligenz, Demokratie und Privatheit, Baden-Baden: Nomos, S. 265ff., https://www.nomos-elibrary.de/10.5771/9783748913344/kuenstliche-intelligenz-demokratie-und-privatheit
46 Dachwitz I (2023) Das sind die 650.000 Kategorien, in die uns die Online-Werbeindustrie einsortiert, netzpolitik.org, https://netzpolitik.org/2023/microsofts-datenmarktplatz-xandr-das-sind-650-000-kategorien-in-die-uns-die-online-werbeindustrie-einsortiert/
47 Pariser E (2012 [2011]) Filter Bubble. Wie wir im Internet entmündigt werden, dt. Übers., München: Hanser
48 Leisegang D (2023) Algorithmen rütteln kaum an politischen Einstellungen, netzpolitik.org, https://netzpolitik.org/2023/studien-zu-facebook-und-instagram-algorithmen-ruetteln-kaum-an-politischen-einstellungen/
49 Aden H, Kleemann S, Hirsbrunner SD (2023) Fairness, Erklärbarkeit und Transparenz bei KI-Anwendungen im Sicherheitsbereich – ein unmögliches Unterfangen?, vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik, Nummer 242 (2/2023), Dezember 2023
50 Futureoflife.org (2017) a. a. O., Punkte 7 (Transparenz bei Fehlfunktionen), 8 (Transparenz bei Rechtsprechung), 9 (Verantwortung), 16 (Menschliche Kontrolle)
51 Holzinger A (2018) Explainable AI (ex-AI), Aktuelles Schlagwort, Informatik-Spektrum Band 41, Heft 2, S. 138ff.
52 Zweig KA (2023) a. a. O., S. 257ff.
53 Schinzel B (2022) Diskriminierung durch digitale Entscheidungsstrukturen, Aus Politik und Zeitgeschichte, Band 72, Ausgabe 10-11, S. 26-34
54 Pumhösel A (2020) Gender-Bias: Schlechtere Jobchancen für Frauen durch Algorithmen, derstandard.at, https://www.derstandard.at/story/2000115720676/gender-bias-schlechtere-job-chancen-fuer-frauen-durch-algorithmen
55 Heikklä M (2022) Dutch scandal serves as a warning for Europe over risks of using algorithms. Politico, 29. März 2022, https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/
56 Wolfangel E (2018) Programmierter Rassismus, Zeit online, https://www.zeit.de/digital/internet/2018-05/algorithmen-rassismus-diskriminierung-daten-vorurteile-alltagsrassismus
57 Europäisches Parlament (2020) Rechte des geistigen Eigentums bei der Entwicklung von KI-Technologien, Entschließung, P9_TA(2020)0277, https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_DE.html.
58 Europäisches Parlament (2020) a. a. O., Absatz 11
59 Europäisches Parlament (2020) a. a. O., Absatz 15
60 Koep-Kerstin W, Dingeldey P (2023) Zwischen Falschinformation, menschenrechtlichen Problemen und kreativer Denkleistung. Ein Interview mit dem Chatprogramm ChatGPT, vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik, Nummer 242 (2/2023), Dezember 2023
61 Europäisches Parlament (2020) a. a. O., Erwägungsgrund D
62 Vgl. hierzu auch Lamberth A (2023) „Intelligente" Roboter und die Humanisierung von Arbeit, und Dingeldey P (2023) Das Recht auf Faulheit im Lichte künstlicher Intelligenz und Koep-Kerstin W, Dingeldey P (2023) a. a. O., in vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik, Nummer 242 (2/2023), Dezember 2023
63 Waas B (2023) Künstliche Intelligenz und Arbeitsrecht, Hans-Böckler-Stiftung, HSI-Schriftenreihe, Band 26, Frankfurt am Main: Bund-Verlag, https://www.hugo-sinzheimer-institut.de/faust-detail.htm?sync_id=HBS-008472, S. 27ff
64 Kreowski HJ, Lye A (2023) Künstliche Intelligenz im Dienst des Militärs, in diesem Heft, S. 19-23
65 Bläsius KH, Siekmann J (2023) Ist die Künstliche Intelligenz gefährlich? FIfF-Kommunikation, 40. Jahrgang, Heft 3, S. 9 bzw. W&F Wissenschaft und Frieden, 41. Jahrgang, Heft 3, S. 28
66 Wikipedia, Stichwort Nuklear-Fehlalarm von 1983, https://de.wikipedia.org/wiki/Nuklear-Fehlalarm_von_1983
67 Deutscher Bundestag (2020) a. a. O.
68 Europäische Union (2021) The Artificial Intelligence Act, https://artificialintelligenceact.eu/the-act/.
69 Europäische Union (2021) a. a. O., Erwägungsgrund 5
70 Europäische Union (2021) a. a. O., Erwägungsgrund 15

Hans-Jörg Kreowski und Aaron Lye

Künstliche Intelligenz im Dienst des Militärs

Einleitung

Die Anfänge der Computertechnik in Deutschland, Großbritannien und dann vor allem in den USA am Ende des Zweiten Weltkriegs und in den ersten Jahren danach waren stark bestimmt von den Wünschen und vom Geld des militärischen Komplexes. Die ersten Jahre der Künstlichen Intelligenz (KI) waren dagegen eher zivil geprägt. 1956 trafen sich zwölf junge aufstrebende Wissenschaftler um John McCarthy für einige Wochen im Dartmouth College in New Hampshire zu einem von der Rockefeller-Stiftung geförderten Workshop und begründeten das Fachgebiet der Künstlichen Intelligenz (vgl. McCarthy et al. 1955).

Erklärtes Ziel war, kognitive Fähigkeiten wie Sehen und Erkennen, Hören und Verstehen, Planen und Entscheiden, Lernen sowie Problemlösen mit Hilfe von Computerprogrammen zu simulieren – seit damals kaum verändert. Im gesamten Antrag findet sich kein Bezug zu Rüstung und Krieg. Bis in die 1980er-Jahre hinein bleibt die KI zivil ausgerichtet. Nachdem die ersten Versuche, einen *General Problem Solver* zu kreieren, gescheitert waren, lag der KI-Fokus lange auf der Entwicklung von Expertensystemen, die für spezielle Anwendungen das Wissen von Fachleuten nutzbar machen sollten – teils auch mit Erfolg.

Nachdem Japan mit dem *Fifth-Generation-Projekt* Anfang der 1980er-Jahre einen großangelegten Versuch gestartet hatte, durch geeignete Computertechnik und Programmiersprachen KI anwendbar zu machen, haben die USA mit der *Strategic Computing Initiative* (SCI) geantwortet, um die technologische Vormachtstellung zu sichern. SCI war gleichzeitig der Einstieg in die militärische Nutzung von KI und soll deshalb im nächsten Abschnitt ausführlicher dargestellt werden. Ein zentrales Projekt von SCI war die Entwicklung autonomer Roboterfahrzeuge. Da solche Systeme – oft Drohnen genannt – bis heute in der weltweiten Rüstung eine prominente Rolle spielen und ihre Entwicklung noch lange nicht ihren Höhepunkt erreicht hat, wird die Thematik in einem eigenen Abschnitt vertieft. Mehr in die Zukunft gerichtet, gehen wir auf die neuesten Entwicklungspläne Deutschlands, Frankreichs und Spaniens für ein Kampfflugzeug ein, bei denen KI insgesamt und Kriegsdrohnen speziell eine übergeordnete Rolle spielen sollen. Der öffentlichen Aufmerksamkeit meist verborgen, aber militärisch mindestens ebenso relevant wie Drohnen sind die Entwicklungen bei Schlachtenmanagementsystemen, die wir als letztes Thema ausführlich behandeln. Der Artikel endet mit einem Ausblick, in dem auch verschiedene Punkte angesprochen werden, die nicht ausgeführt werden konnten.

Strategic Computing Initiative

Bis in die 1980er-Jahre hinein verlief die Entwicklung der KI anders als andere Teilgebiete der Informatik vergleichsweise zivil. Das änderte sich gewaltig, nachdem Japan 1982 das *Fifth Generation Computer Systems* (FGCS) *Project* ins Leben gerufen hat, um eine Führungsrolle im IT-Sektor zu erreichen (siehe z. B. Shapiro 1983 und Odagiri et al. 1997). Das Projekt war auf zehn Jahre angelegt mit dem Ziel, eine neuartige Computertechnologie zu entwickeln, die massive Parallelität mit logischer Programmierung verbindet und als Plattform für zukünftige KI-Anwendungen dienen sollte. Datenverarbeitung sollte mit Hilfe logischen Schließens als Programmierparadigma zu Wissensverarbeitung werden. Um die Benutzungsfreundlichkeit zu steigern, sollten Ein- und Ausgabe über Alltagssprache, Graphiken, Bilder und Dokumente möglich sein. Als Grundlagenforschung war das Projekt ein Erfolg, kommerziell ist es eher gescheitert, weil sich die hohen Ansprüche zu der Zeit technisch nur bedingt umsetzen ließen.

In den USA wurde das japanische FGCS-Projekt als Angriff auf die eigene technologische Vormachtstellung empfunden und mit der *Strategic Computing Initiative* (SCI) beantwortet (siehe z. B. Roland & Shiman 2002). Während FGCS zivil angelegt war, gilt für SCI das genaue Gegenteil: Es war durch und durch militärisch ausgerichtet. Das US-Verteidigungsministerium setzte auf das militärische Potenzial der KI und startete 1983 SCI, wobei drei Aufgaben im Zentrum standen: ein Sprachassistent für die Piloten der Luftwaffe, ein Schlachtenmanagementsystem für die Marine und autonome Landfahrzeuge für das Heer. SCI war auf zehn Jahre angelegt und hatte einen für damalige Verhältnisse gigantischen Finanzrahmen von fast einer halbe Milliarde US-Dollar. Da sich keine schnellen Erfolge abzeichneten, wurde das Programm noch vor dem Ende gekürzt. Dennoch muss SCI wohl als Ausgangspunkt einer beispiellosen Entwicklung der KI gesehen werden.

Da die drei Aufgabenstellungen von SCI sowohl von der Art der Anwendungen als auch von den intendierten kognitiven und intelligenten Leistungen her bis heute typische Entwicklungen im militärischen – aber auch zivilen Kontext – darstellen, soll darauf etwas näher eingegangen werden.

SCI war als generisches Forschungsprojekt mit den Schwerpunkten auf Bildverstehen, Spracherkennen und -verstehen sowie Expertensystemen konzipiert. Ziel des Bildverstehens war, Veränderungen in Bildszenen in Echtzeit zu erkennen, Objekte und Ereignisse in Szenen zu identifizieren und graphisch oder mit Hilfe natürlicher Sprache zu charakterisieren. Auf der Basis neuer Chip-Designs, paralleler Rechnerarchitekturen und Algorithmen sollte ein wissensbasiertes Sehen zur Erkennung von Landmarken, Vermeidung von Hindernissen, Geländeverfolgung, Gelände- und Objektmodellierung und zur Zielfindung realisiert werden. Ziel des Spracherkennens und -verstehens war, das gesprochene Wort als Computereingabe zu ermöglichen, wobei Sprache in Telefonqualität selbst in geräuschvoller Umgebung wissensbasiert mit hoher Genauigkeit verstanden werden sollte. Umgekehrt sollte auch natürlichsprachliche Maschinenausgabe in eingeschränkten Kontexten realisiert werden. Als technische Herausforderungen galten die Wissensakquisition, das Textverstehen, die Sprachgeneration, den Redekontext einzuordnen sowie die Entwicklung paralleler Algorithmen für

diese Anwendungen. Was die Technologie der Expertensysteme angeht, wurde das Ziel einer neuen Generation zur Unterstützung der Schlachtenmanagementsysteme für das US-Verteidigungsministerium ausgegeben. Dabei sollten mächtige Schluss-, Kontroll- und Erklärungsmechanismen zur Verfügung gestellt, automatische Wissensakquisition und Unsicherheit einbezogen sowie Informationen aus verschiedenen Quellen fusioniert werden. Als computertechnische Basis waren fortgeschrittene LISP-Maschinen vorgesehen, die wesentlich mehr Regeln zu verarbeiten erlauben sollten, als bis dahin machbar war.

Methodisch und technisch unterscheidet sich dieser Ansatz nicht vom damaligen Stand der Entwicklung im zivilen Bereich. Das übergeordnete Ziel und die drei Anwendungsprojekte waren dagegen ganz auf militärische Bedürfnisse zugeschnitten. Angestrebt wurde die Entwicklung einer breiten Basis an Maschinenintelligenztechnologien, um die nationale Sicherheit zu verbessern. Die Sprachassistenz für Kampfpiloten setzte vor allem auf Hören, Verstehen und Sprechen auf, wobei die Verarbeitung natürlicher Sprache mit Hilfe von Expertensystemen bewerkstelligt werden sollte. Das Schlachtenmanagementsystem für die Marine hatte die Expertensystemtechnologie als Mittelpunkt, die regelbasiert für das Datensammeln, Schließen und Planen zuständig waren. Essenziell für die autonomen Roboterfahrzeuge sind Sensoren wie Kameras, 3D-Scanner und Sonare, Bildverarbeitung und -verstehen sowie die Fusion mit anderen Sensordaten, ohne die Navigation in unbekanntem und unwegsamem Gelände, Landmarkenerkennung, Routenplanung, Hindernisvermeidung, visuelle Aufklärung und Zielfindung unmöglich sind. Für die Datenverarbeitung selbst waren wieder Expertensysteme gedacht, wobei mit rund 5000 Regeln gerechnet wurde, die hundertmal schneller ausgeführt werden sollten als bis dahin möglich.

Am Ende der 1980er-Jahre blieb man noch weit hinter dem Erreichen dieser Ziele zurück. Inzwischen hat die Verarbeitung von Sprache und Text ein hohes Niveau erreicht, genauso wie die Verarbeitung von Bildern, wie sie insbesondere für das eigenständige Navigieren von Fahrzeugen benötigt wird. Und Managementsysteme sind inzwischen weit verbreitet. Dabei ist festzuhalten, dass sich die Ziele und auch die Methoden seit den Anfängen der KI gar nicht allzu sehr verändert haben, aber der heute verfügbare Speicherplatz und die Rechengeschwindigkeit erlauben solche Anwendungen. So beruht der Erfolg von Sprachverarbeitungssystemen und Bildverarbeitungssystemen auf maschinellem Lernen mit Hilfe neuronaler Netze, die wahrscheinlichkeitstheoretische Verfahren unter Verwendung riesiger Datenmengen nutzen. Die Idee stammt aus den 1940er-Jahren und wurde in vielen Entwicklungsschritten zu heutiger Reife vorangetrieben. SCI war ein Zwischenschritt, aber von großer Schubkraft.

Militärische Drohnen

Die Idee unbemannter Fahr- und Flugzeuge für militärische Zwecke reicht weit zurück. Schon vor dem und im Zweiten Weltkrieg gab es mehrere derartige Entwicklungen mit sehr begrenzter Wirkung. Das SCI-Projekt zum Bau autonomer Robotervehikel in Form unbemannter Landfahrzeuge hat dagegen eine breite, bis heute fortgesetzte Entwicklung von Drohnen aller Art eingeläutet. Die USA planen einen erheblichen Teil ihrer Bewaffnung auf unbemannte Systeme umzustellen, wie es in der *Unmanned Systems Integrated Roadmap* (Department of Defense 2018) dargestellt wird. Andere Nationen folgen diesem Weg.

Im Folgenden soll die Entwicklung unbemannter fliegender Vehikel skizziert werden. Die Entwicklungen bei unbemannten Fahrzeugen, Booten und U-Booten ist etwas anders, darauf wird aber aus Platzgründen verzichtet.

Unbemannte fliegende Systeme sind bereits seit 20 Jahren im Einsatz. Angefangen hat es mit Aufklärungsdrohnen, die feindliches Gebiet überfliegen und ausspionieren können. Es hat dann nicht lange gedauert, bis solche Drohnen mit Raketen ausgestattet wurden. Vorreiter waren Israel mit Einsätzen in Gaza und vor allem auch die USA mit Zehntausenden kriegsvölkerrechtswidrigen Einsätzen in Afghanistan, Pakistan, Jemen, Somalia und im Irak im „Krieg gegen den Terror". Seit einigen Jahren werden Drohnen zunehmend auch in herkömmlichen Kriegen eingesetzt.

Inzwischen sind Aufklärungs- und Killerdrohnen weit verbreitet, und es gibt sie in einer breiten Palette von klein bis groß und von billig bis teuer. Die USA haben ihre Großdrohnen mit erheblicher Reichweite, schwerer Bewaffnung und mit Stückpreis von ein- und zweistelligen Millionenbeträgen nur an Großbritannien und Frankreich verkauft. Deutschland hat eine kleine Zahl von Israel geleast. Zu den Herstellerstaaten gehören inzwischen auch u. a. China, Iran, Russland und die Türkei. Mehrere Dutzend Staaten haben diese Waffen angeschafft und setzen sie teilweise auch bereits ein. Der Krieg Russlands gegen die Ukraine zeigt, dass bewaffnete Drohnen auch eine Rolle spielen, wenn zwei Armeen gegeneinander kämpfen (vgl. Kreowski 2022 und Marischka 2022). Da bewaffnete Drohnen langsam und niedrig fliegen, sind sie relativ gut mit Luftabwehr zu bekämpfen. Greift eine Partei allerdings mit einer größeren Zahl an Drohnen gleichzeitig an, ist die Gefahr groß, dass einzelne Drohnen ihr Ziel erreichen.

Aufklärungs- und Killerdrohnen sind Waffensysteme, die es ohne KI nicht gäbe. Die aktuelle Entwicklung belegt, dass eine gewaltige Rüstungsspirale eingesetzt hat mit einem erheblichen Proliferationsproblem. Es gibt inzwischen ein weitgefächertes Arsenal an Aufklärungs- und Killerdrohnen, das von der Riesendrohne *Global Hawk* bis zu den nur wenige tausend Dollar teuren tragbaren *Switchblades* reicht. Kleine (teils umgebaute kommerzielle) Drohnen mit Sprengsätzen und Kampfdrohnen als einfacher Bausatz stellen Waffen dar, die sich jeder Kontrolle entziehen.

Der nächste Entwicklungsschritt, an dem intensiv gearbeitet wird, ist die vollständige Autonomie, bei der dann die Bordsysteme auch über den Waffeneinsatz entscheiden. Technisch handelt es sich um ein Software-Update, das möglicherweise längst in diverse unbemannte Systeme eingebaut ist, aber bisher nicht (oder nur in Einzelfällen) aktiviert wurde. In Politik und Militär sind autonome Waffen aus ethischer und kriegsvölkerrechtlicher Sicht nicht unumstritten, weil die mit dem Waffeneinsatz verbundene Entscheidung über Leben und Tod nicht mehr direkt von Menschen getroffen wird, sondern indirekt schon bei der

Programmierung (siehe z. B. Williams & Scharre 2015). Ein Verbot ist dennoch nicht absehbar; entsprechende Verhandlungen im Rahmen der *Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems* der Vereinten Nationen scheitern bisher am „Nein" Russlands, der USA und vieler anderer Staaten.

Einige autonome KI-Systeme sind seit Jahrzehnten im Einsatz. Die meisten autonomen Systeme, die derzeit von den NATO-Streitkräften eingesetzt werden, sind regelbasierte Expertensysteme, darunter autonome vorprogrammierte Fahrzeuge, Luft- und Raketenabwehrsysteme und autonome Flugkörper. So verfügen beispielsweise sowohl das amerikanische *Patriot*-Luftabwehrraketensystem als auch das *Aegis*-Verteidigungssystem, die seit den 1990er- bzw. 1980er-Jahren im Einsatz sind, über halbautonome und vollautonome Modi, die es Computern ermöglichen, ankommende Bedrohungen ohne menschliche Zustimmung zu erkennen, anzuvisieren und anzugreifen. Darüber hinaus können eine Reihe unbemannter Luftfahrzeuge, darunter die israelische Drohne *Harop* und die amerikanische *RQ-11 Raven*, die seit Anfang der 2000er-Jahre im Einsatz sind, autonom fliegen, wobei sie sich oft auf vorprogrammierte Flugrouten und Ziele verlassen (vgl. Gray & Ertan 2021).

KI-Automatisierung des Luftkampfs

Flugkampfassistenzsysteme werden kontinuierlich weiterentwickelt und die Überlegungen, die diesbezüglich in der SCI entstanden, sind weitestgehend überholt. Mehrere aktuelle Flugzeugsysteme enthalten KI-Komponenten. Die *F-35 Lightning II* von Lockheed Martin hat beispielsweise Systeme zur Entscheidungsunterstützung und Datenanalysesysteme (vgl. Osborn 2017a). In ähnlicher Weise werden zwei europäische Flugzeugprojekte der nächsten Generation, der *Next Generation Fighter* (NGF) des *Future Combat Air System* (FCAS) und *Tempest* von BAE, KI-Komponenten enthalten. BAE Tempest wird über ein KI-gestütztes autonomes Flugsystem verfügen (vgl. Adams 2018). Das Flugkampfassistenzsystem des FCAS-Projekts wird unter dem Acronym ASTARTES entwickelt. ASTARTES steht für *Air Superiority Tactical Assistance Real Time Execution System*. Die Beschreibung des Ziels für diese Komponente bleibt abstrakt. Bisher ist lediglich die Rede davon, menschliche Erfahrungen zu digitalisieren, um die Operator:innen in Taktik zu unterstützen. Aus anderen KI-unterstützten Feuerkampfsystemen lässt sich aber das eine oder andere als wahrscheinliche Konkretisierung ableiten. Durch das von Rheinmetall entwickelte System *Attac*, welches für zukünftige Gefechtsfahrzeuge entwickelt wird, soll die „Fahrzeugbesatzung im Bereich der Beobachtung und Zielerfassung sowie in der Entscheidungsfindung und Wirkung entlastet werden", indem die Fahrzeuge „die Fähigkeit (erlangen), selbstständig aufzuklären, erkannte Objekte zu klassifizieren und in die Bedrohungslage in Echtzeit einzuordnen" (Marischka 2023).

Allerdings ist derzeit sogar nicht einmal ausgeschlossen, dass der NGF unbemannt sein könnte. Die Wahrscheinlichkeit dafür ist gar nicht so klein. Das zeigen jüngste Entwicklungen: Im August 2020 hat in einer Virtual-Reality-Simulation eines F-16-Luftkampfwettbewerbs eine von *DeepMind* (Google) entwickelte KI einen menschlichen Piloten mit 5 zu 0 besiegt (siehe Knight 2020). Nachfolgend wurden in weniger als drei Jahren die KI-Algorithmen, die im Rahmen des *Air Combat Evolution-Programms* (ACE) der DARPA entwickelt wurden, von der Steuerung simulierter F-16, die Luftkämpfe auf Computerbildschirmen fliegen, zur Steuerung einer tatsächlichen F-16 im Flug weiterentwickelt (DARPA 2023).

Bei FCAS wird mit dem *Next Generation Weapon System* noch weiter gedacht. Beim FCAS handelt es sich um ein gigantomanisches europäisches Rüstungsprojekt, das seit der Unterzeichnung eines gemeinsamen Investitionsplans im September 2021 von Deutschland, Frankreich und Spanien auf den Weg gebracht wird und dessen Kosten voraussichtlich im dreistelligen Milliardenbereich liegen werden. Das FCAS soll ein Verbundsystem mehrerer Luftkampfeinheiten werden, bei dem die KI eine zentrale Rolle spielen soll. Das FCAS gilt als wichtiger Schritt zur KI-Automatisierung des Krieges (siehe Kreowski & Lye 2022). Die Hauptkomponente des FCAS ist das in Teilen schon oben beschriebene Kampfflugzeug (NGF). Das Projekt umfasst aber weit mehr. Es soll ein System von Systemen werden, ein sogenanntes *Next Generation Weapon System*, bei dem Kampfjets zusammen mit Schwärmen von autonomen Lenkflugkörpern und bewaffneten und unbewaffneten Drohnenschwärmen in einer Weise eingesetzt werden sollen, die Europa dem Plan gemäß in allen Teilen der Welt die Überlegenheit im Luftkampf sichern soll (Renn 2021).

Die Interaktion von Soldat:innen mit Robotern und unbemannten Systemen während des Kampfgeschehens wird als *Manned-Unmanned-Teaming* (MUM-T) bezeichnet (siehe Department of Defense, 2013, S. 139 und Pletsch 2020). Airbus demonstrierte im Jahr 2018 die Möglichkeiten des MUM-T mit fünf von dem Unternehmen gebauten Drohnen des Typs Do-DT 25, die autonom agierten und von einem Piloten befehligt wurden, der sich in einem bemannten Führungsflugzeug (Learjet) in der Luft befand (Airbus 2018). Die Tests beinhalteten Formationsflüge, das Ausweichen vor einer simulierten Bedrohung, simulierte Aufklärung und die Kompensation einer im Flug ausgefallenen Drohne. Autonom handelnde bewaffnete Drohnen sollen sich als Schwärme formieren und gemeinsam operieren können. Letzteres wird als *Swarming* bezeichnet. Um *Teaming* und *Swarming* zu realisieren, sind mehrere Fähigkeiten und Techniken erforderlich, so unter anderem Teaming-/Swarming-Algorithmen, neue Sensoren oder Missionsmanagementsysteme für die Führungsunterstützung durch die Besatzung des bemannten Flugzeugs. Das von Airbus bisher entwickelte Flugmanagementsystem für unbemannte Luftfahrzeuge kombiniert schon heute vollautomatische Lenkung und Navigation mit Schwarmfähigkeiten.

In Zukunft wird die F-35 wahrscheinlich auch KI einsetzen, um unbemannte Drohnen zu steuern, die Waffen tragen, Überwachung und Aufklärung betreiben oder feindliche Luftabwehr testen können (vgl. Osborn 2017b). Gleiches gilt für FCAS.

Es deutet sich eine Transformation zu neuen militärischen Taktiken an, die zum einen auf kooperativem Kampf vieler einzelner autonomer Systeme und zum anderen auf dem Einsatz von Täuschung und zahlenmäßiger Überlegenheit beruhen. Ziel scheint es hier zu sein, den Gegner möglichst großflächig zu stören oder durch zahlenmäßige Übermacht zu überwältigen. Wenn die Pla-

nungen in die Realität umgesetzt werden, versprechen sie dem Militär vor allem auch eine verbesserte Effizienz, indem sichergestellt wird, dass die für einen bestimmten Einsatz erforderliche Mischung der Fähigkeiten auch zielgenau zur Verfügung steht. Zudem könnte das System die Risiken für den Einsatz menschlicher Kampfeinheiten reduzieren, da die bemannten Einheiten in sicherer Entfernung bleiben könnten, während die Drohnen konkret die Front bilden (Airbus 2020a). In welche Richtung gedacht wird, zeigt das Positionspapier zur Anwendung Künstlicher Intelligenz in den Landstreitkräften der Bundeswehr (Amt für Heeresentwicklung 2019). In einem geschilderten Zukunftsszenario wird eine Situation erdacht, in der bis zu 15.000 autonome unbemannte Luftfahrzeuge unterschiedlicher Bauart und Größe sich zu Schwärmen formieren, um verschiedene Aufgaben auszuführen.

Schlachtenmanagement und Combat Clouds

Das in der SCI vorgeschlagene Schlachtenmanagementsystem bestand aus unterschiedlichen Expertensystemen (vgl. Roland & Shiman 2002, S. 194f). Das *Force Requirements Expert System* (FRESH) würde die Einsatzbereitschaft der Flotte überwachen und bei der Aufteilung der Kräfte entsprechend den Fähigkeiten und dem Status der einzelnen Schiffe helfen. Das *Capabilities Assessment Expert System* (CASES) würde die relative Stärke der US-amerikanischen und der feindlichen Streitkräfte vergleichen. Das *Campaign Simulation Expert System* (CAMPSIM) würde das Ergebnis verschiedener Handlungsoptionen simulieren und entsprechende Empfehlungen benennen. Das *Operations Plan Generation Expert System* (OPGEN) würde Operationspläne nach vorgegebenen Strategien entwickeln. Das *Strategy Generation and Evaluation Expert System* (STRATUS) schließlich würde bei der Entwicklung von Plänen für die Strategie auf der Ebene des Einsatzgebietes helfen.

Grundsätzlich hat sich an dieser Ratio bis heute wenig geändert und auch prinzipiell sind die Anforderungen an die Systeme auf andere klassische Militärdomänen übertragbar. Aber es gibt auch wesentliche Weiterentwicklungen. Mittlerweile sind bei vielen Militärs die Systeme, Sensoren und Wirkmittel miteinander in (nahezu) Echtzeit vernetzt. In großangelegten Projekten wird dies aber immer noch weiter vorangetrieben. Ein Projekt der Bundeswehr ist die *Digitalisierung landbasierter Operationen* (D-LBO), durch das perspektivisch „bis zu 25.000 Fahrzeuge und 155.000 Soldaten" untereinander vernetzt agieren können sollen (Marischka 2023). Wesentlich hierfür sind Informationsverbünde, für die sich der Begriff *Combat Clouds* etabliert hat. Die Erfassung von Daten von einem Luftfahrzeug und aus dem Weltraum aus ist wesentlich für militärische Aufklärung und Entscheidungsprozesse. Das FCAS-Projekt umfasst daher die Entwicklung einer vernetzten und verteilten Architektur von Sensoren, den Entwurf zukünftiger Sensorarchitekturen und die Entwicklung der zugehörigen Sensortechnologien. Beim FCAS soll die sogenannte *Air Combat Cloud* alle Flugsysteme verbinden, um die Verarbeitung und Verteilung von Informationen und Instruktionen nahezu in Echtzeit zu ermöglichen (Airbus 2020b). Die dafür wesentliche Infrastruktur stellen Satelliten im Weltraum dar – insbesondere Satelliten in niedriger Erdumlaufbahn, welche die Hochgeschwindigkeitsdatenübertragung zwischen den Systemen ermöglichen. Des Weiteren wird in den Planungsunterlagen für FCAS der Zugang zu satellitengestütztem Bildmaterial nahezu in Echtzeit für die Lageeinschätzung als Ziel benannt. Dazu passt auch die geplante neue Satellitengeneration für Informationserfassung, Satellitenaufnahmen und Kommunikationsübertragung.

Eine andere Entwicklung ist ebenfalls bemerkenswert. Sie lässt sich allerdings vor allem in den USA nachzeichnen. Seit einer Dekade findet eine gravierende Verlagerung geheimdienstlicher und militärischer Daten in die Cloud kommerzieller Anbieter statt (siehe Schelling 2019 und Lye 2021). Die Beziehungen von Big-Tech zu Geheimdiensten und Pentagon wurden über Jahrzehnte aufgebaut. Nachdem sich vor allem der Marktführer *Amazon Web Services* im Aufbau, dem Umzug und der Etablierung positioniert hatte, ist jetzt eine deutliche Diversifizierung bemerkbar, sodass mittlerweile alle großen KI- und Clouddiensteprovider entsprechende Verträge haben. Die Projekte stellen nicht einfach nur weitere und neue Kommunikations- und Infrastrukturprojekte der Geheimdienste und des Pentagons dar. Stattdessen zeigen sie auf, dass die Großkonzerne des Silicon Valley eine wesentliche Rolle in der Kriegsführung spielen. KI-gestützte Kriegsführung mit der Auswertung von immer mehr Sensordaten, um militärische Gesamtsituationen zu erfassen, logistische und taktische Optimierung sowie automatisierte Entscheidungen von Truppenbewegungen und Angriffen werden forciert entwickelt und einsetzbar gemacht.

Ausblick

In diesem Artikel haben wir ausgehend von SCI die militärischen KI-Entwicklungen in den letzten 40 Jahren skizziert. In der Artikelsammlung Künstliche Intelligenz zieht in den Krieg (Kreowski & Lye 2021b) sind diverse Einzelaspekte diskutiert, die wir hier nicht ausführen konnten. Mit Hilfe von KI ist ein breites Arsenal an neuen und weiterentwickelten Waffensystemen sowie an militärischen Informations- und Kommunikationssystemen für Aufklärung, Kommando, Steuerung und Management hervorgebracht worden. Wie Wirtschaft und Politik sich weltweit von KI die Sicherung zukünftiger Wertschöpfung versprechen, so erwägen Militärs auf der ganzen Welt den Einsatz fortgeschrittener KI in militärischen Systemen und haben begonnen, fortgeschrittenere KI-Fähigkeiten zu entwickeln. Mehrere Länder haben militärische KI-Strategien veröffentlicht, in denen sie darlegen, wie KI in ihren militärischen Systemen genutzt werden soll, und haben darüber hinaus militärische KI-Forschungszentren eingerichtet (vgl. Gray & Ertan 2021). Es ist manchmal die Rede davon, dass der Einsatz von KI im Krieg dazu zwingt, Krieg neu zu denken. Tatsächlich aber werden neue Kriegssysteme mit oder oder KI entwickelt und eingesetzt, sobald sich die Militärs Vorteile und Überlegenheit versprechen. So bedeutet Drohnenkrieg nicht, dass nun Krieg nur noch mit Drohnen geführt wird, sondern dass die Drohnen zu all den Kriegsgeräten, die sonst schon Tod und Zerstörung bringen, als neues Element dazukommen und allenfalls mittelfristig andere Waffensysteme ersetzen. Es ist noch nicht abzuschätzen, was die jüngsten aufsehenerregenden Entwicklungen bei generativen Sprachmodellen und Systemen wie ChatGPT für militärische Zwecke bringen werden. Googles DeepMind, welches in einigen Strategiespie-

len, wie Schach, Go, Starcraft II menschliche Expert:innen besiegte, löste auch ein 50 Jahre offenes Problem der Biologie (vgl. Jumper 2020). Es werden bereits Befürchtungen laut, dass beispielsweise neuartige Bio- und Chemiewaffen möglich werden. Auf jeden Fall bleibt die militärische Seite der KI ein Thema, das man nicht aus den Augen verlieren darf.

Der Beitrag erschien zuerst in der Zeitschrift vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik, Nummer 242 (2/2023), Dezember 2023. Wir danken Autoren und Redaktion für die freundliche Genehmigung zum Wiederabdruck.

Referenzen

Adams, Eric 2018: Meet The UK's New, Very British Fighter Jet, Wired, 6 August 2018

Airbus 2018: Airbus demonstrates manned-unmanned teaming for future air combat systems. Website. 02.10.2018

Airbus 2020a: Manned-Unmanned Teaming and Remote Carriers: transcending individual assets' capabilities. Website. 08.10.2020

Airbus 2020b: Airbus and Thales join forces to develop the Air Combat Cloud for Future Combat Air System. Website. 19.02.2020.

Amt für Heeresentwicklung 2019: Künstliche Intelligenz in den Landstreitkräften: Ein Positionspapier des Amts für Heeresentwicklung

DARPA 2023: ACE Program's AI Agents Transition from Simulation to Live Flight: DARPA completes first flight tests of air combat algorithms on specialized F-16 fighter jet DARPA. 13 Feb. 2023

Department of Defense 2018: Unmanned Systems Integrated Roadmap. FY2017-2042

Gray, Maggie; Ertan, Amy 2021: Artificial Intelligence and Autonomy in the Military: An Overview of NATO Member States' Strategies and Deployment. NATO CCDCEO.

Jumper, John et al. 2020: High Accuracy Protein Structure Prediction Using Deep Learning. In Fourteenth Critical Assessment of Techniques for Protein Structure Prediction

Knight, Will 2020: A Dogfight Renews Concerns About AI's Lethal Potential, Wired, Conde Nast, 25 August 2020.

Kreowski, Hans-Jörg 2022: Drohnenkrieg in der Ukraine: Fakten und erste Folgenabschätzung. Wissenschaft und Frieden 2022/3

Kreowski, Hans-Jörg; Lye, Aaron 2021a: Future Combat Air System: Künstliche Intelligenz fliegt und kämpft mit. Wissenschaft und Frieden Dossier 93

Kreowski, Hans-Jörg; Lye, Aaron 2021b: Künstliche Intelligenz zieht in den Krieg. FIfF-Kommunikation 4/2021 und Wissenschaft und Frieden Dossier 93

Lye, Aaron 2021: Transformation geheimdienstlicher und militärischer Serverinfrastruktur in den USA durch kommerzielle Cloud-Provider am Beispiel Amazon. In. FIfF-Kommunikation 4/2021

Marischka, Christoph 2022: Drohnen im Ukraine-Krieg: Technologietransfer als Gamechanger – und Kriegsgrund? IMI-Studie 2022/03

Marischka, Christoph 2023: Noch schreibt KI für uns Texte – bald wird sie Kriege führen. In: Telepolis. 29. Mai 2023

McCarthy, J. et al. 1955: A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. August 31, 1955, http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

Odagiri, Hiroyuki et al. 1997: Research consortia as a vehicle for basic research: The case of a fifth generation computer project in Japan. Research Policy. 26 (2): 191–207. doi:10.1016/S0048-7333(97)00008-5

Osborn, Kris 2017a: ‚Air Force Chief Scientist Confirms F-35 Will Include Artificial Intelligence', Defense Systems, 20. Januar 2017

Osborn, Kris 2017b: ‚The F-35 Will Soon Be Equipped with Artificial Intelligence to Control Drone Wingmen', Business Insider, 20. Januar 2017.

Pletsch, Marius 2020: Mensch-Maschine: EU Großprojekte zum Manned-Unmanned-Teaming. IMI-Analyse 2020/11, in: IMI-AUSDRUCK (März 2020). 16. März 2020

Renn, U. 2021: Unbemannte Helfer: Zur Bedeutung der Remote Carrier im Future Combat Air System, in: Europäische Sicherheit und Technik, 4-2021, S. 38-42.

Roland, Alex; Shiman, Phlip 2002: Strategic Computing – DARPA and the Quest for Machine Intelligence, 1983-1993. The MIT Press 2002.

Schelling, Arkadi 2019: Künstliche Intelligenz als Cloud Service: Folgen für Gesellschaft, Geheimdienst und Militär. IMI-Analyse 16/2019

Shapiro, Ehud Y. 1983: The fifth generation project – a trip report. Communications of the ACM. 26 (9): 637–641. doi:10.1145/358172.358179. S2CID 5955109.

Trakimavičius, Lukas 2021: The Future Role of Nuclear Propulsion in the Military. NATO Energy Security Centre of Excellence. 10/2021

Williams, Andrew P.; Scharre, Paul D. (Eds.) 2015: Autonomous Systems – Issues for Defence Policymakers. NATO Headquarters Supreme Allied Commander, Transformation, Norfolk, Virginia, United States, https://apps.dtic.mil/sti/citations/AD1010077

Hans-Jörg Kreowski und Aaron Lye

Hans-Jörg Kreowski (Jahrgang 1949) ist Professor (i. R.) für Theoretische Informatik an der Universität Bremen. Er ist Mitglied im Vorstand des Forums InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FIfF) und vertritt das FIfF im Vorstand der Zeitschrift Wissenschaft und Frieden. Er ist Mitglied der Leibniz-Sozietät der Wissenschaften zu Berlin, wo er zusammen mit Wolfgang Hofkirchner in Wien den Arbeitskreis Emergente Systeme, Information und Gesellschaft organisiert. Von 2019 bis 2022 war er außerdem Mitherausgeber des Grundrechte-Reports.

Aaron Lye hat an der Universität Bremen Informatik studiert und dort auch Ende 2021 seine Promotion abgeschlossen. Seit etwas zehn Jahren engagiert er sich beim FIfF mit den Schwerpunkten Überwachung und informationstechnische Kriegsführung.

Weiterentwicklung einer universellen Umweltsensorstation

2. November 2023 – *FIfF und TDRM starten eine programmatische Initiative zu Citizen Sensing und Citizen Science.*

Im Jahr 2022 entwickelten Mitglieder des FIfF e. V.[1] und des TDRM-Projektes[2] in einem von der Deutschen Stiftung für Engagement und Ehrenamt (DSEE)[3] geförderten Projekt die Open-Source-Hardware des Prototypen einer Umweltsensorstation[4]. Einfacher Geräteaufbau, niedrige Kosten und flexible Sensorauswahl zeichnen die Station aus. Lizenzfreie Kommunikation per LoRaWAN[5] und TTN-Netzwerk[6], große Reichweite und Versorgung durch Solarmodul und Akku ermöglichen eine freizügige Aufstellung. Die Software der Station wurde innerhalb des ersten Projektes prototypisch bis zur Funkübertragung implementiert.

Nun erhalten das FIfF und TDRM in einem Folgeprojekt, wiederum gefördert durch die DSEE bis Ende 2024, die Chance, den Grundstein für eine programmatische Initiative zu legen: Citizen Sensing – Bürger:innen beteiligen sich aktiv am Monitoring unserer Umwelt und übernehmen damit eine gesellschaftliche Verantwortung. Sie setzen sich mit konstruktiven Anwendungen der Informationstechnik auseinander und demonstrieren deren gesellschaftlichen Nutzen. Die Datenerhebung dient als Grundlage zur öffentlichen wissenschaftlichen Datenauswertung und Verknüpfung mit aktuellen gesellschaftsrelevanten Fragen, zu Citizen Science.

Im Folgeprojekt soll nun eine anwendungsreife Software entwickelt und die Hardware überarbeitet werden, außerdem ein Feldtest durchgeführt sowie die öffentliche Dokumentation dieser neuen Soft- und Hardware erarbeitet werden. Für die Entwicklung der Stationssoftware wird das embedded RTOS Zephyr[7] eingesetzt. Zur serverseitigen Anbindung wird ebenfalls Open-Source-Software genutzt.

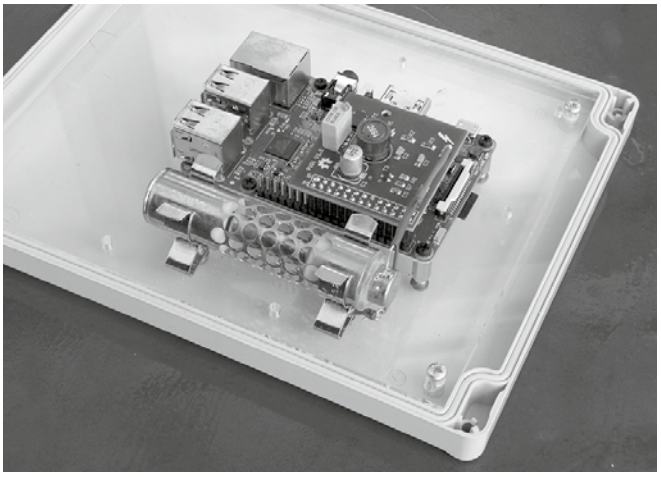
Die neue Station soll in seiner zukünftigen Version u. a. atmosphärische Radioaktivität, Lärm, zivile Überflüge (ADS-B), Luftqualität, Feinstaub, Temperatur, Luftfeuchte, Regen usw. erfassen können, so dass diese im Internet öffentlich dargestellt werden können.

Mit unserem Vorhaben möchten wir zivilgesellschaftliche Aktivitäten stärken. Eine Beteiligung von Bürger:innen an der Beobachtung, Erfassung und Auswertung der Parameter unserer Lebenswelt fördert Umweltbewusstsein und -verantwortung, vermittelt Wissen über Umwelt und Technik und ebnet den Weg zu einem informierten politischen Engagement. Vielfach sind solche Initiativen bereits behördlichen Maßnahmen zuvor gekommen und haben politisches Handeln beschleunigt oder sogar erst veranlasst.

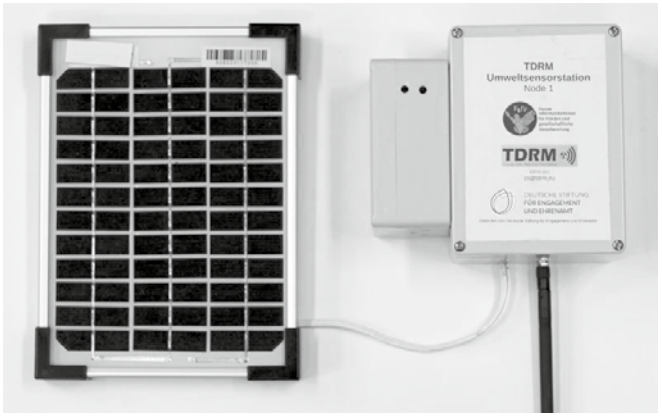
TDRM besteht seit der offiziellen Inbetriebnahme des TDRM-Netzes im Rahmen einer Pressekonferenz am 16. Dezember 2016. TDRM ist grenzüberschreitend im niederländisch-belgisch-deutschen Raum aktiv und die einzig verbliebene Gruppeninitiative zu den belgischen AKWs Thiange und Doel.

Anmerkungen

1 *https://fiff.de/*
2 *http://tdrm.eu/*
3 *https://www.deutsche-stiftung-engagement-und-ehrenamt.de/*
4 *https://github.com/PeterKamm/Environmental-Citizen-Sensor-Station-FIfF-DSEE/*
5 *https://de.wikipedia.org/wiki/Long_Range_Wide_Area_Network/*
6 *https://www.thethingsnetwork.org/*
7 *https://www.zephyrproject.org/*



TDRM-Station von 2020. Foto: TDRM Dietrich Meyer-Ebrecht



Prototyp der TDRM-Sensor-Station von 2023. Foto: TDRM

TDRM
Thiange Doel Radiation Monitoring

Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung

DEUTSCHE STIFTUNG FÜR ENGAGEMENT UND EHRENAMT

Offener Brief

Joint statement of scientists and NGOs on the EU's proposed eIDAS reform

2nd November 2023

Dear Members of the European Parliament,
Dear Member States of the Council of the European Union,

We the undersigned are cybersecurity experts, researchers, and civil society organisations from across the globe.

We have read the near-final text of the eIDAS digital identity reform which has been agreed on a technical level in the trilogue between representatives from the European Parliament, Council and Commission. We appreciate your efforts to improve the digital security of European citizens; it is of utmost importance that the digital interactions of citizens with government institutions and industry can be secure while protecting citizens' privacy. Indeed, having common technical standards and enabling secure cross-border electronic identity solutions is a solid step in this direction. However, we are extremely concerned that, as proposed in its current form, this legislation will not result in adequate technological safeguards for citizens and businesses, as intended. In fact, it will very likely result in less security for all.

Last year, many of us wrote to you to highlight some of the dangers in the European Commission's proposed eIDAS regulation. After reading the near-final text, we are deeply concerned by the proposed text for Article 45. The current proposal radically expands the ability of governments to surveil both their own citizens and residents across the EU by providing them with the technical means to intercept encrypted web traffic, as well as undermining the existing oversight mechanisms relied on by European citizens. Concretely, the regulation enables each EU member state (and recognised third party countries) to designate cryptographic keys for which trust is mandatory; this trust can only be withdrawn with the government's permission (Article 45a(4)). This means any EU member state or third party country, acting alone, is capable of intercepting the web traffic of any EU citizen and there is no effective recourse. We ask that you urgently reconsider this text and make clear that Article 45 will not interfere with trust decisions around the cryptographic keys and certificates used to secure web traffic.

Article 45 also bans security checks on EU web certificates unless expressly permitted by regulation when establishing encrypted web traffic connections (Article 45(2a)). Instead of specifying a set of minimum security measures which must be enforced as a baseline, it effectively specifies an upper bound on the security measures which cannot be improved upon without the permission of ETSI. This runs counter to well established global norms where new cybersecurity technologies are developed and deployed in response to fast moving developments in technology. This effectively limits the security measures that can be taken to protect the European web. We ask that you reverse this clause, not limiting but encouraging the development of new security measures in response to fast-evolving threats.

The current text also mentions in multiple places the need for the European Digital Identity Wallet to protect privacy, including data minimization, and prevention of profiling. Yet, the legislation still allows relying parties like governments and service providers to unnecessarily link together and gain full knowledge about the uses of credentials in the new European Digital Identity System. Given the broad intended uses of this system, which span all areas of life from health, finance, commerce, online activity up to public transport, we believe that failing to require both unlinkability and unobservability will severely compromise the privacy of EU citizens. Article 6a(7)(a) should be aligned with the negotiation mandate from the European Parliament lead Industry Committee and thereby prevent technologically that such information can be obtained by governments and other parties without the explicit consent of users. Article 6a(7a)(b) should "mandate" instead of "enable" that interactions cannot be linked by relying parties or other actors, where identification of the user is not mandatory. Lastly, forum-shopping from 'Big Tech' and other bad actors can only be prevented by a harmonised implementation of the Regulation that allows national eIDAS agencies to be overruled should they fail to act.

Finally, we would like to highlight our frustration that decisions crucial for the security and privacy of citizens, businesses, and governments, are being taken behind closed doors in trilogue negotiations without public consultation of experts about the potential consequences of the proposed regulations. We urge the European Parliament, Commission, and Council to reconsider their legislative processes and commit to greater transparency so that experts and the public can effectively contribute to the development of new regulations.[1]

In summary, we strongly warn against the currently proposed trilogue agreement, as it fails to properly respect the right to privacy of citizens and secure online communications; without establishing proper safeguards as outlined above, it instead substantially increases the potential for harm.

1. Undermining website authentication undermines communications security.

The current text of Article 45 mandates that browsers must accept any root certificates provided by any Member State (and any third party country approved by the EU) and will have severe consequences for the privacy of European citizens, the security of European commerce, and the Internet as a whole.

Root certificates, controlled by so-called certificate authorities, provide the authentication mechanisms for websites by assuring the user that the cryptographic keys used to authenticate the website content belong to that website. The owner of a root certificate can intercept users' web traffic by replacing the website's cryptographic keys with substitutes he controls. Such a substitution can occur even *if the website has chosen to use a different certificate authority with a different root certificate.* **Any root cer-**

tificate trusted by the browser can be used to compromise any website. There are multiple documented cases of abuse, because the security of some certificate authorities has been compromised. To avoid this, there exists legislation that regulates certificate authorities, complemented by public processes and continuous vigilance by the security community to reveal suspicious activities.

The proposed eIDAS revision gives Member States the possibility of inserting root certificates at will, with the aim to improve the digital security of European citizens by giving them new ways to obtain authentic information of who operates a website. In practice, this does exactly the opposite. Consider the situation in which one of the Member States (or any of the third party states recognized now or in the future) were to add a new authority to the EU Trusted List. The certificate would have to be immediately added to all browsers and distributed to all of their users across the EU as a trusted certificate. By using the substitution techniques explained above, the government-controlled authority would then be able to **intercept the web traffic of not only their own citizens, but all EU citizens**, including banking information, legally privileged information, medical records and family photos. This would be true even when visiting non-EU websites, as such an authority could issue certificates for any website that all browsers would have to accept. Additionally, although much of eIDAS2.0 regulation carefully gives citizens the capability to opt out from usage of new services and functionality, this is not the case for Article 45. **Every citizen would have to trust those certificates**, and thus every citizen would see their online safety threatened.

Even if this misbehaviour was discovered, under the current proposal it would not be possible to remove this certificate without the ultimate approval of the country having introduced the certificate authority. Neither eIDAS's article 45 nor any provisions in adjacent EU legislation such as the NIS2 Directive provide any independent checks and balances on these decisions. Further, European citizens do not have an effective way to appeal these decisions. This situation would be unacceptably damaging to online trust and safety in Europe and across the world. We believe this legislative text must be urgently reworked to avoid these serious consequences by clarifying that eIDAS does not impose obligations to trust cryptographic keys used for encrypted web traffic.

The proposed legislation also prevents the introduction of security checks when verifying the certificates used for encrypted web traffic in Article 45 (2a). As written, this language requires that the EU's website certificates not be subjected to any mandatory requirements beyond those specified in ETSI standards. Mandatory requirements on certificates are essential when browsers validate certificates presented for use in encrypted web connections. Preventing these additional security checks has no useful purpose and only hampers the improvement of cybersecurity for European citizens. The detailed rules on certificate validation and display are constantly being adapted based on new research results and consensus in the security community. Existing security mechanisms, well-studied and accepted by the security community at large, such as TLS 1.3 and certificate transparency logs currently enable browsers to quickly adapt to changing threats and improve global web security. It is essential that this regulation establishes a mandatory minimum set of security standards, but does not impose a limited set of requirements which would hamper the adoption of new security technology within the EU.

While Article 45 could be understood as reducing the power of the large companies behind the major web browsers, from a technical perspective, this is not the case. There already exists a large number of certificate authorities capable of issuing certificates trusted in every web browser, many of which are European and also recognised under the EU's existing eIDAS legislation. Websites have a free choice about which certificate authority they use and all of the approved certificate authorities are treated equally in the browser. Should issues arise, the EU is already well-equipped to tackle them through the recently passed Digital Markets Act, which specifically identifies popular browsers and cloud services and bans self-preferencing behaviour by gatekeepers.

Article 45 itself does nothing to assist this process or to enable European scrutiny of trust decisions by 'Big Tech', instead it only enables the interception of EU citizens' web traffic by European governments. It further prevents concerned users, who may have serious and substantiated concerns about being subject to state surveillance, from choosing, or even creating, a browser that has stricter security checks.

In summary, this regulation allows misbehaviour by any individual Member State (or approved third party countries) to compromise the safety and security of other Member State's citizens. If it is implemented, it would result in citizens having to, **without a choice**, trust **all** certificate authorities defined by Member States (and recognized third countries) **in addition** to the parties they trust today. This regulation does not eliminate any existing risk. Instead, by undermining the existing secure web authentication processes, introduces new risks with no gain by European citizens, businesses, and institutions. Moreover, if this regulation becomes a reality, it is only to be expected that **other countries will put pressure on browsers to obtain similar privileges** as EU Member States – as some have unsuccessfully attempted in the past – globally endangering web security.

In order to address these concerns and avoid the security issues introduced by the current legislation proposal which could result in incalculable damage, we recommend:

- The text be clarified to ensure that this legislation will not interfere with trust decisions around the cryptographic keys or certificates used to secure web traffic and the consequent impact on privacy and security of European citizens.

- Additional checks independent from those envisioned in the legislation are not only permitted but encouraged to enable browsers to rapidly incorporate advances made by the security community to improve the security of communications.

In particular:

- The re-introduction of text to Article 45 (2) limiting its scope: "Such recognition, support and interoperability means solely that web-browsers shall ensure that the identity data attested in the certificate provided using any of the methods is displayed in a user-friendly manner."

- The deletion of Article 45 (2a) so that new security checks can be implemented effectively

- In Recital 32: Adding clarification that the obligations of recognition, interoperability and support in Article 45 do not extend to the use of encryption and authentication technologies for securing web traffic.

We also explicitly note that established processes clearly allow new certificate authorities to be added to browser root trust stores; nation states wishing to establish a new CA legitimate and lawful purposes need to go through the same security certification procedures that existing authorities do, without requiring new regulation. Fostering the development of an EU-native browser, or strengthening the supervision of certificate authorities across the EU, would have a much more positive impact on the overall security of European citizens than attempting to change the status quo of web security from within the eIDAS regulation.

2. A complex system only provides the security and privacy guarantees of its weakest component

The European Digital Identity Wallet (EDIW) is designed to identify and authenticate users with a high level of assurance. The Wallet includes identity information from national IDs (age, sex, etc), and can be extended with additional attributes. These attributes could include very sensitive information such as medical certificates, or important information for the future of European citizens such as their professional qualifications. The eIDAS regulation foresees the creation of an ecosystem of public and private entities that will benefit from the Wallet to have access to certified personal information about citizens.

We welcome the provisions crafted in the legislation, which advocate for strong protections to preclude tracking and profiling, that enable the option of revealing attributes in a selective manner or via zero-knowledge attestation, that attribute providers should not learn about with whom users share their attributes, or that mention that the wallet should allow for unlinkability when identification is not needed. These are essential to promote the use of technologies that can provide these properties by design, and we commend the legislators for including them.

Yet, the legislation only enables the existence of privacy-preserving technologies, but does not mandate them (Article 6a(7a)(b)). We are concerned that this legal ambiguity could lead to adeterioration of privacy-safeguards that ultimately leaves too much room for technical implementation on member state level. Importantly, operators of the EDIW can still obtain knowledge about concrete user behaviour even when the user has not consented to this. With a privacy-respecting architecture such information is not necessary for the provision of the EDIW. With the current legal text the architecture of the whole system risks undermining trust from citizens in the whole system (Article 6a(7)). A fully harmonised European system for the benefit of the private sector also needs a fully harmonised level of safeguards European citizens can rely upon. Moreover, relying parties (service providers with access to the wallet) can also register in any of the Member States, thus the effective regulatory regime that bad actors and 'Big Tech' can exploit is the weakest of all Member States as we have seen with the GDPR and DSA. This is particularly challenging because of the necessity of cross-border interoperability. Hence, we recommend in Article 46e to empower the European Digital Identity Cooperation Group to overrule the decisions of national eIDAS regulators in order to prevent the circumvention of these important protections.

In order to address these concerns and avoid that the eIDAS regulation results in a new privacy problem with no security gain in terms of authentication, we recommend:

- Make unlinkability a mandatory rather than optional requirement by Replacing "enable" with "mandate" in Article 6a(7a)(b).

- Align the technical architecture with the strong protections established in the lead Industry committee of the European Parliament in Article 6a(7).

- Provide a majority in the European Digital Identity Cooperation Group according to Article 46e the power to overrule the decision of national eIDAS regulators in order to ensure a harmonised enforcement of this regulation.

Without these necessary amendments the eIDAS regulation risks becoming a gift to Google and other Big Tech actors. A European solution to the central question of handling sensitive identity information needs to protect citizens against surveillance capitalism through strong technical mechanisms and be resilient against attempts to exploit the regulatory system through jurisdiction-shopping.

Anmerkungen

1 T-540/15 – De Capitani v Parliament

As of the 26th November 2023, the letter has been signed by 551 scientists and researchers from 42 countries, as well as numerous NGOs: *AG Nachhaltige Digitalisierung, Associação de Empresas de Software Open Source Portuguesas (ESOP), Associação Portuguesa para a Promoção da Segurança da Informação (AP2SI), Asociatia pentru Tehnologie si Internet (ApTI), Center for Democracy and Technology, Chaos Computer Club, Council of European Professional Informatics Societies, D64 – Zentrum für Digitalen Fortschritt e. V., Defesa dos Direitos Digitais (D3), deSEC, Digitalcourage, Digitale Gesellschaft, Electronic Frontier Finland (Effi), Electronic Frontier Foundation (EFF), Emerald Onion, Entropia e.V., Epicenter.works, European Digital Rights (EDRi), Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FIfF) e.V., Foundation for Information Policy Research (FIPR), Gesellschaft für Informatik e.V., Homo Digitalis, IETF Internet Architecture Board (IAB), Innovationsverbund Öffentliche Gesundheit e.V., Internet Governance Project, Internet Architecture Board, Internet Society, Internet Society Catalan Chapter, Internet Society Switzerland Chapter, Internet Society UK Chapter, IT-Pol, LOAD e.V., La Quadrature du Net, Özgür Yazılım Derneği, Petites Singularités, Privacy & Access Council of Canada, Privacy First, Rhizomatica, SICEH Foundation, SUPERRR Lab, The Digital Freedom and Rights Association, The Document Foundation, The Matrix.org Foundation, The Syncthing Foundation, The Tor Project, Trust in Digital Life*

Dagmar Boedicker

Konferenz über Nachhaltigkeit und Digitalisierung

an der Universität Augsburg

Das war eine besondere Tagung: interdisziplinär, international und dabei fast ein Familientreffen. Denn so groß ist die Gemeinschaft der teils noch recht jungen Forschenden nicht, die sich in Augsburg zwischen IKT und Ökologie bewegen. In der Schlussrunde am dritten Tag äußern sie sich zufrieden mit dem Austausch, aus dem eine neue Gemeinschaft entstanden sei. Die Politikwissenschaftlerin Angela Oels und ihr Team hatten ein vielfältiges und spannendes Programm zusammengestellt, in dem es auch Raum für Vernetzung gab. Dieser Raum ist nötig und willkommen in einer Forschungslandschaft, in der viele recht isoliert an diesen Zukunftsthemen arbeiten. Die Interdisziplinarität war ein Erfolg und die Herausforderung der anderen Disziplinen „gut für den Kopf". Es wird hoffentlich eine Folgekonferenz geben. Hier zunächst ein Bericht mit Schlaglichtern von der Veranstaltung, die ich auch als eine Vertiefung von Aspekten der Tagung Bits & Bäume erlebt habe. Leider konnte ich das reichhaltige Angebot an Vorträgen und Diskussionen nur in Teilen wahrnehmen.

Forschungsfragen

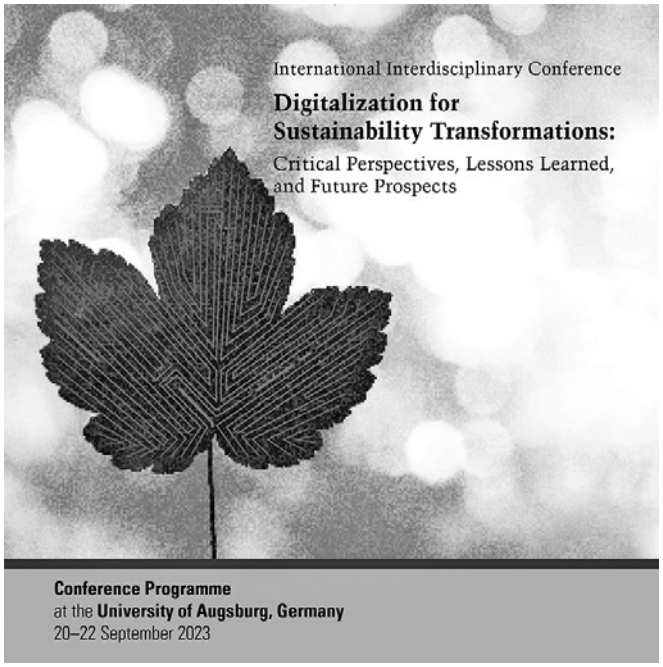
Die beiden übergreifenden Fragen, die die Konfererenz an die Forschenden gestellt hat:

- Wie verändert die Digitalisierung Erkenntnisse und Praktiken der Umwelt- und Nachhaltigkeits-Politik?
- Wie sollte der Rahmen für ihre Regulierung und Handhabung aussehen, um ihr Potenzial in den Dienst der Nachhaltigkeits-Transformation zu stellen?

Einige der Details aus den Forschungsvorhaben: Welche Vorstellung herrscht in dem einen der beiden Bereiche über den jeweils anderen? Mit welchem Verständnis gehen die Akteure (hoch-) komplexe Probleme der Nachhaltigkeit an? Wie lässt sich der Zusammenhang zwischen Digitalisierung und Nachhaltigkeit in konkrete Politik umsetzen? Wer sind die Akteure und wie hat sich ihr Diskurs im Lauf der Zeit gewandelt? Gibt es immer noch koloniale Muster? Welche technischen Lösungen für Nachhaltigkeits-Probleme werden angeboten oder eben nicht und warum? Wie sollten sie gestaltet sein? Welche Entscheidungsprozesse bestimmen in welchen Machtverhältnissen und Institutionen den Zusammenhang von Digitalisierung und Nachhaltigkeit? Wie unterscheiden sie sich in unterschiedlichen Regionen und welche Chancen bieten sie für Beteiligung? Wo befördern sie dagegen ausbeuterische Mechanismen? Welche digitalen Techniken unterstützen den sozial-ökologischen Transformations-Prozess, wo sind ihre Grenzen und wie hilfreich kann KI sein?[1]

Subversive Daten und Partizipation

Daten zur nachhaltigen Transformation können subversiv sein, in manchen Regionen sind sie gefährlich. Das wird deutlich bei *Radhika Krishnans* Vortrag mit dem Titel *Planning a just Transition from coal in India: Digitisation of data and the possibility of subversions in mineral resource governance*. In ihrem Forschungsprojekt untersucht sie die Möglichkeiten für einen gerechten und partizipativen graduellen Ausstieg aus dem indischen Kohlebergbau mit dem Ziel, einen übertragbaren Rahmen für solche Transformationen zu entwerfen. Seit die Zivilgesellschaft in Indien aktiver wird, lassen sich aus Projekten mit (auch



Laien-) Wissenschaftlerïnnen und gestützt auf Satelliten-Aufnahmen verlässlichere Daten zum Grundflächenverbrauch und den Umweltschäden des Kohlebergbaus gewinnen. Krishnans Beispiel: In der Kohleregion *Korba* leben 45 % der ursprünglichen Bevölkerung von Waldwirtschaft, die Kohle bestimmt aber das BIP und die Arbeitsplätze. Konflikte um die Landnutzung sind programmiert. Die Grenzen zwischen Wirtschaft und Regierung sind porös, staatliche Institutionen strukturell schwach und der Willkür ausgesetzt.

Was ist, wenn dann eine Mine schließt und das Land zurückgegeben werden muss? Ein Stilllegungs- und Schließungsplan ist zwar gesetzlich vorgeschrieben, unzuverlässige Daten führen aber – unbeabsichtigt oder nicht – oft zu nicht nachhaltiger Wiedernutzbarmachung und Bewirtschaftung. Wie schon bei der Erschließung können auch Maßnahmen im Rahmen der Stilllegung wie eine Aufforstung dazu führen, dass Menschen vertrieben werden. Deshalb sammeln die Forschenden Schließungspläne *(best practices)* aus anderen Ländern. Die Daten aus Radhika Krishnans Projekt sind hilfreich, aber heikel, es ist schwierig, sie in die richtigen Hände zu geben. Sie müssen in

die lokalen Sprachen übersetzt, visualisiert und verständlich gemacht werden, eine aufwändige Eins-zu-eins-Vermittlung, die über eine sichere Website im Netz erfolgen soll. Damit können die Beteiligten Ärger bekommen, es ist also Vorsicht geboten.

Auch *Joshua Zeunerts* Vortrag weist auf das Konfliktpotenzial hin. In seinem Projekt *Foodlandscapes*[2] zeigt er Videos aus der Vogelperspektive, visualisiert aus 881 *Geolocation*-Einträgen typischer Produktions-Landschaften, die er in Australien untersucht und gefilmt hat. Foodlandscapes wirkt in spielerischer Art und Weise aufklärerisch und als kritischer Kommentar darüber, was Landwirtschaft und Landschaft verbindet. Im Vortrag „What the f#$d? Landscape digitisation linking food choices and production" zeigt Zeunert die riesigen Flächen, auf denen in Australien Nahrungsmittel produziert werden, vorwiegend Viehzucht, die mindestens so weit vom Bauernhof-Idyll entfernt ist wie in Europa. Er ist dafür 35 000 km gefahren. Gefragt, ob das Bildungsziel, mehr über den Zusammenhang zwischen unseren Ernährungs-Präferenzen und ihren negativen Folgen zu erfahren, nicht durch einige Aufnahmen aus Schlachthöfen hätte veranschaulicht werden können, weist Zeunert darauf hin, dass die vorherige australische Regierung Tierschützer sanktionierte, die sich exponiert hatten. Auch wenn die jetzige Regierung das nicht tue, rufe Kritik an der industriellen Landwirtschaft sehr wohl Aggressionen hervor.

Nachhaltigkeit – ein verbrauchter Begriff?

Diesen Begriff – als *Sustainability* Schlüsselwort des Tagungstitels – führen die Anwesenden gar nicht oft im Munde. Nicht nur *Roy Bendor* stellt fest, dass der Begriff der nachhaltigen Entwicklung heute seltener benutzt wird. Möglicherweise hat er sich abgenutzt. Es könnte aber auch daran liegen, dass die Welt noch sehr weit davon entfernt ist, die 17 Nachhaltigkeitsziele (*Sustainable Development Goals*, SDGs) und die dazugehörigen 169 Unterziele zu erreichen. Bendor meint, dass der Begriff der Resilienz den der Nachhaltigkeit abgelöst habe. Auch *Ruth Machen,* sie lehrt Stadtplanung an der Newcastle University (UK), befasst sich in ihrem Vortrag *Algorithmic climate governance: Reproducing hegemony or radical transformation?* mit Umwelt-Regimen am Übergang zwischen Klima-Wissenschaft und Politik. Sie beleuchtet den Umgang mit Nachhaltigkeitswissen und welche Werte und Perspektiven Eingang in das politische Handeln finden. Machen fragt, wie digitale Techniken Prozesse prägen und die Ergebnisse von Entscheidungen in der Klima-Politik beeinflussen. Vorhandenes Wissen mit den betreffenden Daten bestimmt, in welchen Zusammenhang die Probleme von den Entscheiderïnnen gestellt werden und welche Algorithmen bei den Lösungsansätzen greifen sollen. Die algorithmische Herangehensweise steht für den status quo: weiß, kapitalistisch, militärisch. Das Spezielle wird zum allgemein Gültigen. Algorithmische Entscheidungen fallen in einer Hierarchie, in der die Beziehung zu den Kundïnnen oder Verbraucherïnnen Vorrang hat – solche Regime sind schwerlich nachhaltig. Machen arbeitet mit *Laclaus* und *Mouffes* Definition von Hegemonie: Hegemonie ist kulturell und ideologisch, die Dominanz entsteht durch Zustimmung, nicht durch Gewalt. Grundlage der Organisation ist eine Vielfalt von Institutionen, wie es auch die westlichen Tech-Monopole sind. Sie sollen dominieren und kontrollieren und beeinflussen

Radhika Krishnan, Ph.D.

dazu den sozialen, politischen und kulturellen Diskurs. Algorithmen bilden Sedimente, wie jede Infrastruktur verhärten sie, und da ist zu fragen, welche Idee vom *gesunden Menschenverstand* sie auf dem Gebiet der Nachhaltigkeit erzeugen. Machen ermutigt mit diesem Zitat von Chantal Mouffe: „Hegemony is a claim that […] can always be contested." Einerseits kann KI Verantwortung fragmentieren, bestimmte Optionen ausschließen und koloniale Muster und andere Verzerrungen reproduzieren. Energie- und Ressourcenverbrauch sind problematisch. Sie kann aber auch Perspektiven eröffnen. Wir sollten die Technik nicht ablehnen, sondern nach ihren Möglichkeiten suchen, und dabei nicht vergessen, dass nicht das Modell agonistische Diskussionen auslöst, sondern die Beziehung zum Modell und zu seinen Ergebnissen. Dieser Aufgabe müssen sich all die Forscherïnnen stellen, die für ihre Forschung brennen, sie müssen sich fragen: Ist das Ziel wirklich Nachhaltigkeit?

Auch die Diskussion zu Machens Vortrag ist spannend. So wird sie gefragt, ob Algorithmen beispielsweise bei Verteilungskonflikten helfen können oder eher nicht. Ihre Antwort: Algorithmen können diese Agonismen aufzeigen, dabei ist nicht sicher, dass dadurch ungehörte Stimmen hörbar werden. Davon müsse sie erst noch überzeugt werden. Als Politik-Beraterin hört sie oft, dass fehlende Information ein Problem sei, es liege aber eher am Willen zur Umsetzung, Datensammeln diene der Auf-

Roy Bendors Frage nach der Nachhaltigkeit

Atmosphäre im Saal …

schieberitis. Wenn es um die Prozesse des Schärfens mit Modellierung und Umsetzung geht, bestimmen darüber die Entscheiderïnnen. Dort müssen wir die Prioritäten prüfen und kontrollieren.[3]

Wer über Politiken entscheidet, muss alle Dimensionen von Nachhaltigkeit berücksichtigen, sagt *Daniel Wurm* vom Wuppertal Institut für Klima, Umwelt, Energie gGmbH. Nachhaltigkeit hat mehrere Dimensionen: ökologisch, politisch, technisch, sozial/kulturell und ökonomisch. Die Europäische Union will Digitalisierung und Nachhaltigkeit parallel (als Zwillinge) verfolgen. In seinem Vortrag *Selecting transformative policies in the twin Transition* fragt Wurm, wie eine Gesellschaft die richtige Vorgehensweise *(policy)* wählt. Woher kommt die Orientierung, damit die EU nicht in die falsche Richtung läuft? Wurm zeichnet das Rahmenwerk für seine Untersuchung mit bisher zwei Fallstudien und 20 Experten-Interviews, zu denen weitere 20 bis 40 kommen sollen: Probleme in einer gesellschaftlichen Aufgabe *(mission)* sind unterschiedlich und komplex, umstritten und unsicher. Es gibt mehrere Pfade zum Ziel, abhängig von den gesellschaftlichen Perspektiven, Interessen und Gewichtungen. Dabei ist nicht nur unklar, auf welche Lösungen sich die Gesellschaft in diesem Raum der Probleme und Lösungen einigen kann, möglicherweise muss sie die Pfade zum Ziel bei Digitalisierung sowie Nachhaltigkeit aufeinander abstimmen.

Wurms Beispiele machen es deutlich: In einer Fallstudie wird eine Lösung für die Digitalisierung gesucht mit dem Ziel der Kreislaufwirtschaft, in der anderen mit dem Ziel einer nachhaltigen Energieversorgung. Im Fokus der einen stehen die Ressourcen, Schwerpunkt der anderen ist die Energie; das eine Phänomen ist bekannter als das andere, beide stehen in unterschiedlichen Phasen der öffentlichen Debatten. Die interviewten Expertïnnen identifizierten die wesentlichen aktuellen Probleme und Lösungen in vier systemischen Dimensionen (technisch, ökonomisch, sozio-kulturell/verhaltensbezogen, regulatorisch) und es entstanden Diagramme für Probleme und Lösungen, die in den vier Dimensionen konsolidiert wurden. In zunächst zwei Workshops entstanden Bewertungen der Komplexität entlang der gesellschaftlichen Skalen von Zustimmungsgrad, Rollenverständnis und (Un)Gewissheit. Wie umstritten sind Probleme und ihre Wahrnehmung sowie mögliche Lösungen? Wie klar sind die Aufgaben bzw. Rollen? Wie viel wissenschaftliche Evidenz gibt es?

Bisher gibt es keine Blaupause für die parallele Transformation (*twin transition)* der EU. Wurm kann Hinweise dazu geben, welche Kriterien für die Auswahl der Transformations-Politik wesentlich sein dürften: Die sehr unterschiedlichen Ausgangslagen sind zu beachten und die Herangehensweise muss flexibel sein. In sehr vielen Fällen ist die öffentliche Debatte entscheidend, um mehr Wissen zu sammeln, neue Institutionen zu gründen oder zu stärken, in die Betroffene oder die Öffentlichkeit einbezogen werden können, um Lösungsansätze in verschiedenen Zusammenhängen auszuprobieren und zu kommunizieren. Wurms Fazit: Diese gesellschaftliche Herausforderung muss ein definiertes Ziel haben, und die *twin transition* braucht eine klare Orientierung, sonst läuft sie in die falsche Richtung.

Design für Nachhaltigkeit

Auch wenn viele unterschiedliche Auffassungen über die Bedeutung des Begriffs herrschen, sind sich die meisten Menschen wohl darüber einig, dass uns Nachhaltigkeit wichtig ist. Es ist ein sehr reichhaltiger, vieldeutiger Begriff, auf dieser Konferenz wird er in mehreren Dimensionen betrachtet, als soziale, politische, ökologische, technische und wirtschaftliche Nachhaltigkeit. Roy Bendor aus Delft bietet zwei Definitionen an: Nachhaltigkeit ist das Lösen komplexer sozial-ökologischer (socio-environmental) Probleme, und: Nachhaltigkeit ist die Beziehung zwischen ideellen und materiellen Praktiken. Design verbinde diese beiden Bereiche. Im Vortrag unter dem Titel *Sustainability imaginaries by design* wirft Bendor einen Blick auf Kommunikation für Nachhaltigkeit mit interaktiven, audiovisuellen Medien. Unter den Beispielen für Schubser *(Nudges),* durch die Menschen zu einem nachhaltigen Konsum veranlasst werden sollen, zeigt er den H_2O-Tracker *Track your water use and score,* die *Bottle Bank Arcade* für Flaschen als Wertstoff und andere Anwendungen der *fun theory*[4]*.* Bendors anregende Präsentation führt nicht die üblichen Designs von industriellen Produkten vor. Ihm geht es um neue Vorstellungen und Bildwelten[5], um ein *regeneratives Design* mit drei Säu-

… und während der Pause

len: einer ontologischen von Interdependenz und Vernetztheit, einer epistemologischen, die *das Wissen der anderen* sucht, und einer politischen mit Reziprozität und Achtsamkeit/Sorge. Ein partizipatives, kooperatives Design, dessen Evolution einzelne Designerïnnen variieren.

Bendors Vortrag ist der erste Hauptvortrag und stimmt bereits an, was sich durch die Konferenz ziehen wird: Kommunikation muss gestaltet werden. Wer im Bereich der Sorge arbeitet, weiß um die Vernetztheit unserer Beziehungen, die, die das nicht tun, können durch solches Design etwas lernen. Aber auch das Verlernen *(Unlearning)* ist ein Ziel, denn die westliche rationale Epistemologie und Kultur sind voller rücksichtsloser Muster. Mir kam allerdings die Frage zu kurz, wie weit die doch etwas paternalistischen Konzepte einer Technik tragen können, die menschliches Verhalten ändern und Nachhaltigkeit oder Empathie durch *gamification* bewirken möchte.

Stößt die Nachhaltigkeits-Transformation an eine gläserne Decke?

Mein Eindruck ist, dass jetzt, nach Jahrzehnten der Untätigkeit, die niedrig hängenden Früchte geerntet sind. Wir trennen ein bisschen Müll und sparen ein bisschen Energie und Papier. Wie kommen wir nun zu der Kreislaufwirtschaft, die Ressourcen und das Klima schont? Was kann die Digitalisierung dazu beitragen? *Lena Brüch* und ihre Kollegïnnen vom Fraunhofer FIT haben drei kleinen und mittleren Unternehmen diese Frage gestellt und 41 Antriebs- und 38 bremsende Kräfte identifiziert und in Kategorien eingestuft. Zwar betrachten die Firmen Kreislaufwirtschaft als eine große Chance, aber der Weg ist schwierig für KMUs, die angesichts des starken Konkurrenzdrucks wohl schon ins weniger teure Ausland ausgewichen wären, wenn sie nicht so spezialisiert wären. Auch die Kunden können umweltbewusst sein, und wenn Ausgangsmaterialien sehr langlebig sind, wie in der stoffabhängigen Baubranche, dann kann es sich lohnen, auf Kreisläufe zu setzen, obwohl dazu die Prozesse umstrukturiert werden müssen. FIT will den KMU als Hilfe eine strukturierte Liste der helfenden und hindernden Faktoren bieten: Oft fehlen Unterstützer. Staatliche Regulierung ist nicht vorhanden (so beim internationalen Abfall-Management), Kunden oder Lieferanten ziehen nicht mit, oder es fehlt an Kapital. Verwaltungs-Hindernisse können den Unternehmen das Leben schwer machen, aber auch ein Mangel an Information oder technischem Knowhow. Eine umweltbewusste Firmenkultur will gepflanzt und gepflegt werden. Für manche Produkte gestaltet sich das Kreislauf-taugliche Design zu kompliziert, wenn beispielsweise ein Gerät aus Sicherheitsgründen nicht geöffnet werden darf. Treibende Kräfte der Transformation spiegeln oft die Barrieren: Eine ökologische Kultur im Unternehmen und seinem Umfeld motiviert die Beschäftigten, die gesteigerten Anforderungen ihres Arbeitsalltags in den Kreislauf-Prozessen zu erfüllen. Sie ermöglicht Netzwerke, auch mit externen Auditoren, und schafft finanziell einträgliche Nachfrage, wachsendes Renommee, staatliche Förderung, helfende Lieferanten. Immer mehr Firmen entdecken die Vorteile kostensparender Prozesse: weniger Energie- und Ressourcenverbrauch, Lieferanten, die beispielsweise Aus- und Überschüsse an Material aus der Fertigung zurücknehmen, haltbare

Diskussion Digitalization for sustainability: Top-down or bottom-up transformations? *mit Radhika Krishnan (IIIT Hyderabad), Teresa Cerratto Pargman (Stockholm University), Andrea Hamm (Weizenbaum Institute for the Networked Society), Moderation Angela Oels*

Maschinen, die es möglich machen Investitionen zu strecken *(Retrofitting)*. Die IT kann hilfreiche Werkzeuge bieten, sowohl um Fertigungsabfall zu vermeiden, als auch um die Materialmenge zu reduzieren.

Greenwashing

Sustainability Narratives in Digitalization Research Funding, der Vortrag von *Mario Angst* (Universität Zürich), bestätigt Zweifel an der Wahrhaftigkeit des Transformations-Willens. Für diesen Willlen wären Forschungsmittel ein wesentlicher Indikator und Hebel. Der Forscher analysierte 37 000 Abstracts in vier Sprachen (er arbeitet schließlich in der Schweiz) nach ihrem Anteil von Nachhaltigkeit und Digitalisierung. Die nationale Forschungseinrichtung der Schweiz, SNF, veröffentlicht die Daten über Fördermittel[6]. Sie gibt damit natürlich nur Auskunft über geförderte Projekte, nicht aber darüber, was nicht gefördert wurde und warum. Angst kann verschiedene Erzählungen des Nachhaltigkeits-Diskurses identifizieren, begleitet von den jeweiligen *buzz words* als Traumfängern für Nachhaltigkeit ohne Verhaltensänderung oder Verzicht. *Smart* ist so ein allgegenwärtiger Begriff, am liebsten mit Energie verbunden. Effizienz rangiert weit oben, berücksichtigt aber den strukturellen Wandel auf Makro-Ebene nicht. Suffizienz dagegen fehlt, sowohl bei Forschungsvorhaben zum Lebenszyklus als auch bei Ermächtigung oder Ertüchtigung.[7]

Botschaft vom kollektiven Handeln

Al-Amin Dabo (University of Northampton) berichtet von entstehender und wachsender Kreislaufwirtschaft in der sich entwickelnden Welt. Schon lange sammeln und sortieren Menschen Wertstoffe auf den Mülldeponien. Kommunale Aktion und gemeinsames Handeln erweisen sich als effektiv. In seiner For-

schung untersucht Dabo, welche Rolle digitale Werkzeuge wie Plattformen und andere Instrumente für die Vernetzung, als Unterstützung bei der Daten-Analyse, in der Bildung oder für soziale Gemeinschaften spielen können. Das Problem der Trittbrettfahrer, auf das Olsons *theory of collective action* hinweist, lässt sich durch die Technik abmildern: Wo personelle und materielle Ressourcen aufgetrieben werden müssen, bietet die Digitalisierung ein hilfreiches Kommunikations-Instrument, ebenso um den Zusammenhalt der Gruppen zu fördern, gemeinsame Ziele zu finden und abzustimmen und auch bei der Entscheidungsfindung durch Abstimmungen oder andere Formen Konsens zu erzielen. In seinem Vortrag *Understanding digital enablers for a circular future* erklärt Dabo, wie er in einem Methoden-Mix die am Abfall-Recycling Beteiligten in Interviews und Fragebögen befragte, Fokus-Gruppen untersuchte und als Beobachter selbst teilnahm, um herauszufinden, wie die Technologie im und auf das soziale Umfeld wirkt, wie sie sich am effektivsten nutzen lässt und die unterschiedlichen Akteure am besten einbezieht und mobilisiert.

Es bleiben viele Fragen, wenn der informelle Sektor der Abfallwirtschaft in ein kommunales Management einbezogen werden soll: Welche Rolle können dabei digitale Plattformen spielen? Kann die Technik das Vertrauen in den Sammler-Gemeinschaften fördern? Würde *IoT* im Sammelprozess dabei unterstützen, Arten und Mengen der Stoffe effizienter zu überwachen und zu verarbeiten? Welche Wirkung haben kollaborative Werkzeuge auf Arbeitsteilung und Koordination? Kann die Analyse der Daten beim Planen der Routen und Zeiten helfen?

Und die KI?

In seinem Vortrag *Artificial Intelligence for real politics?* macht der stellvertretende FIfF-Vorsitzende *Rainer Rehak* wenig Hoffnung, dass eine KI-Revolution die technische Lösung bringen könnte, die kollektives Handeln überflüssig machen und die nachhaltige Transformation durch optimale und gleichzeitig neutrale Entscheidungen im Feld der Politik ganz *objektiv* herbeiführen könnte. So wird es nicht gehen, die politische Auseinandersetzung wird uns nicht erspart bleiben.

Eine mit Bewusstsein und Kreativität begabte KI ist nirgends in Sicht, neben der schwachen KI versammelt die von Rehak so betitelte *Zeitgeist-KI (Artificial Zeitgeist Intelligence)* zahlreiche informatische Methoden und Werkzeuge. Sie können weder Entscheidungs- noch Optimierungskriterien definieren noch Fragen aufwerfen oder Lösungen außerhalb ihrer ursprünglichen Domäne suchen. Hilfreich kann KI dennoch sein. Rehak nennt konkrete Einsatzgebiete: Erdbeobachtung, das Erken-

Die VeranstalterÏnnen, in der Mitte Prof. Dr. Angela Oels. Alle Fotos A. Kaltenberg, Zentrum für Klimaresilienz der Universität Augsburg (CC-BY 4.0)

nen und Messen von Kontaminierungen, bessere Kreislauf-Verfahren, Abfalltrennung und -bearbeitung, das Minimieren von Rohstoff- und Energieeinsatz, vorausschauende Wartung, *grüne* Mobilität und Reduzierung des CO_2-Ausstoßes. Rehak erinnert daran, dass KI dabei nie neutral ist. Sie ist auch kein Akteur. Optimierungsziele müssen politisch vorgegeben werden. Die Macht der so beeindruckenden großen Sprach-Modelle (LLMs) lässt sich nicht demokratisieren. Er kritisiert die Entpolitisierung durch KI und plädiert dafür, KI zu entmystifizieren. Technische Lösungen gesellschaftlicher Probleme sind keine Lösungen.

Anmerkungen

1 *ausführlich auf der Konferenz-Website: https://www.uni-augsburg.de/de/fakultaet/philsoz/fakultat/powi-klimapolitik/conference/ (abgerufen 1.11.2023)*
2 *foodlandscapes.com.au (abgerufen 1.11.2023)*
3 *Nach meinem Eindruck kein einfaches Unterfangen, denn dort sind sie selten transparent.*
4 *So nutzt VW als Marketing-Instrument Thefuntheorycompany.*
5 *https://www.xrtoday.com/mixed-reality/top-xr-arts-entertainment-vendors-for-2023/; https://www.xrmust.com/xrmagazine/?category=Creative (abgerufen 1.11.2023)*
6 *Data.snf.ch/datasets (abgerufen 1.11.2023)*
7 *zu Meta-Reflektionen über Computergestützte Sozialwissenschaft: sustainability.discourses, deutschsprachig https://sustainability.discourses.ch/de/ (abgerufen 1.11.2023)*

Dagmar Boedicker

Dagmar Boedicker ist Journalistin, technische Redakteurin und langjährige Redakteurin der FIfF-Kommunikation.

Hans-Jörg Kreowski

Wissenschaft für den Frieden aus naturwissenschaftlich-technischer Sicht

Editorial zum Schwerpunkt

Die Zeitschrift Wissenschaft und Frieden *hat anlässlich ihres 40-jährigen Bestehens am 6. und 7. Oktober 2023 in Bonn das Symposium* Wissenschaft für den Frieden *mit 120 Teilnehmer:innen veranstaltet. Das Programm umfasste 20 Vorträge, 14 Panels und 14 Workshops sowie einen Festakt mit einem wissenschaftshistorischen Dialog zwischen Eva Senghaas-Knobloch und Jürgen Altmann. Das FIfF hat sich als eine von elf Trägerorganisationen in bescheidenem Umfang an der Finanzierung, aber stark an der Organisation beteiligt. Wie die Zeitschrift hat auch das Symposium die Friedens- und Konfliktforschung im weitesten Sinne und in einem breiten interdisziplinären Spektrum widergespiegelt. Neben Beiträgen aus Friedenspädagogik, Friedenspsychologie, Friedenspolitik, Soziologie, Philosophie und Recht sowie Kunst und Kultur wurden auch aktuelle politische Themen wie die feministische Außenpolitik, die nationale Sicherheitsstrategie und mit zwei Panels Russlands Krieg gegen die Ukraine diskutiert. Ein großer Block des Vortragsprogramms galt naturwissenschaftlichen und technischen Aspekten – vielfach mit Informatikbezug. Dieser Teil des Symposiumsprogramms wird in diesem Heft als Schwerpunkt dokumentiert.*

Der Schwerpunkt setzt sich zusammen aus vier schriftlichen Ausarbeitungen:

- Anja-Liisa Gonsior, Thea Riebe, Stefka Schmid, Thomas Reinhold, Christian Reuter: *Friedensinformatik: heute und morgen*
- Karl Hans Bläsius, Jörg Siekmann: *Computergestützte Frühwarn- und Entscheidungssysteme für nukleare Bedrohungen*
- Christian Heck: *Kritik des gläsernen Gefechtsfeldes – Was Sprachmodelle und Massendaten im Krieg bedeuten*
- Ryan R. Swan: *Neubewertung einer unabhängigen Satellitenüberwachungseinheit: Ein Mechanismus, um Transparenz in Kooperation umzuwandeln?*

Der Beitrag von Christian Heck ist ein Nachdruck aus dem Jubiläumsheft 4/2023 der Zeitschrift *Wissenschaft und Frieden* mit freundlicher Genehmigung der W&F-Redaktion und des Autors.

Um einen vollständigeren Überblick zu vermitteln, werden fünf weitere Beiträge kurzgefasst wiedergegeben. Die Texte sind die Beschreibungen der Beiträge zu dem Symposium:

- Timothy Williams: *Frieden und Konflikt in der digitalen Ära*
- Jens Hälterlein: *ELSA zieht in den Krieg – Zur Rolle der Kritik an autonomen Waffensystemen für deren Legitimationsstrategien*
- Dieter Engels, Jürgen Scheffran, Ekkehard Sieker: *Krieg im Weltraum? Es ist mal wieder Fünf vor Zwölf*
- Lukas Rademacher: *Wie verifiziert man nukleare Abrüstung?*
- Rebecca Froese, Daniela Pastoors, Jürgen Scheffran und Melanie Hussak: *„Frieden verbessert das Klima" – Wie Konflikttransformation zur Bewältigung der Klimakrise beitragen kann*

Da die Einreichungsfrist dieses Schwerpunkts nur wenige Wochen nach dem Symposium lag, konnten die Autor:innen aus zeitlichen Gründen keine Langfassungen schreiben. Es besteht die Chance, dass das in dem einen oder anderen Fall für eine der nächsten FIfF-Kommunikationen nachgeholt wird.

Außerdem ist im Schwerpunkt das Grußwort von Stefan Hügel als FIfF-Vorsitzendem zum 40-jährigen Bestehen von W&F abgedruckt sowie ein kurzer Text zu Sinn und Zweck der Zeitschrift und des Symposiums. Aus meiner Sicht spielt die Mitwirkung des FIfF bei W&F eine wichtige Rolle. In den Anfangsjahren war die Zeitschrift stark geprägt von naturwissenschaftlichen Fragestellungen insbesondere im Zusammenhang mit der Atomrüstung und den daraus resultierenden Gefahren eines Atomkriegs. Dann aber hat sich eine weitgehende sozial- und geisteswissenschaftliche Orientierung durchgesetzt. Dem FIfF fällt damit die Aufgabe zu, die gravierenden Auswirkungen der Digitalisierung des Kriegsgeschäfts und die damit verbundenen Gefährdungen einer friedlichen Entwicklung der Gesellschaft thematisch einzu-

bringen. Das ist in den letzten Jahren oft gelungen, könnte aber noch intensiviert werden. W&F ist ein passendes Publikationsorgan für das Thema Rüstung und Informatik in all seinen Facetten, wenn eine Leser:innenschaft weit jenseits der Informatik erreicht werden soll.

Wenn eine seit 40 Jahren bestehende Zeitschrift wie auch das fast so alte FIfF „Frieden" im Namen führen, dann muss man leider konstatieren, dass diese Thematik weiterhin schreckliche Aktualität besitzt, weil Krieg eine alltägliche Realität darstellt, weil Konflikte in vielen Teilen der Welt zu eskalieren drohen und weil Politik regional bis weltweit wenig erfolgreich dagegen arbeitet, wenn sie sich überhaupt darum kümmert. Ein betrübliches, ja skandalöses Beispiel hat gerade der deutsche Verteidigungsminister Boris Pistorius gegeben, der bei *Berlin direkt*[1] sagt: „Wir müssen kriegstüchtig werden." Sein Versuch, diesen Begriff als synonym zu „verteidigungsfähig" hinzustellen, ist irreführend. Denn hätte er „verteidigungsfähig" gemeint, hätte er das ja auch sagen können. „Kriegstüchtig" bedeutet laut Duden, für einen Krieg gut gerüstet zu sein. Für welchen Krieg? Wo und gegen wen soll Deutschland Krieg führen? Krieg ist nach der UN-Charta verboten. Wozu müssen wir, Deutschland, die Bundeswehr „tüchtig" sein, Verbotenes zu tun? Es gibt allerdings zwei Ausnahmen: Wenn ein Staat angegriffen wird, darf er sich verteidigen; und wenn die UN einen Krieg beenden oder verhindern will, darf sie militärisch eingreifen. „Verteidigungsfähig" ist

Hans-Jörg Kreowski auf dem Jubiläumssymposium

also das viel richtigere Wort, das auch durch das Grundgesetz gedeckt ist. Der Kriegsminister Pistorius muss also wohl etwas anderes meinen, und er steht da in der Regierung anscheinend nicht allein. Wie wäre es mit „friedenstüchtig"?

Um einen berühmten Dichter des 20. Jahrhunderts zu zitieren: *„Imagine all the people living life in peace"* (John Lennon 1971).

1 *Verteidigungsminister Pistorius im Interview mit Dominik Rzepka, zdfheute am 29.10.2023 um 20:37 h*

Anja-Liisa Gonsior, Thea Riebe, Stefka Schmid, Thomas Reinhold und Christian Reuter

Friedensinformatik: heute und morgen

Fortschritte in Wissenschaft und Technologie spielen eine entscheidende Rolle im Zusammenhang mit Frieden, Konflikt und Sicherheit (Reuter 2019). Insbesondere die Rolle der Informatik in der Friedens- und Konfliktforschung hat sich durch die Digitalisierung der Gesellschaft gewandelt. Die Bewältigung der damit verbundenen Herausforderungen für Frieden und Sicherheit durch die akademische Forschung erfordert die Anwendung und Etablierung interdisziplinärer Ansätze (Reuter et al. 2020). An dieser Stelle kann die naturwissenschaftlich-technische Friedens- und Konfliktforschung entscheidende Beiträge leisten, um aktuelle Themen und damit verbundene Problemstellungen aus verschiedenen disziplinären Perspektiven zu analysieren und zu bewerten. So müssen beispielsweise Fragen im Kontext von Cyberangriffen oder Cyberwaffen sowohl aus der Perspektive der Informatik als auch der Politikwissenschaft betrachtet werden (Reuter et al. 2020).

Einleitung

Das Forschungsfeld der *Friedensinformatik* kann als Subdisziplin der naturwissenschaftlich-technischen Friedens- und Konfliktforschung bezeichnet werden. Diese beschäftigt sich mit der Rolle informationstechnischer Artefakte in Konflikten und Kriegen sowie mit der Gestaltung von informationstechnischen Systemen, welche zur Gewaltvermeidung oder friedlichen Transformation von Konflikten beitragen können. Dazu gehören Aspekte wie die Widerstandsfähigkeit von informationstechnischen Infrastrukturen, beispielsweise als Zielscheibe in Konfliktsituationen, aber auch die Rolle von IT-Anwendungen bei der Prävention und Bewältigung von Konflikten, Krisen und Katastrophen (Reuter 2019). Darüber hinaus behandelt die Friedensinformatik unter anderem Themen in den Bereichen Cyber-Rüstungskontrolle, *Dual-Use*-Forschung, künstliche Intelligenz (KI) und autonome Systeme. Es werden interdisziplinär Konzepte und Methoden aus der Informatik und den Sozialwissenschaften zusammengeführt, wie der Mensch-Computer-Interaktion, IT-Sicherheit, Technikfolgenabschätzung und Politikwissenschaft. In den *Empfehlungen zur Weiterentwicklung der Friedens- und Konfliktforschung* wies der deutsche Wissenschaftsrat 2019 auf den dringenden Handlungsbedarf zur Stärkung der naturwissenschaftlich-technischen Friedens- und Konfliktforschung hin, die in Deutschland derzeit strukturell zu schwach ist, um den massiven Bedarf an Politikberatung zu decken (Wissenschaftsrat 2019). Dieser Beitrag möchte einen Einblick in ausgewählte aktuelle Arbeitsthemen und Bereiche der Friedensinformatik am Lehrstuhl *Wissenschaft und Technik für Frieden und Sicherheit* (PEASEC) der Technischen Universität Darmstadt geben.

Cyberwaffen, die Militarisierung des Cyberspace und Cyber-Rüstungskontrolle

Mit Blick auf die zunehmende Militarisierung des Cyberspace und die damit verbundenen Gefahren und Bedrohungen erge-

ben sich in diesem Bereich vielfältige Herausforderungen. In diesem Kontext wappnen sich weltweit bereits viele nationale und internationale Sicherheitsdoktrinen gegen Software, die entwickelt wurde, um in IT-Systeme eingespeist zu werden mit dem Ziel der Spionage oder Sabotage (Reinhold & Reuter 2022). Das prominenteste Beispiel ist sicher der Einsatz der Schadsoftware *Stuxnet*, die im Jahr 2010 entdeckt wurde und die das industrielle Kontrollsystem einer Nuklearanlage im Iran manipulierte mit dem Ziel, Schwellenwerte und Parameter der Kontrollsoftware heimlich zu ändern, um dadurch den Produktionsprozess zu sabotieren (Langner 2013). Das Beispiel demonstriert, wie Bedrohungen aus dem Cyberraum Einfluss auf physische Infrastrukturen haben können. Vor dem Hintergrund derartiger Bedrohungen versuchen nationale Regierungen und Geheimdienste nicht nur, geeignete Verteidigungsmaßnahmen gegen Schwachstellen von Computersystemen zu etablieren, sondern teilweise gleichzeitig auch, dies für die offensive Planung von eigenen Cyberangriffen zu nutzen (Reinhold & Reuter 2022).

Durch die steigende Abhängigkeit von Informationstechnologien in ökonomischen, gesellschaftlichen und politischen Bereichen ergeben sich zahlreiche Herausforderungen. Dennoch bzw. gerade deswegen besteht derzeit kein einheitliches international anerkanntes Verständnis über die Bedrohung durch Cyberwaffen und ihre Verhinderung, geschweige denn über ein verbindliches Rechtsinstrument. Eine weitere Herausforderung stellt das sogenannte *Attributionsproblem* dar, also der forensische und politische Prozess der Gewinnung gesicherter Erkenntnisse über den Ursprung eines Cyberangriffs (Saalbach 2019). Attribution stellt ein wichtiges Instrument dar, um die Wirksamkeit und Durchsetzbarkeit völkerrechtlicher Normen und Grundlagen auch im Cyberspace zu verbessern, indem der Akteur hinter einem Angriff klar benannt und in Verantwortung genommen werden kann. Dies gilt beispielsweise für den gebotenen Schutz der Zivilbevölkerung oder die Rechtmäßigkeit von Verteidigungsmaßnahmen eines angegriffenen Staates. Auch für die Rüstungskontrolle ist es sinnvoll, den Einsatz einer bestimmten Cyberwaffe und deren Ursprung zu ermitteln. Neben einer fehlenden Definition von Cyberwaffen und dem Attributionsproblem erschweren zudem die Virtualität des Cyberspace, die fehlende physische Form von Sicherheitslücken und Schadsoftware, ständige technologische Weiterentwicklung, fehlender politischer Wille, Verifizierung sowie der Einfluss zahlreicher Akteure den Umgang mit Cyberwaffen (Reinhold et al. 2023). Daher versagen vielfach etablierte Konzepte der Rüstungskontrolle für den Cyberspace und neue Ansätze, die den technischen Besonderheiten dieses Raumes Rechnung tragen, sind dringend von Nöten. Die *Friedensinformatik* kann hier wertvolle Beiträge leisten. So etwa bei der Frage danach, wie sich *Cyberwaffen* bewerten lassen (Reinhold & Reuter 2021), aber auch anhand welcher technischen mess- und erfassbaren Parameter eine Schadsoftware vor deren Einsatz und unabhängig von mutmaßlichen Absichten klassifiziert und reguliert werden kann.

Autonomie in Waffensystemen

Die Anwendung von KI in vielfältigen gesellschaftlichen Bereichen ist derzeit in aller Munde. Ein besonders schnell voranschreitender und kritischer Anwendungskontext ist die Integration von Autonomie in Waffensystemen – sogenannten *(lethal) autonomous weapon systems* ((L)AWS). Die Debatte über die Entwicklung und den Einsatz von AWS als Technologie gewinnt zunehmend an Bedeutung, wobei internationale Verhandlungen ins Stocken geraten, während gleichzeitig die technologische Entwicklung immer weiter voranschreitet (Riebe et al. 2020). Stetige technologische Weiterentwicklungen sind nur ein Grund dafür, warum es bis heute keine allgemeingültige bzw. international anerkannte Definition von AWS gibt. Das *International Committee of the Red Cross* (ICRC) schreibt dazu: „Autonomous weapon systems select and apply force to targets without human intervention" (ICRC 2021). Das US-Verteidigungsministerium vertritt ein ähnliches Verständnis und definiert ein autonomes Waffensystem als „weapon system that, once activated, can select and engage targets without further intervention by an operator" (DoD 2023). Diese Definitionen verdeutlichen den Unterschied zwischen AWS und konventionellen Waffen, da Autonomie in Waffensystemen viel stärker im Hinblick auf den Anwendungs- bzw. Einsatzkontext definiert werden muss, weil es sich nicht um eine klar abgrenzbare Waffenkategorie handelt. Neben der definitorischen Problematik werfen autonome Waffensysteme technische, (völker-)rechtliche, sicherheitspolitische, ethische und humanitäre Fragen auf. Hier ist noch viel Forschung nötig, deren Ergebnisse zwischen Wissenschaft, Politik und Zivilgesellschaft diskutiert werden müssen. Außerdem sollten auch Vertreter:innen aus Militär und Industrie in Gespräche integriert werden.

Mit Blick auf eine angemessene politische und rechtliche Regulierung solcher neuartigen Waffensysteme müssen daher neue Konzepte im Rahmen der Rüstungskontrolle gefunden werden (Sauer 2021). Die Entwicklungen und möglichen Regulierungen rund um AWS werden seit einigen Jahren innerhalb der *UN-Konvention über bestimmte konventionelle Waffen* (engl. *Convention on Certain Conventional Weapons* (CCW)) im Rahmen einer *Group of Governmental Experts* (GGE) zwischen Mitgliedsstaaten, Zivilgesellschaft und Fachexpert:innen diskutiert. Das Element der menschlichen Kontrolle, Fragen im Kontext der Mensch-Maschine-Interaktion sowie die *human security* rücken dabei immer mehr in den Vordergrund. In diesem Kontext weist die internationale NGO-Kampagne *Campaign to Stop Killer Robots* auf die Bedeutung intersektionaler Ansätze hin und verdeutlicht beispielsweise, in welchem Ausmaß *Gender*- und *Race*-Bias in der Konzeption, Technologie und Anwendung von LAWS in der Kriegsführung vorhanden sind (Stop Killer Robots 2020). Auf dieser Basis skizziert die Kampagne zukünftige Herausforderungen für die Einordnung dieser Waffensysteme. Auch in der GGE werden diese Aspekte in den letzten Jahren vermehrt in den Gesprächen aufgegriffen, auch wenn technische, rechtliche und sicherheitspolitische Themen weiterhin im Fokus stehen.

Dual-Use-Technologien

Die Einordnung von Autonomie und KI in Waffensystemen wird durch ihre *Dual-Use*-Eigenschaften erschwert, denn Autonomie ist eine Eigenschaft, die potenziell in unterschiedliche Systeme integriert werden kann. Daher wird auch immer häufiger von *Autonomie in Waffensystemen* statt von *autonomen Waffen-*

systemen gesprochen, beispielsweise in der neuen Richtlinie des US-Verteidigungsministeriums (DoD 2023). *Dual-Use* bezeichnet hierbei die Möglichkeit, eine Technologie sowohl für militärische als auch für zivile Zwecke zu nutzen. Der Dual-Use-Charakter vieler Informations- und Kommunikationstechnologien (IKT) wirft neue Fragen für Forschung und Entwicklung sowie für die nationale, internationale und menschliche Sicherheit auf (Riebe 2023).

Die Herausforderungen für die Forschung ergeben sich hierbei durch die Kombination von konventionellen mit automatisierten oder autonomen Trägersystemen, aber auch durch völlig neuartige Systeme zum Mensch-Maschine-Teaming und zur Unterstützung von Menschen, zum Beispiel durch autonome Systeme wie Drohnen, Roboterhunde oder auch Exoskelette. Innerhalb der Disziplinen der Technikfolgenabschätzung, kritischer Sicherheitsforschung und Mensch-Computer-Interaktion können Fragen zum Beispiel im Hinblick auf die Einhaltung des Humanitären Völkerrechts, die Diffusion von Technologien und Wissen und effektive Rüstungskontrollmaßnahmen analysiert werden. Beispielweise müssen vor dem Hintergrund verantwortungsvoller Forschung und Entwicklung technologische Diffusionen von KI vom zivilen auf den militärischen Bereich berücksichtigt werden (Schmid et al. 2022). Auch in anderen Bereichen der IT – unter anderem in Bezug auf Software oder Kryptographie – spielt der Dual-Use-Charakter eine Rolle.

Maßnahmen zur Bewältigung der Risiken, die mit verschiedenen Dual-Use-Technologien einhergehen, beispielweise durch Proliferationskontrollen oder politische Maßnahmen, gestalten sich mitunter sehr unterschiedlich. Dabei werfen Innovationen in verschiedenen Bereichen – so etwa im Kontext von KI, Robotik und Cybersicherheit – neue Fragen zu jeweiligen Dual-Use-Risiken auf.

Internationale Visionen und Governance von Künstlicher Intelligenz

Jüngste Forschung und Entwicklung im Bereich der *KI* wird gesellschaftlich breit, auch im Hinblick auf Herausforderungen, diskutiert. Beispiele wie der Konversationsbot ChatGPT, *predictive maintenance* in der Logistik oder militärische KI verweisen darauf, dass Schlüsseltechnologien Anwendung in verschiedenen Formen und Kontexten finden und gesellschaftlich transformativ wirken können. Gleichzeitig können neben den Herausforderungen von KI für die Rüstungskontrolle einige dieser Verfahren aber auch als Werkzeug *für* die Rüstungskontrolle verstanden und eingesetzt werden (Reinhold & Reuter 2022), insbesondere wenn zum Teil enorme Informationssammlungen und Sensordaten analysiert werden müssen.

In Bezug auf internationale Sicherheit zeigt sich zudem der Trend der Geopolitisierung von Innovationen, welcher die Förderung von Technologien für nationale und machtpolitische Zwecke umfasst. Staatliche Akteure prägen durch Regulierung Innovationsprozesse entscheidend. Es lohnt daher eine Analyse ihrer Innovationspolitiken und damit verknüpften Visionen von Technologien sowie der Mensch-KI-Beziehung. In Bezug auf die Europäische Union (EU) zeigt sich beispielsweise, dass diese konkrete politische Maßnahmen zur Erforschung und Entwicklung von KI ergriffen hat. Hinsichtlich Debatten über die Technikfolgenabschätzung, die sich auf die Risiken für den Menschen und Fragen der Kontrolle von KI konzentrieren, vertritt die EU in diesem Zusammenhang einen ethischen, auf den Menschen ausgerichteten Ansatz für die Anwendung von KI. Um in diesem Zusammenhang die Entstehung von Normen zu verstehen, muss vor dem Hintergrund der Mensch-Computer-Interaktion das Verständnis des Akteurs für die Interaktion zwischen Mensch und KI analysiert werden, wobei die Konzeptualisierungen von Erklärbarkeit, Interpretierbarkeit und Risiken eine wichtige Rolle spielen. Hier wird zudem erneut die Relevanz interdisziplinärer Ansätze für das detailliertere Verständnis der Visionen unterschiedlicher Akteure über die menschliche Kontrolle von KI verdeutlicht, vor allem hinsichtlich der Bewertung und Governance solcher Technologien (Schmid 2022).

Ausblick auf zukünftige Fragen der Friedensinformatik

Der Einsatz von Informatik-Artefakten in Kriegen und Konflikten wird zunehmen. Daher wird auch die Friedensinformatik in Zukunft unter anderem Fragen zur Attribution und Verifikation von Cyberwaffen, zur Mensch-Maschine-Interaktion und deren Auswirkung auf die internationale Sicherheit erforschen. Konfliktdynamiken im Cyberspace, sowie der Einsatz autonomer Waffensysteme insbesondere vor dem Hintergrund der Nutzung von KI, stellen uns weiterhin vor relevante Forschungsfragen. Aus der *Governance*-Perspektive ist zudem von großer Bedeutung, welche Schwerpunkte bei der Innovationsförderung im Bereich der zivil-militärischen *Dual-Use*-Güter gesetzt werden, um technologische Trends zu beeinflussen. Da IT eine Querschnittstechnologie darstellt, werden durch neue Entwicklungen zahlreiche andere Bereiche mitbeeinflusst, so beispielsweise auch im Kontext konventioneller Waffen. Daher müssen auch Einflüsse auf bereits etablierte Forschungsbereiche berücksichtigt werden.

In Zukunft werden IT-Systeme – auch vor dem Hintergrund von KI – zunehmend untereinander und mit dem Cyberspace verbunden sein, was bedeutet, dass die Verteidigung gegen Cyberangriffe eine immer größere Bandbreite an verteilten digitalen Geräten umfassen wird, die entsprechend noch widerstandsfähiger gegen Angriffe und Abhängigkeiten gemacht werden müssen. Dadurch sowie durch die zunehmende Menge an Informationen wird sich auch das Spektrum möglicher Angriffsvektoren vergrößern und diversifizieren (Reinhold & Reuter 2022). In Bezug auf Autonomie in Waffensystemen und andere Dual-Use-Technologien muss in Zukunft weiterhin erörtert werden, wie eine potentielle Rüstungskontrolle aussehen könnte oder ob eventuell weichere Formen der Regulierung – beispielweise in Form sogenannter *soft-law*-Mechanismen – wirkungsvoller sein können hinsichtlich der Bildung neuer starker Normen (Rosert 2021). Dabei können politische Aushandlungsprozesse, bei denen verschiedene Akteursperspektiven aufeinandertreffen, maßgeblich für eine wertgeleitete und zukunftsgerichtete Technikfolgenabschätzung und verantwortungsbewusste Designprozesse sein.

Referenzen

DoD. (2023): DoD Directive 3000.09. Autonomy in Weapon Systems, https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf.

ICRC. (2021): ICRC position on autonomous weapon systems, https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems.

Langner, R. (2013): A Technical Analysis of What Stuxnet's Creators Tried to Achieve, https://www.langner.com/wp-content/uploads/2017/03/to-kill-a-centrifuge.pdf.

Reinhold, T.; H. Pleil & C. Reuter (2023). Challenges for Cyber Arms Control: A Qualitative Expert Interview Study. Zeitschrift Für Außen- Und Sicherheitspolitik, 16(3), 289–310. https://doi.org/10.1007/s12399-023-00960-w.

Reinhold, T. & C. Reuter (2022): Cyber Weapons and Artificial Intelligence: Impact, Influence and the Challenges for Arms Control. In Reinhold, T. & N. Schörnig (Eds.), Armament, arms control and artificial intelligence: The janus-faced nature of machine learning in the military realm. Springer. https://doi.org/10.1007/978-3-031-11043-6.

Reinhold, T. & C. Reuter (2021): Towards a Cyber Weapons Assessment Model – Assessment of the Technical Features of Malicious Software. IEEE Transactions on Technology and Society.

Reuter, C. (Ed.). (2019): Information Technology for Peace and Security: IT Applications and Infrastructures in Conflicts, Crises, War, and Peace. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-25652-4.

Reuter, C.; J. Altmann; M. Göttsche & M. Himmel (2020): Natural Science and Technical Peace Research: Definition, History, and Current Work. Sicherheit Und Frieden (S+F), 38(1).

Riebe, T. (2023): Technology Assessment of Dual-Use ICTs—How to Assess Diffusion, Governance and Design, https://doi.org/10.26083/TU-PRINTS-00022849.

Riebe, T.; S. Schmid & C. Reuter (2020): Meaningful Human Control of Lethal Autonomous Weapon Systems: The CCW-Debate and Its Implications for VSD. IEEE Technology and Society Magazine, 39(4), 36–51, https://doi.org/10.1109/MTS.2020.3031846.

Rosert, E. (2021): Autonomie in Waffensystemen: Menschliche Kontrolle verbindlich vorschreiben, die UNCCW stärken. In U. Kühn (Ed.), Research Report: Rüstungskontrolle für die nächste Bundesregierung. Ein Empfehlungsbericht (pp. 48–53). IFSH. https://ifsh.de/publikationen/research-report-006.

Saalbach, K.-P. (2019): Attribution of Cyber Attacks. In C. Reuter (Ed.), Information Technology for Peace and Security: IT Applications and Infrastructures in Conflicts, Crises, War, and Peace.

Sauer, F. (2021): Stepping back from the brink: Why multilateral regulation of autonomy in weapons systems is difficult, yet imperative and feasible. International Review of the Red Cross, 102(913), 235–259. https://doi.org/10.1017/S1816383120000466.

Schmid, S. (2022): Trustworthy and Explainable: A European Vision of (Weaponised) Artificial Intelligence. Die Friedens-Warte, 95(3–4), 290. https://doi.org/10.35998/fw-2022-0013.

Schmid, S.; T. Riebe, & C. Reuter (2022): Dual-Use and Trustworthy? A Mixed Methods Analysis of AI Diffusion Between Civilian and Defense R&D. Science and Engineering Ethics, 28(2), 12. https://doi.org/10.1007/s11948-022-00364-7.

Stop Killer Robots (2020): Intersectionality and Racism, https://www.stopkillerrobots.org/resource/intersectionality-and-racism/.

Wissenschaftsrat. (2019): Empfehlungen zur Weiterentwicklung der Friedens- und Konfliktforschung, (Drs. 7827-19), pp. 1–178, https://www. wissenschaftsrat.de/download/2019/7827-19.html.

Anja-Liisa Gonsior, Thea Riebe, Stefka Schmid, Thomas Reinhold und Christian Reuter

Anja-Liisa Gonsior ist wissenschaftliche Mitarbeiterin und Doktorandin am Lehrstuhl *Wissenschaft und Technik für Frieden und Sicherheit* (PEASEC) am Fachbereich Informatik der Technischen Universität Darmstadt. Ihre Forschungsinteressen liegen in den Bereichen autonome Waffensysteme, Meaningful Human Control, (Cyber-) Rüstungskontrolle, (naturwissenschaftlich-technische) Friedens- und Konfliktforschung sowie kritische Sicherheitsstudien.

Dr. **Thea Riebe** ist wissenschaftliche Mitarbeiterin und Postdoktorandin am Lehrstuhl PEASEC und forscht zu Technikfolgenabschätzung von Dual-Use-Technologien in der Informatik und verbindet Ansätze aus Technikfolgenabschätzung, Kritischer Sicherheitsforschung und Mensch-Computer-Interaktion.

Stefka Schmid ist wissenschaftliche Mitarbeiterin am Lehrstuhl PEASEC. Ihre Forschungsinteressen sind Innovationspolitiken als Gegenstand kritischer Sicherheitsstudien, die naturwissenschaftlich-technische Friedens- und Konfliktforschung sowie Mensch-Computer-Interaktion in Krisenszenarien. In ihrer Promotion setzt sie sich mit der Nutzung von Technologien seitens kollektiver Akteure im Kontext globaler Sicherheitspolitiken auseinander.

Dr. **Thomas Reinhold** ist wissenschaftlicher Mitarbeiter sowie Postdoktorand im gemeinsamen Projekt *Cluster Natur- und Technikwissenschaftliche Rüstungskontrollforschung (CNTR)* des Leibniz-Institut für Friedens- und Konfliktforschung (PRIF) sowie des Lehrstuhls PEASEC. Im Mittelpunkt seines wissenschaftlichen Interesses stehen die Bedrohungen im Cyberspace und das Problem der zunehmenden Militarisierung dieser Domäne.

Prof. Dr. Dr. **Christian Reuter** ist Universitätsprofessor und Dekan am Fachbereich Informatik der Technischen Universität Darmstadt. Sein Lehrstuhl Wissenschaft und Technik für Frieden und Sicherheit (PEASEC) verbindet Informatik mit Friedens- und Sicherheitsforschung. Er hält Doktorgrade in Wirtschaftsinformatik (Universität Siegen) sowie in Sicherheitspolitik (Radboud Universiteit Nijmegen).

Karl Hans Bläsius und Jörg Siekmann

Computergestützte Frühwarn- und Entscheidungssysteme für nukleare Bedrohungen

Computergestützte Frühwarn- und Entscheidungssysteme dienen der Erkennung eines Angriffs mit Atomwaffen. Hierbei kann es aber zu Fehlalarmen kommen, bei denen nukleare Angriffe gemeldet werden, obwohl kein Angriff vorliegt. Solche Fehler könnten zu einem Atomkrieg aus Versehen *führen. Immer kürzere Vorwarnzeiten erfordern zunehmend den Einsatz von Techniken der KI. Da die verfügbaren Daten unsicher und unvollständig sind, können auch KI-Systeme in solchen Situationen nicht zuverlässig entscheiden.*

Rebellion oder Untergang

Der große amerikanische Wissenschaftler Noam Chomsky kommt zu dem Schluss, dass unsere Erde und unser Überleben als Menschheit in unserer langen Evolutionsgeschichte mehrfach von außen bedroht war, aber zum ersten Mal sind wir Menschen selbst in der Lage, diesen Planeten und unser Überleben als Spezies auszulöschen. Dabei sieht er den Klimawandel und den durch technisches Versagen ausgelösten Atomkrieg als größte Bedrohungen (Chomsky 2020).

Nukleare Abschreckung

Die Sicherung der atomaren Zweitschlagfähigkeit ist die Grundlage der Abschreckungsstrategie, die bis heute jeden potenziellen Angreifer abgehalten hat, einen atomaren Angriff zu starten: „Wer als erster schießt, stirbt als zweiter." Um auch bei einer Gefährdung der Zweitschlagfähigkeit reagieren zu können, haben die Atommächte computergestützte Frühwarn- und Entscheidungssysteme entwickelt und installiert, mit dem Ziel, einen Angriff rechtzeitig zu erkennen, um die eigenen atomaren Trägerraketen vor dem vernichtenden Einschlag aktivieren zu können. Eine solche Strategie wird als *launch-on-warning*-Strategie bezeichnet.

Wesentliche Komponenten eines Raketenwarnsystems sind:

- Sensoren zur Feststellung eines atomaren Angriffs,
- Computerzentren und Kommunikationsnetzwerke zur Analyse und Übermittlung von Daten,
- Kommandostellen zur Bewertung von Warninformationen und der Gefährdungslage sowie zur Planung und Anordnung von Gegenaktionen.

Wesentliche Grundlage für das Erkennen von Raketenstarts sind Satelliten, die auch über Sensoren verfügen, um nukleare Explosionen zu entdecken. Zur Beobachtung und Berechnung von Flugbahnen sind Radarstationen gebaut worden, und über die strategisch wichtigen Bereiche der Weltmeere sind Horchsensoren verteilt, die die Bewegungen von U-Booten erfassen. Die Daten von Sensoren werden in hochkomplexen redundant ausgelegten Computer-Netzwerken verarbeitet (siehe z. B. Bläsius und Siekmann 1987; Schlosser 2013).

Fehlalarme

In hochkomplexen Systemen treten Fehler auf, und es ist unmöglich ein solches System fehlerfrei zu realisieren. Fehler in Frühwarnsystemen können aber bedeuten, dass ein Angriff mit atomaren Raketen gemeldet wird, obwohl keine Bedrohung vorliegt, und ein solcher Fehler kann zu gefährlichen Situationen und unter Umständen sogar zu einem Atomkrieg führen. Ursachen für solche Fehler können z. B. falsche oder falsch interpretierte Sensordaten, Übertragungsfehler oder Computerfehler sein.

Fehler in Frühwarnsystemen werden üblicherweise nicht bekannt. Aufgrund von vielen Berichten über Fehlalarme wurden Anfang der 1980er Jahren Untersuchungen zu solchen Fehlern durchgeführt. In einem Bericht von Hart und Goldwater an den Senat sind für den Zeitraum 1. Januar 1979 bis 30. Juni 1980 insgesamt 147 Fälle von Anzeichen einer Bedrohung der USA bekannt geworden, die zur Auslösung der Alarmstufe 1 (*Missile Display Conference*) von drei möglichen Alarmstufen führten. Folgende fünf Fälle führten zur zweiten Alarmstufe (*Threat Assessment Conference*):

- **3. Oktober 1979:** Ein Radar, zuständig für das Erfassen U-Boot-gestützter Raketen, entdeckte einen Raketenkörper auf niedriger Umlaufbahn und verursachte einen falschen Alarm und eine Treffermeldung.
- **9. November 1979:** Es wurde ein Massenangriff durch Atomraketen gemeldet; dessen Ursache war ein Simulationsprogramm zum Testen von Systemkomponenten, das im Raketenwarnsystem von NORAD aktiviert wurde, ohne das Bedienungspersonal hierüber zu informieren.
- **15. März 1980:** Im Rahmen sowjetischer Übungen wurden vier Raketen von U-Booten aus gestartet. Eine dieser Raketen entwickelte eine Flugbahn, die ein Ziel in den USA zu ergeben schien.
- **3. Juni 1980 und 6. Juni 1980:** Es wird ein Massenangriff mit Raketen auf die USA gemeldet. Grund war ein defekter Chip in einer Kommunikationseinheit, die permanent Daten sendete, wobei an bestimmten Stellen im Normalfall Nullen stehen müssen. Durch den Hardwarefehler wurden an diesen Stellen andere Werte gesendet und damit angreifende Raketen gemeldet.

Auch während der Kuba-Krise gab es einige sehr kritische Situation, wie z. B. am 27. Oktober 1962: Ein russisches U-Boot,

das sich vor Kuba in internationalen Gewässern befand, wurde von der amerikanischen Marine eingekesselt und attackiert. Die Amerikaner wollten es zum Auftauchen zwingen. Aufgrund der Attacken glaubte die russische Besatzung, der Krieg sei bereits ausgebrochen, und musste über den Einsatz der Atomwaffe an Bord entscheiden. Der Kapitän des U-Boots hielt die Situation des U-Boots und der Besatzung für aussichtslos und entschied, den nuklearen Torpedo abzuschießen. Der Torpedo-Offizier stimmte dem Abschuss zu. Für die Entscheidung über den Atomwaffeneinsatz waren auf diesem Boot drei Offiziere zuständig, da hier auch der Flottenkommandant anwesend war. Nur wenn alle drei zustimmten, war ein Einsatz zulässig. Der dritte Offizier, Wassili Archipow, verweigerte die Zustimmung für den Abschuss und verhinderte damit möglicherweise einen atomaren Krieg.

Besonders bekannt geworden ist ein Vorfall vom 26. September 1983: Ein Satellit des russischen Frühwarnsystems meldet fünf angreifende Interkontinentalraketen. Da die korrekte Funktion des Satelliten festgestellt wurde, hätte der diensthabende russische Offizier Stanislaw Petrow nach Vorschrift die Warnmeldung weitergeben müssen. Er hielt einen Angriff der Amerikaner mit nur fünf Raketen aber für unwahrscheinlich und entschied trotz der Datenlage, dass es vermutlich ein Fehlalarm sei und verhinderte damit eine Katastrophe mit atomarem Schlag und Gegenschlag. Der Vorfall ereignete sich während einer instabilen politischen Lage: Die Nachrüstung durch Mittelstreckenraketen stand an und wenige Wochen vorher hatten die Sowjets aus Versehen eine koreanische Passagiermaschine über internationalen Gewässern abgeschossen. Möglicherweise hätte eine Maschine aufgrund der Fakten den Angriff eher als echt eingeschätzt und Gegenreaktionen eingeleitet. Petrow hatte gefühlsmäßig auf einen Fehlalarm gehofft, wollte nicht für den millionenfachen Tod von Menschen verantwortlich sein und hat sich entsprechend entschieden.

Risiken steigen

Es ist zu erwarten, dass das Risiko eines Atomkriegs in den nächsten Jahren und Jahrzehnten stark steigen wird. Der Klimawandel wird zu mehr Krisen führen, und neue technische Entwicklungen werden die Komplexität von Frühwarnsystemen und Bedrohungssituationen so stark erhöhen, dass die Beherrschbarkeit solcher Systeme immer schwieriger wird.

In den letzten Jahren hat ein neues Wettrüsten in verschiedenen militärischen Dimensionen begonnen. Die meisten dieser Entwicklungen sind noch am Anfang und die Folgen kaum kalkulierbar. Dies gilt für neue Trägersysteme von Atomwaffen, wie etwa die Hyperschallraketen, die geplante Bewaffnung des Weltraums, den Ausbau von Cyberkriegskapazitäten und die zunehmende Anwendung von Systemen der Künstlichen Intelligenz (KI) bis hin zu autonomen Waffensystemen. Besonders gefährlich könnten navigierbare Marschflugkörper sein, die nur schwer erkennbar sind, aber eventuell eine große Reichweite haben, z. B. mit einem Nuklearantrieb. Alle diese Aspekte haben auch Wechselwirkungen mit Frühwarnsystemen zur Erkennung von Angriffen mit Atomraketen und werden die Komplexität dieser Systeme deutlich erhöhen.

Die Weiterentwicklung von Waffensystemen mit höherer Treffsicherheit und immer kürzeren Vorwarnzeiten wird zunehmend Techniken der Künstlichen Intelligenz erforderlich machen, um für gewisse Teilaufgaben Entscheidungen automatisch zu treffen. Es gibt im Zusammenhang mit Frühwarnsystemen bereits Forderungen, autonome KI-Systeme zu entwickeln, die vollautomatisch eine Alarmmeldung bewerten und gegebenenfalls einen Gegenschlag auslösen, da für menschliche Entscheidungen keine Zeit mehr bleibt.

Automatische Entscheidungen

In manchen Situationen könnten KI-Entscheidungen auch hilfreich sein und zu besseren Resultaten führen als menschliche Entscheidungen. Vielleicht hätte 2020 im Iran ein versehentlicher Abschuss einer Passagiermaschine mit Hilfe von KI verhindert werden können. Im Januar 2020 hatten die USA den iranischen General Soleimani mit einem Drohnenangriff getötet. Als Vergeltung hat Iran wenige Tage später amerikanische Stellungen im Irak angegriffen. Kurz danach wurde im Iran ein ukrainisches Verkehrsflugzeug aus Versehen abgeschossen. Die Bedienungsmannschaft kam zu dem Ergebnis, dass es sich bei dem Flugobjekt um einen angreifenden Marschflugkörper handeln könnte. Die Fehlentscheidung kam vor allem dadurch zu Stande, dass die Bedienungsmannschaft mit Krieg oder einem Angriff der USA gerechnet hatte. In dieser Situation hätte eine Maschine möglicherweise besser entschieden. Denn die reinen Fakten, wie die Größe des Radarsignals, hätten vermutlich gegen einen Marschflugkörper gesprochen. Vielleicht hätte eine Maschine in der Kürze der Zeit auch mehr Informationen, wie z. B. Flugpläne, berücksichtigen können. Die Bedienungsmannschaft hatte den politischen Kontext vermutlich überbewertet.

Vagheit, Unsicherheit und Unvollständigkeit

In Zusammenhang mit Atomwaffen könnten solche automatischen Entscheidungen aber fatal sein. Die im Falle einer Alarmmeldung für eine Entscheidung verfügbaren Daten sind in der Regel vage, unsicher und unvollständig. Bei der Bewertung von Sensorsignalen spielen vage Werte wie Helligkeit und Größe eine Rolle, wobei es ein kontinuierliches Spektrum zwischen „trifft nicht zu“ und „trifft zu“ geben kann. Signale werden auch nicht immer auftreten, können also unvollständig sein. Dies kann insbesondere für neue lenkbare Raketensysteme gelten, die einer Erfassung ausweichen können. Des Weiteren sind für die elektronische Kampfführung Störsysteme wie *Kalaetron Attack* entwickelt worden, die es ermöglichen sollen, eine Erkennung durch die gegnerische Flugabwehr abzuwehren (Behördenspiegel 2020: 45). Im Falle einer Angriffsmeldung kann also nicht sichergestellt werden, dass die Daten auf Basis mehrerer unabhängiger Signalquellen überprüft werden können.

Deshalb können auch KI-Systeme in solchen Situationen nicht zuverlässig entscheiden. In der kurzen verfügbaren Zeit wird es kaum möglich sein, Entscheidungen der Maschine zu überprüfen. Dem Menschen bleibt nur zu glauben, was die Maschine liefert. Aufgrund der unsicheren und unvollständigen Datengrundlage werden weder Menschen noch Maschinen in der Lage sein, Alarmmeldungen zuverlässig zu bewerten. Solche Unsicherhei-

ten können auch bei normalen Waffensystemen relevant sein, allerdings sind die Auswirkungen in der Regel begrenzt, während es im Falle eines Atomkriegs um das Überleben der gesamten Menschheit gehen kann.

Internationale Krisen

Auch wenn die nukleare Abschreckung einen bewussten Atomwaffeneinsatz bisher verhindert hat, gibt es keine Garantie, dass dies immer so bleibt. Die Abschreckungsstrategie schützt nicht vor einem *Atomkrieg aus Versehen*. Das Wissen um die gravierenden Auswirkungen eines Atomkriegs bildet auch in Krisen- und Kriegszeiten eine große Hemmschwelle für den Einsatz von Atomwaffen. Dennoch sind verschiedene Szenarien denkbar, in denen es zu einem Einsatz kommen kann:

- Bewusster Einsatz von Atomwaffen: Eine Seite setzt Atomwaffen ein, um einen Vorteil zu erzielen, ein bestimmtes Ziel zu erreichen oder Vergeltung zu üben.
- Atomkrieg aus Versehen: Aufgrund eines Fehlalarms in einem Frühwarnsystem für nukleare Bedrohungen kommt es durch Fehleinschätzungen zu einem Atomkrieg.
- Kombination von bewusstem und versehentlichem Atomkrieg: Ein Fehlalarm in einem Frühwarnsystem könnte als Anlass für einen nuklearen Angriff gewählt werden, wenn ein solcher ohnehin schon in Erwägung gezogen wurde. Die Aspekte 1 und 2 könnten sich entscheidend verstärken.

Seit dem Krieg in der Ukraine ist auch das Risiko eines möglichen Atomkriegs in der Diskussion, und es wurde sogar mit dem Einsatz von Atomwaffen gedroht. Auch dadurch erhöhen sich die Risiken, denn im Falle einer Alarmmeldung könnte diese, z. B. aufgrund vorheriger Drohungen, als echt eingeschätzt werden.

Maßnahmen zur Reduzierung der Risiken

Die Atomkriegsrisiken sind in den letzten Jahren auch deshalb gestiegen, weil wichtige Vereinbarungen wie INF und *open skies* gekündigt wurden. Zur Reduzierung der Risiken wären dringend Maßnahmen und neue Vereinbarungen erforderlich, wie zum Beispiel:

- Verbesserung von Vertrauen, Kommunikation und Zusammenarbeit zwischen Atommächten,
- Vereinbarungen zur Reduzierung von Atomwaffen und zum Alarmmodus von Atomwaffen,
- Verbesserung des Informationsaustauschs in Zusammenhang mit Frühwarnsystemen,
- Keine automatischen Entscheidungen zum Einsatz von Atomwaffen.

Referenzen

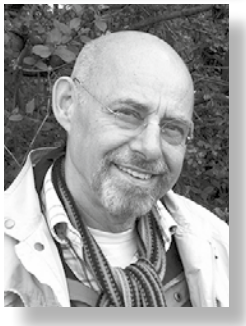
Behördenspiegel (Mai 2020): https://issuu.com/behoerden_spiegel/docs/2020_mai

Bläsius, Karl Hans und Jörg Siekmann (1987): Computergestützte Frühwarn- und Entscheidungssysteme. Informatik-Spektrum, Band 10, Heft 1, 24-39

Bläsius, Karl Hans und Jörg Siekmann (2022): KI in Frühwarnsystemen für nukleare Bedrohungen. In: Bläsius, Karl Hans/ Reiner Schwalb/ Michael Staack (Hrsg.): Künstliche Intelligenz und nukleare Bedrohungen – Risiken eines Atomkriegs aus Versehen, WIFIS-aktuell Band 73, Verlag Barbara Budrich

Chomsky, Noam (2020): Internationalism or Extinction, Routledge (deutsche Ausgabe, 2021: Rebellion oder Untergang!, Westend Verlag)

Schlosser, Eric (2013): Command and Control, Verlag C.H.Beck



Karl Hans Bläsius und Jörg Siekmann

Karl Hans Bläsius war bis 2017 Professor an der Hochschule Trier, Fachbereich Informatik. Fachgebiete: Logik, Funktionale Programmierung, Dokumentanalyse, Künstliche Intelligenz. Er beschäftigt sich mit Frühwarn- und Entscheidungssystemen und *Atomkrieg aus Versehen* seit 1983 und ist Mitinitiator der Seite *www.atomkrieg-aus-versehen.de*.

Jörg Siekmann, geb. 1941, war von 1991 bis 2006 Professor für Informatik und Künstliche Intelligenz an der Universität des Saarlandes und ist seitdem dort Seniorprofessor. Er promovierte 1976 in Artificial Intelligence an der University of Essex und wurde 1983 auf die erste deutsche Professor für Informatik und Künstliche Intelligenz an der Technischen Universität Kaiserslautern berufen. Er war maßgeblich beteiligt am Aufbau der KI-Forschung in Deutschland, ist Gründer und erster Sprecher der KI-Fachgruppe in der Deutschen Gesellschaft für Informatik (GI) und war Sprecher des SFB-378 Ressourcenadaptive kognitive Prozesse. Von 1991 bis 2006 war er Direktor des 1989 von ihm mitgegründeten Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI) und war Koordinator der Universität des Saarlandes für Digitale Bildung. Er wurde 2019 von der GI zu einem der zehn einflussreichsten KI-Forscher gewählt.

Christian Heck

Kritik des gläsernen Gefechtsfeldes[1]

Was Sprachmodelle und Massendaten im Krieg bedeuten

Für die Kriegsführung 4.0 ist das gläserne Gefechtsfeld ausschlaggebend. Doch das Internet of Military Things *(IoMT) und* Battle Management Systeme *sind nicht nur militärisch, sondern auch aufgrund ihrer Operationslogik hochgradig kritikwürdige Instrumente. Der Trend zu immer mehr Komponenten des Maschinellen Lernens, die in diese Systeme implementiert werden, scheint derzeit unaufhaltbar. Doch KI ist entgegen der öffentlichen Meinung keine* Blackbox. *Sie besteht aus vielen* Whiteboxes*, in die wir hineinsehen können. Einzig sie zu erschließen, um ihre inneren Funktionsweisen zur maschinellen Bedeutungsgenerierung verstehen zu lernen, dazu sind wir noch nicht in der Lage. So gilt es, die grundsätzlichen Prämissen dieser Systeme adäquat zu kritisieren. Insbesondere die Bedeutung, die ihnen mittlerweile für kriegerisches Handeln zugemessen wird, muss umso mehr Anlass für Kritik sein, die in diesem Beitrag ausgeführt wird.*

Die letzte sogenannte *Revolution in Military Affairs*, die vernetzte Kriegsführung (*Network Centric Warfare* NCW) der späten 1990er-Jahre, in der technologische Innovationen gezielt zur Vernetzung vieler Informations- und Aufklärungssysteme miteinander gestaltet wurden, wurde im 2. Golfkrieg erstmals durch die US-Streitkräfte als Testfeld unter Beweis gestellt. Es entwickelte sich in der Folge eine systemische Denkweise von Militärtheoretiker:innen, dass durch ein integratives Verständnis von Einzelkomponenten exponentielle Leistungssteigerungen des Gesamtsystems in seiner militärischen Wirksamkeit erreicht werden können. Dieses frühe *Internet der Dinge* (IoT) verwob sich zunehmend mit dem Internet, das wir heute kennen: mit sozialen Netzwerken, IT-Monopolen und Clouds, die riesige Rechenzentren weltweit zu Big Data durch Deep Learning[2] und weitere neue technische kognitive Systeme verrechnen. Ein Auftakt ins neue Jahrtausend.

Ein Jahrtausend hybrider Konflikte und *gläserner Gefechtsfelder* mit Kampfhandlungen in unseren Städten, im Netz sowie auch im Weltraum. Das *Internet of Military Things* (IoMT) und die *Kriegsführung 4.0* (vgl. W&F 4/2019) entwickelten sich aus zahlreichen Spin-in- und Spin-off-Effekten, d. h. Innovationen aus Industrie, Wirtschaft und Gesellschaft, die vom Militär übernommen wurden und umgekehrt. Militärtechniken fanden auf diese Weise Eingang in unsere zivilen Räume. In öffentliche und private Räume, in denen wir uns bewegen und miteinander sprechen bzw. chatten.

Eine technisch erzeugte Wirklichkeit

Um *Multi Domain Operationen* (MDO) auszuführen, das heißt sich innerhalb dieser neu ausgeschriebenen Gefechtsfelder zurechtzufinden, Objekte und Situationen zu erkennen, sie adäquat einzuschätzen um daraufhin beste Gewissensentscheidungen zu treffen, benötigen Soldatinnen und Soldaten in Operationszentralen technische kognitive Systeme, sogenannte *Battle Management Systeme* (BMS), zu deutsch: digitale Führungssysteme. Diese Systeme dienen in erster Linie der Entscheidungsunterstützung, da die notwendigen Erkenntnisse innerhalb dieser Art der High-Tech-Kriegsführung nur durch die Unterstützung durch Technologien erlangt werden können. Diese digitalen Führungssysteme versprechen Einsatzschnelligkeit – und in der Vision von MDOs soll dies heißen, nicht nur möglichst schnell, d. h. in Echtzeit, sondern gar seiner Zeit voraus zu handeln. Diese innerhalb dieser Gefechtsfelder durch und durch technische Zeiteinheit ist jedoch nicht für den Menschen gemacht, sondern zur möglichst fehlerfreien Funktionstüchtigkeit der technischen Systeme selbst. So auch die Interpretation anfallender Datenströme, die im gläsernen Gefechtsfeld in Echtzeit in militärisch-technische Handlungen überführt werden muss. Auch diese ist nicht für den Menschen gemacht und er wird die undenkbare Masse an Daten nicht allein bewältigen können, sei er noch so gut ausgebildet. Auch hierfür braucht es technische kognitive Systeme auf aktuellem Stand und ein ausgefeiltes *Man Machine Teaming* (MMT).

Dies bedeutet jedoch auch, dass im gläsernen Gefechtsfeld kognitive Technologien das meiste, das man wahrnehmen und erkennen kann, auch erst herstellen, dass das gläserne Gefechtsfeld also eine technisch erzeugte Wirklichkeit ist. Unter anderem deshalb ist es in MDOs zwingend notwendig, zur eigenen technischen Handlung während der militärischen Operationen Distanz einzunehmen bzw. einnehmen zu können. Dies erfordert, Trennlinien zu setzen zwischen technischer und menschlicher kognitiver Leistung.

Da es bei militärischen Operationen fast immer um Leben und Tod geht, muss ethisch die letzte Entscheidung beim Menschen liegen. Zugleich müssen Führungskräfte, Beamt:innen und politische Entscheidungsträger:innen aber auch anerkennen, dass diese Entscheidungsfindung ohne technische kognitive Systeme nicht realisierbar ist. Sie müssen hierfür also mehr eine innere als eine analytische oder formale Grenze setzen. Eine Grenze zwischen dem Gewissen und der Entscheidung, basierend auf technischer Kognition.

Denn die spezielle Art dieser technischen Systeme räumt menschlichen Akteuren und nicht-menschlichen Artefakten eine gänzlich neue aktive politische Handlungskraft ein. Auch jenen, die in früheren Operationspraxen eine unwichtige, bisweilen gar keine, zumindest aber eine andere Rolle spielten: Dazu zählen unter anderem Betreiber:innen von Cloudplattformen, Rechenzentren und Satellitenanlagen, kriegspropagandistische Influencer:innen oder eben BMS.

Die Integration von großen Sprachmodellen

Erste Komponenten für MD-Operationen wurden bereits in den frühen 1980er Jahren in den Militärapparat implementiert. Hans-Jörg Kreowski erinnerte in seinem Artikel *Die militärische Seite der Digitalisierung* (Kreowski 2023) bspw. an die *Strategic*

Computing Initiative (SCI) aus dem Jahr 1983, die das US-Verteidigungsministerium bereits Jahrzehnte vor der Notwendigkeit von Multi-Domain-Operationen in gläsernen Gefechtsfeldern startete. Mit SCI sollten KI-Projekte entwickelt werden, bei denen neben dem Design autonomer Landfahrzeuge auch die Konzeption eines frühen BMS und eines Sprachassistenten für die Pilot:innen der Luftwaffe zur Aufgabe standen.

Die jüngste Generation von Software-Produkten, die aus diesem Ansatz heraus entwickelt wurden, ist im April diesen Jahres auf dem Markt erschienen. Im September 2023 unterzeichnete das erste Rüstungsunternehmen einen Vertrag mit dem Hersteller, dem US-amerikanischen Datenanalyse-Unternehmen *Palantir* (vgl. Palantir Technologies 2023a). Mit ihrer digitalen Plattform *AIP for Defense* (Palantir Technologies 2023b) werden große vortrainierte Sprachmodelle mit künstlichen neuronalen Einbettungen (LLMs) für militärische Operationen nutzbar gemacht.

Das Unternehmen selbst wurde 2004 gegründet und nahm in diesem Jahrtausend u. a. im „Krieg gegen den Terror" eine nicht zu unterschätzende Rolle ein. Es spezialisierte sich ziemlich schnell auf die Überwachung von Individuen und die Zusammenführung eigentlich getrennter Datenbestände und wurde somit ein wichtiger Akteur in hybriden und asymmetrischen Kriegsführungsstrategien.

Ihre Datenbankvisualisierungs- und Analyse-Software *Gotham* wurde unter anderem zuerst von der *Joint Improvised-Threat Defeat Organization* (JIDO) getestet, einer Einheit des US-Verteidigungsministeriums (DoD), die eingerichtet wurde, um einem neuen Phänomen in dieser Art des Krieges entgegenzuwirken: dem von Anschlägen mit improvisierten Sprengsätzen (*Improvised Explosive Devices* (IED)), mit Autobomben, Paketbomben, Selbstmordattentaten etc. Die CIA, die NSA und das FBI wurden ziemlich schnell zu Kunden von Palantir. Auch die Europäische Polizeibehörde Europol nutzt inzwischen Produkte von Palantir für die Datenauswertung. In der Bundesrepublik wird Palantirs *Gotham* u. a. in Bayern als *Verfahrensübergreifende Recherche- und Analyseplattform* (Vera) eingesetzt. Auch die Polizei in NRW hat Software von Palantir in Betrieb, mit der *Datenbanküber-greifenden Analyse- und Recherche-Software* (DAR). In Hessen wurde *Gotham* in dem System *HessenData* seit 2017 eingesetzt. Die Landesregierung in Hessen hatte die Anschaffung der Software freigegeben, doch der Einsatz wurde im Februar 2023 vom Bundesverfassungsgericht in Karlsruhe als verfassungswidrig eingestuft. Soweit bekannt, führt das System Daten aus sozialen Medien mit Einträgen in verschiedenen polizeilichen Datenbanken sowie Verbindungsdaten aus der Telefonüberwachung zusammen, um mögliche Straftäter:innen zu ermitteln. Zudem spielte Palantir eine nicht unerhebliche Rolle im Skandal um *Cambridge Analytica*. Das Unternehmen soll Facebook bei der Auswertung der illegal weitergegebenen Daten geholfen haben.

Palantirs neuestes Softwarepaket *AIP for Defense* wird nun in naher Zukunft Einsätze unterstützen, indem es u. a. feindliche Stellungen erkennt und durch eine Chatfunktion (ähnlich dem Interface der international viel diskutierten KI *ChatGPT*) Gegenmaßnahmen vorschlägt und gegebenenfalls autonom ausführt – wie z. B. das Starten einer Aufklärungsdrohne ins Zielgebiet. Doch nicht nur das Interface mit integrierter Chatfunktion erinnert an ChatGPT, auch das maschinelle Lernverfahren in AIP ist ein ähnliches wie bei OpenAIs Künstlicher Intelligenz hinter ChatGPT: das generative Sprachmodell GPT-3 (Generative Pretrained Transformer 3).

GPT-3 ist ein LLM, dessen 175 Milliarden Parameter auf Clouds trainiert werden, die über viele Rechenzentren verteilt sind und dessen Entwicklung derzeit an physische und auch ökologisch tragbare Grenzen stößt. Der Stromverbrauch für das Training entspricht dem von 3.000 europäischen Durchschnittshaushalten, eine Frage an ChatGPT benötigt 1.000 Mal mehr Strom als eine Suchanfrage bei Google und für jede Antwort, die man von dem Bot erhält, könnte man ein Smartphone bis zu 60 Mal aufladen. Vermutlich ist *AIP for Defense* eines der ersten *Battle-Management-Systeme*, die LLMs implementiert haben.

Wie maschinelle Bedeutung generiert wird

Bisher hatten LLMs wie eben GPT-3 bzw. GPT-4, LaMDA und PaLM von Google oder LlaMA von Meta in digitalen Führungssystemen eine eher kleine bis gar keine Rolle gespielt. Etablierteren Ansätzen, wenn auch nicht zwingend minder experimentellen bei der komputativen Verarbeitung natürlicher Sprache (NLP[3]), kann auf der anderen Seite schon seit Jahren eine Art Schlüsselrolle zugeschrieben werden.

Der Sozialpsychologe und Sozialwissenschaftler James W. Pennebaker fasste eines der Hauptargumente hierfür wie folgt zusammen: *„Die Worte, die wir im täglichen Leben verwenden, spiegeln wider, worauf wir achten, woran wir denken, was wir zu vermeiden versuchen, wie wir uns fühlen und wie wir unsere Welt organisieren und analysieren."* (Tausczik und Pennebaker 2010:25)

Die Entwicklung von *Word embeddings*[4], zu deutsch: Worteinbettungen, eröffnete hierfür einen gänzlich neuen Handlungsspielraum in der komputativen Sprach- und Netzwerkanalyse.

Zur Extraktion inhaltlicher Strukturen, Features, Organisationseinheiten etc. werden aus einer Vielzahl von generierten semantischen Beziehungen zwischen Wörtern und Sätzen symbolische Repräsentationen in Sprachmodellen mit künstlichen neuronalen Einbettungen errechnet. Eingebettet in digitale Führungssysteme, kommen die Wörter und Sätze, die es zu berechnen gilt, u. a. aus *Open-Source-Intelligence*-Datensätzen (Medienberichten, Social Media, wissenschaftlichen Arbeiten, öffentlich zugänglichen Statistiken etc.), aus Berichten von Geheimdiensten und des militärischen Nachrichtenwesens, Analysen von Sicherheitsbehörden, Erkenntnissen aus der signalerfassenden Aufklärung, der Bild- und Satellitenaufklärung, von Bots, Drohnen und anderen technischen kognitiven Systemen bzw. unbemenschten Fahrzeugen. Die „Lern"-Kriterien, nach denen maschinell Bedeutungen generiert werden, liegen in diesen Sprachmodellen verankert. Nach ihnen werden die Millionen von Gewichtungen, die an den einzelnen künstlichen Neuronen liegen, eingestellt. Eine undenkbare Masse an Informationen aus den unterschiedlichsten Quellen wird für diesen Prozess gesammelt. Sie wird übersetzt, selektiert, kategorisiert und mit strukturierten Daten angereichert, sprich, sie wird enkodiert. Sie wird maschinenlesbar gemacht (*machine readable*), um sie dann nach ihrer Verarbeitung wieder für den Menschen

lesbar zu machen (*human readable*). Es findet auf diese Weise eine maschinelle Vorinterpretation statt. Eine Menschenlesbarmachung von maschinell erlernten symbolischen Repräsentationen, nach vorgegebenen Regeln, die im Code verankert sind.

Als im Jahr 2013 das vortrainierte Sprachmodell *Word2vec* von einem Google-Forscherteam veröffentlicht wurde, galt dies als ein Durchbruch in den NLP-Forschungsgemeinden. Es wurde recht zügig in Technologien der inneren und äußeren Sicherheit implementiert und ist bis heute noch eines der gebräuchlichsten Modelle zur Generierung von Worteinbettungen mittels Deep Learning. Die von Word2vec erzeugten Worteinbettungen können einfach und bequem zur Weiterverarbeitung verwendet werden und zugleich konnte das KI-Modell Ergebnisse auf dem neuesten Stand der Technik (ca. 2013-15) liefern.

Die inneren Funktionsweisen vortrainierter Sprachmodelle mit künstlich neuronalen Einbettungen, egal ob LLMs oder Word-2vec, beruhen auf der Idee, dass im Gegensatz zur formalen Linguistik und zur Chomskyschen Tradition allein die Kontextinformation eine brauchbare Darstellung sprachlicher Elemente darstellt. Je nach Modell werden hier einzelne Wörter eines Korpus (Textes) als symbolische Repräsentationen in einem semantischen Vektorenraum (Wortraum) mit etwa 300 Dimensionen dargestellt. Zum Vergleich: in heutigen Transformer-Architekturen à la GPT-3 werden 1.536 Dimensionen pro Wort errechnet. Worteinbettungen repräsentieren also den innertextlichen Kontext eines Datensatzes, in dem das jeweilige Wort vorkommt.

Niemanden interessiert, wie es funktioniert, solange es funktioniert

Das Studium aktueller NLP-Forschungsarbeiten zeigt, dass trotz der weit verbreiteten Anwendung von künstlich neuronalen Worteinbettungen noch immer erstaunlich wenig über die Struktur und die Eigenschaften – und folglich auch die Konsequenzen – dieser Einbettungsräume bekannt ist. Dennoch werden zunehmende Teile von Welt durch sie in formale, computerlinguistisch verarbeitbare Informationen und Beschreibungsebenen (Morphologie, Syntax, Semantik, Aspekte der Pragmatik etc.), das heißt in symbolische Repräsentationen umgewandelt.

Auf diese Weise fließen aber auch immer wieder auftauchende, bisweilen diskriminierende Tendenzen bis hin zu Rassismen in digitale Führungssysteme mit ein. Sogenannte „Verzerrungen", die diesen textanalytischen Machine-Learning-Verfahren zwar nicht explizit, auch nicht vorsätzlich eingeschrieben werden, die aber dennoch im realweltlichen Gebrauch in Erscheinung treten. Denn trotz dieser bekannten und in den letzten 3-4 Jahren auch in der Öffentlichkeit verhandelten Defizite von KI-Sprachmodellen implementieren Firmen diese weiterhin in ihre Produkte und verkaufen sie an Sicherheitsbehörden und an das Militär.

Der Trend hin zu immer größer werdenden Modellen und immer mehr (unüberprüften) Trainingsdaten in den letzten Jahren führt auch zu einer immer schlechter werdenden Kontrollierbarkeit der inneren Funktionsweisen von technischen kognitiven Systemen. Funktionsweisen, durch die Minderheiten diskriminiert und gesellschaftliche Gruppen marginalisiert werden, und das auf eine Weise, die meist den Entwickler:innen selbst nicht bekannt bzw. bewusst, schlimmstenfalls egal ist (vgl. Bender et al. 2021).

Auch aus diesem Grund ist es für Soldatinnen und Soldaten in Operationszentralen oder in Ämtern und Firmen, die die analytische Ausarbeitung für Führungskräfte unterstützen, unabdingbar, sich ein grundlegendes Verständnis des inneren Aufbaus dieser technischen kognitiven Systeme anzueignen. Denn die Bedeutung der jeweiligen militärisch-technischen Handlung muss aus diesen Systemen heraus, also in deren inhärenter Pragmatik, erschlossen werden. Aus technischen kognitiven Systemen, die neben ihrer natürlichen Begrenztheit zugleich auch eine scheinbare Grenzenlosigkeit zu Tage fördern.

Eine Grenzenlosigkeit, die in ihrem Lern-Vermögen liegt, die uns derzeit auch bis zur Entgrenzung unseres menschlichen Denkens führt. Denn diese technischen Systeme können *„sowohl lernen, dass die Erde flach ist, als auch rund"* (Chomsky et al. 2023), so Noam Chomsky (linker Intellektueller, Anarchist und emeritierter Professor für Linguistik am Massachusetts Institute of Technology, MIT) im März 2023 in der *New York Times*. Sie können auch lernen, dass seit weit über einem Jahr ein von Russland geführter Angriffskrieg gegen die ukrainische Bevölkerung wütet. Im nächsten Moment können sie dies jedoch auch wieder verlernen, egal ob es der Wirklichkeit entspricht oder nicht. Was diese technischen kognitiven Systeme eben nicht können, ist, ihre innere Grenzenlosigkeit durch ethische Prinzipien einzuschränken. Eine Eigenschaft, die wir als „moralisches Denken" bezeichnen. Die technischen kognitiven Systeme sind, während sie prozessieren, getrennt von der Außenwelt. In ihren inneren Entscheidungsfunktionen liegen keine Modellierungen von dem, was Worte, was Dinge, was Taten und Ereignisse für uns in der Welt bedeuten. Dennoch werden sie über Leben und Tod von Soldatinnen und Soldaten, von Jugendlichen, von Großeltern, von Eltern und unseren Kindern algorithmisch mitentscheiden.

Christian Heck

Christian Heck ist künstlerisch-wissenschaftlicher Mitarbeiter für Ästhetik & neue Technologien / Experimentelle Informatik und Doktorand an der Kunsthochschule für Medien Köln. Seine Forschungs- und Arbeitsschwerpunkte liegen auf Friedensforschung, Ästhetische Praxis und Ethik der Künstlichen Intelligenz mit Fokus auf Generative Systeme, ADM, IT-Sicherheitstechnologien, Kampfdrohnen und autonome Waffensysteme. Er ist Mitglied im Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FIfF) e. V. und der Gesellschaft für Informatik (GI).

Referenzen

Bender EM, Gebru T, McMillan-Major A and Shmitchell S (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, S. 610-623.

Chomsky N, Roberts I, Watumull J (2023) The False Promise of ChatGPT. New York Times, 8.3.2023.

Kreowski HJ (2023) Die militärische Seite der Digitalisierung. IMI-Ausdruck 113, S. 25-27.

Mikolov T, Sutskever I, Chen K, Corrado G, Dean J (2013) Distributed Representations of Words and Phrases and their Compositionality. NIPS'13: Proceedings of the 26th International Conference on Neural Information Processing Systems – Volume 2, S. 3111–3119.

Palantir Technologies (2023a) Palantir Technologies Signs Partnership With Titan Defence Firm, Babcock. Pressemitteilung, 13.9.2023.

Palantir Technologies (2023b) AIP for Defense. Homepage, URL: palantir.com/aip/defense/.

Tausczik YR, Pennebaker JW (2010) The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. Journal of Language and Social Psychology 29 (1), S. 24-54.

Anmerkungen

1 *Nachdruck aus dem Jubiläumsheft 4/2023 der Zeitschrift Wissenschaft und Frieden mit freundlicher Genehmigung der W&F-Redaktion und des Autors.*

2 *Die Technologie des „Deep Learning" begann sich um die Jahrtausendwende zu entfalten. Es begann bald darauf die Zeit von Big Data (dem Anstieg der Datenmengen durch die Verbreitung der Internettechnologien), und es wurden erhebliche Fortschritte in den Computertechnologien (in der Rechenkapazität, GPUs und preiswerten Speichertechnologien) erzielt. Erst durch diese technische Infrastruktur wurde die Weiterentwicklung der Künstlichen Neuronalen Netze (KNN) hin zum Deep Learning im Forschungs- und vermehrt auch im Anwendungsbereich möglich. Von dieser Technologie sprechen wir heute in erster Linie, wenn von Künstlicher Intelligenz zu hören ist: der subsymbolischen Künstlichen Intelligenz.*

3 *Das Kürzel NLP steht für Natural Language Processing. Eine Mischwissenschaft, die anteilig aus der Computerlinguistik, den Computerwissenschaften und der Künstliche Intelligenz Forschung besteht. Sie ist eine Wissenschaft der algorithmischen Verarbeitung von Sprache, der Verarbeitung von Daten und des künstlich intelligenten Verhaltens zugleich.*

4 *Der Sammelbegriff „Word embeddings" steht für eine Reihe von Sprachmodellierungs- und Feature-Learning-Techniken in NLP, bei denen Wörter oder Phrasen aus dem Vokabular auf Vektoren mit reellen Zahlen abgebildet werden: z. B. eine globale Korpusstatistik (GloVe: Globale Vektoren für Wortdarstellung) oder eine Wortkontextdarstellung (Word2vec) (siehe Mikolov et al 2013).*

Ryan R. Swan

Neubewertung einer unabhängigen Satellitenüberwachungseinheit

Ein Mechanismus, um Transparenz in Kooperation umzuwandeln?

Transparenz gilt seit langem als wesentlicher Baustein für Vertrauen und Kooperation zwischen Staaten (Glaser 1997; Keohane & Martin 1995). Der technologische Fortschritt treibt eine Informationsrevolution voran, die zu einer erhöhten Zugänglichkeit von Informationen führt (Hanson 2008). Diese Realität des sich ständig weiterentwickelnden Informationszeitalters wirft eine wichtige Frage für die Friedensforschung auf: Wie können diese Informationen in Transparenz auf eine solche Weise umgewandelt werden, um die zwischenstaatliche Zusammenarbeit zu fördern?

Einleitung

Dieser Beitrag legt nahe, dass die Wahrung von Objektivität ein zentraler Schlüssel ist. Obwohl die Explosion in in der Entwicklung der Open-Source-Intelligenz traditionelle Informationsmonopole der Großmächte erodiert hat, behalten diese Staaten immer noch weitgehend die Kontrolle über die Verwendung der von Privatunternehmen produzierten Satellitendaten und verhindern dadurch, dass diese zur Überprüfung voreingenommener staatlicher Narrative genutzt werden (Bennett et al. 2022; Zhao et al. 2021). In diesem Beitrag wird eine Möglichkeit vorgestellt, diese Daten zu objektivieren und dadurch ihr Potential zu maximieren, um vertrauensbildende Transparenz und letztendlich erhöhte Kooperationschancen zu verwirklichen: eine unabhängige Satellitenüberwachungseinheit (USE).

Die Vorstellung einer solchen Struktur ist nichts Neues. Die Vereinten Nationen prüften bereits Anfang der 1980er-Jahre die mögliche Umsetzbarkeit der Kernidee. Eine praktische Verwirklichung scheiterte jedoch am Widerstand der Großmächte und unerschwinglichen Vorlaufkosten. Dennoch besteht auch heute noch das Bedürfnis nach objektiver Transparenz. Die selektiven Narrative russischer und westlicher Seite im Zusammenhang mit dem Krieg in der Ukraine macht den Mangel an objektiven Informationsquellen deutlich. Da die Spannungen zwischen den Großmächten weiter eskalieren mögen und ihre Beziehungen in einen neuen, diesmal dreigliedrigen Kalten Krieg zunehmend auszuarten drohen, ist mit einer Verschärfung von Informationskriegstaktiken zu rechnen. Eine USE könnte eventuell dazu dienen, den schädlichen Folgen von Informationskriegen entgegenzuwirken und eine stärkere Rechenschaftspflicht für staatliche Akteure herzustellen.

Hintergrund

Im Frühjahr 1978 legte Frankreich den Vereinten Nationen ein Memorandum vor, in dem die Gründung einer Internationalen Satellitenüberwachungsagentur (International Satellite Monitoring Agency, ISMA) vorgeschlagen wurde. Das Memorandum

wies darauf hin, dass die im Bereich der Beobachtungssatelliten erzielten Fortschritte eine „neue Entwicklung in der Verwaltung internationaler Angelegenheiten“ darstellen und schlug vor, diese neuartigen Fähigkeiten in den Dienst der internationalen Gemeinschaft zu stellen (United Nations 1983: 1). 1979 wurde gemäß VN-Resolution 33/71 J eine Expertengruppe ernannt, um die technischen, rechtlichen und finanziellen Auswirkungen des ISMA-Vorschlags zu bewerten. Diese Gruppe „erkannte den Beitrag an, den die Überwachung durch Satelliten zur Überprüfung der Rüstungskontrolle und zur Verhütung und Beilegung von internationalen Konflikten leisten könnte, wodurch die Vertrauensbildung zwischen Staaten gefördert würde“ (United Nations 1983: 3).

Trotz anfänglicher Unterstützung von über 120 Nationen ist nichts Konkretes aus dem Vorschlag entstanden. Dieses Ergebnis resultierte aus zwei Hauptgründen. Erstens machten die beiden Supermächte deutlich, dass sie nicht bereit wären, sich auf die ISMA-Idee einzulassen und ihre Vorherrschaft im Bereich der Satellitenbeobachtung aufzugeben (Voute 1984). Zweitens wären die Vorabkosten für die Einrichtung der Agentur ohne finanzielle Unterstützung durch die Großmächte unerschwinglich hoch gewesen. Infolgedessen wurde der Vorschlag nicht weiterverfolgt und geriet allmählich in Vergessenheit. Obwohl internationale Einrichtungen wie das Büro der Vereinten Nationen für Weltraumfragen (United Nations Office for Outerspace Affairs, UNOOSA) gegründet wurden, um die friedliche Nutzung des Weltraums zu erleichtern, wurde die Idee einer USE nicht wieder ernsthaft geprüft. Angesichts der anhaltenden Relevanz von einer USE im gefährlichen aktuellen geopolitischen Umfeld und der technologischen Fortschritte, die effektivere Satellitenüberwachungsfähigkeiten ermöglichen, zusammen mit sinkenden Kosten ist eine erneute Betrachtung der USE-Idee angebracht.

Gegenwärtige Relevanz

Die sich verschlechternden Beziehungen zwischen den Großmächten führen heute zu einer Verschärfung der Rivalität und einem sich beschleunigenden militärischen Wettbewerb. Der Zusammenfluss dieser Entwicklungen führt zu einem internationalen Sicherheitsumfeld, das sehr gefährlich ist – vielleicht gefährlicher als je zuvor. Die konkurrierende Militarisierung einer Reihe aufkommender Technologien droht neue Eskalationswege zu schaffen – sowohl absichtliche als auch unbeabsichtigte – in einer Zeit, in der zunehmende Spannungen für Bedingungen sorgen, die einem Pulverfass ähneln. Unter solchen Umständen ist objektive Transparenz von hoher Bedeutung, gleichwohl aber schwer zu erreichen. Staatliche Konkurrenten sind bestrebt, Informationen zu ihrem eigenen einseitigen Vorteil auszunutzen und gleichzeitig voreingenommene Narrative zu zeichnen, die die Realität verdrehen. Dies schürt weiteres Misstrauen und treibt das Worst-Case-Szenario-Denken voran.

Mehr Informationen können jedoch auch zur gegenseitigen Überprüfung genutzt werden, was der Förderung der Kooperation durch die Verringerung der Unsicherheit und die Erhöhung der Rechenschaftspflicht dient. Dies mindert den Anreiz, Informationen zur Ausnutzung von anderen zu nutzen, da die Erfolgschancen geringer sind. Insbesondere in Zeiten des aktiven Wettbewerbs neigen Staaten dazu, die Ausnutzung vorzuziehen, um die Möglichkeiten einseitiger Gewinne auszuschöpfen, bevor sie sich der kooperativen Verifizierung zuwenden, wenn die Kosten und Gefahren der anhaltenden Konkurrenz zu hoch steigen. Eine USE könnte gefährliche Wettbewerbsspiralen abwenden, indem sie die Kultivierung von Informationsasymmetrien – einem Grundpfeiler des Wettbewerbs – von Anfang an erschwert. Dies könnte Anreize dafür setzen, früher die Positivsummenkooperation anzustreben.

Technologische Fortschritte und sinkende Kosten

Seit den 1990er-Jahren ist das Monopol der Großmächte im Bereich der Satellitenüberwachung stark zurückgegangen. Die Zahl kommerzieller Satelliten mit immer ausgefeilteren Beobachtungskapazitäten ist dagegen konsequent gestiegen (Moric 2023). Dies hat die unmittelbare Kontrolle durch Staaten über den Zugriff auf Satellitenbeobachtungsintelligenz geschwächt. Privatunternehmen wie Planet, Airbus und Maxar betreiben mittlerweile große Satellitenflotten mit hochauflösenden optischen Fähigkeiten. Darüber hinaus setzen diese Konzerne zunehmend Systeme mit synthetischem Aperturradar (*synthetic aperture radar*, SAR) ein, die eine fortgeschrittene radiowellenbasierte Beobachtung ohne Beeinträchtigung durch die Wolkendecke sowie *Radiofrequency Mapping* (RFM) zur Erkennung bodengestützter Radar- und Kommunikationssysteme ermöglichen (Moric 2023; Erwin 2023). Diese Ausweitung von kommerziellen Satellitenfähigkeiten, betrieben mittlerweile von Unternehmen in mehr als 30 Ländern, hat zu einem globalen Satellitenintelligenzmarkt im Wert von über 270 Milliarden US-Dollar geführt (Satellite Industry Association 2021). Bis vor Kurzem waren diese Informationen nur einigen Militärmächten zugänglich.

Ryan R. Swan

Ryan R. Swan ist Fellow beim Bonn International Centre for Conflict Studies und Stipendiat (Promotionsstipendium) der Gerda Henkel Stiftung im Sonderprogramm Sicherheit, Gesellschaft und Staat. Sein Forschungsschwerpunkt liegt auf neuen Technologien und deren Auswirkung auf Sicherheitsdynamiken. Er schloss ein Masterstudium in Politikwissenschaften an der University of Cambridge (Trinity Hall) und ein Jurastudium an der UCLA School of Law ab.

Außerdem ist die Größe von Satelliten dramatisch zurückgegangen. Gegenwärtige Geräte sind von der Größe eines Müllwagens bis auf die eines Toasters geschrumpft. Diese Satelliten, die mit Sensoren ausgestattet sind, die eine genauere Sichtbarkeit als das menschliche Auge erzielen, können eine nahezu lückenlose Abdeckung der Erdoberfläche ermöglichen, wenn sie mit SAR- und anderen wolkenunterdrückenden Systemen ausgerüstet sind und in großen Konstellationen mit fortschrittlichen KI-gestützten Informationsverarbeitungskapazitäten konzentriert sind (Moric 2022). Darüber hinaus sind die Kosten dieser miniaturisierten Satelliten dramatisch von etwa 400 Millionen auf nur noch 1 Million US-Dollar gesunken (Davenport 2021). Dadurch werden die in vergangenen Zeiten bestehenden Kostenhindernisse erheblich gesenkt.

Fazit

Anhand der deutlichen Fortschritte in der Satellitentechnologie, verbunden mit erweiterter Zugänglichkeit und geringeren Kosten dieser Technologien, ist eine Neubewertung von USE-Konzepten angemessen. Die wesentliche Relevanz bleibt bestehen, während die praktische Umsetzbarkeit zu wachsen scheint. In den kommenden Jahren werden immer robustere Satellitenfähigkeiten zusammen mit immer präziseren Informationsverarbeitungskapazitäten ein bisher unerreichtes Maß an Transparenz bieten. Während diese Informationsrevolution derzeit einen Wettbewerb in Form eines High-Tech-Wettrüstens anregt (Wright 2022; Raska 2020), ist die Möglichkeit, die staatliche Rechenschaftspflicht zu stärken und abschreckend auf konkurrierende Ausnutzungsversuche zu wirken, realistischer als je zuvor. Die Aufgabe besteht darin, Mechanismen zu konzipieren, die erhöhte Transparenz in erhöhte Kooperation umsetzen. Eine USE könnte ein solcher Mechanismus sein. Die nächsten Schritte wären die Ausarbeitung einer Forschungsagenda, die die technischen und institutionellen Designmerkmale und deren Zusammenführung in eine optimale Struktur untersucht.

Referenzen

Bennett, Mia M./Janice K. Chan and Colin J. Gleason (2022),: The Politics of Pixels: A Review and Agenda for Critical Remote Sensing, Progress, in Human Geography 46(3):729-52.

Davenport, Christian (2021): The Revolution in Satellite Technology Means there are Swarms of Spacecraft No Bigger than a Loaf of Bread in Orbit, The Washington Post, April 6, https://www.washingtonpost.com/technology/2021/04/06/small-satellites-growth-space/.

Erwin, Sandra (2023): Satellite Imaging Industry Responds to Demand for Intelligence Fusion, Space News, April 17, https://spacenews.com/on-national-security-satellite-imaging-industry-responds-to-demand-for-intelligence-fusion/.

Glaser, Charles (1997): The Security Dilemma Revisited, World Politics 50(1): 171-201.

Keohane, Robert and Lisa Martin (1995): The Promise of Institutionalist Theory, International Security 20(1): 39-51.

Moric, Igor (2022): How Commercial Satellite Imagery Could Soon Make Nuclear Secrecy Very Difficult – If Not Impossible, Bulletin of the Atomic Scientists, July 5, https://thebulletin.org/2022/07/how-commercial-satellite-imagery-could-soon-make-nuclear-secrecy-very-difficult-if-not-impossible/.

Moric, Igor (2023): Nuclear Stability in a World With Overhead Transparency, Comparative Strategy 42(5): 621-54.

Raska, Michael (2020): The Sixth RMA Wave: Disruption in Military Affairs, Journal of Strategic Studies 44(4): 456-79.

Satellite Industry Association (2021): State of the Satellite Industry Report. https://sia.org/news-resources/state-of-the-satellite-industry-report/.

United Nations (1983): The Implications of Establishing an International Satellite Monitoring Agency, https://digitallibrary.un.org/record/49898?ln=en.

Voute, Caesar (1984): Agreement and Disagreement on an International Satellite Monitoring Agency, Remote Sensing 5(2): 479-83.

Wright, Chris (2022): A High-Tech Arms Race is Shaping the Future of Warfare, Wired, April 13, https://wired.me/technology/security/a-high-tech-arms-race-is-shaping-the-course-of-war/.

Zhao, Bo/Shaozeng Zhang/Chunxue Xu/Yifan Sun and Chengbin Deng (2021): Deep Fake Geography? When Geospatial Data Encounter Artificial Intelligence, Cartography and Geographic Information Science 48(4): 338-52.

Timothy Williams

Frieden und Konflikt in der digitalen Ära

Digitale Technologien, insbesondere soziale Medien und Cyberfähigkeiten, haben soziale Beziehungen, gesellschaftliche Strukturen und politische Dynamiken grundlegend verändert. Menschen kommunizieren, diskutieren, argumentieren und interagieren in diesem digitalen Zeitalter mit anderen auf der ganzen Welt. Globalisierte Netzwerke aus bekannten Offline-Kontakten und neuen digitalen Kontakten entstehen, und zunehmend überlappen sich online/digitale und offline/physische Räume, sodass hybride Räume entstehen, in denen soziale Realitäten kommunikativ konstruiert und soziale Identitäten geschaffen werden. Hiermit haben sich auch Dynamiken in Bedingungen für und Konsequenzen von Frieden und Konflikt grundlegend verschoben und eine Reihe ganz neuer Herausforderungen, aber auch Chancen ergeben. Diesen neuen Gegebenheiten stellen sich neue Ansätze der Friedens- und Konfliktforschung und analysieren zunehmend digitale Gewalt, d. h., „die Verwendung von Worten, Bildern, Computercode oder Anweisungen innerhalb eines digitalen Raums, um Individuen oder Gruppen psychischen oder physischen Schaden zuzufügen.“[1] Hierbei kristallisieren sich zwei Schwerpunkte in der Forschung zu digitaler Gewalt heraus:

1. Online-Praktiken zur Organisation und Vorbereitung physischer Gewalt durch Mobilisierung, Legitimation und Motivation und

2. digitale Gewalt als diskursive Formationen, die sich in diskriminierenden Diskursen, Desinformation und Hassreden manifestieren.

Große Social-Media-Plattformen und Messaging-Dienste spielen eine wichtige Rolle, ebenso wie die vielen Nischenplattformen für spezialisierte Gruppen wie 8kun, 4chan oder Reddit, wo rechtsradikale Gedanken und Verschwörungsmythen verbreitet werden. Maßgeblich hängt die Kommunikation von der Architektur der Plattformen und der Struktur der Algorithmen ab, *Filterblasen* und *Echokammern* können entstehen, bei denen bestimmte Gruppen einen unterschiedlichen Zugang zu Informationen und Nachrichten, einschließlich Fehlinformationen und Fake News, erhalten, wodurch bereits bestehende politische Einstellungen verstärkt und die Polarisierung gefördert werden.

Dieser Vortrag führt zunächst in die digitale Welt ein und diskutiert, wie sich das Konzept der Gewalt konzeptionell in der digitalen Ära verändert hat und welche Konsequenzen das für die Erforschung haben könnte. Weiter zeigt der Vortrag auf, wie Konflikt und Gewalt in ihrer Entstehung und Dynamiken durch ihre digitale Einbettung verändert werden, bevor diskutiert wird, welche Chancen sich hiermit auch für die Prävention und Aufarbeitung ergeben.

Anmerkung

1 *Kilger M (2016): Interventions, Policies, and Future Research Directions in Cybercrime, in: C.A. Cuevas CA and C.M. Renniso (eds). The Wiley Handbook on the Psychology of Violence. Chichester: Wiley, 604–622: 606, eigene Übersetzung.*

Timothy Williams

Prof. Dr. **Timothy Williams** hat die Juniorprofessur für Unsicherheitsforschung und gesellschaftliche Ordnungsbildung in der Fakultät für Staats- und Sozialwissenschaften und ist Sprecher des Forschungszentrums RISK, beides an der Universität der Bundeswehr München.

Jens Hälterlein

ELSA zieht in den Krieg – Zur Rolle der Kritik an autonomen Waffensystemen für deren Legitimationsstrategien

Im ersten Teil des Vortrags möchte ich zunächst kurz das Projekt *Meaningful Human Control* (MEHUCO) vorstellen. Es sollen die Gesamtziele des Projekts, zentrale konzeptuelle Ansätze, Forschungsfragen und Arbeiten der einzelnen Teilprojekte beleuchtet werden.

Danach möchte ich mich einer Forschungsfrage widmen, die wir derzeit im Paderborner Teilprojekt verfolgen. Dabei geht es um die Beobachtung, dass seit kurzem Konzepte verantwortungsvoller und vertrauenswürdiger KI aus dem zivilen in den militärischen Bereich wandern. Nachdem ich diese Entwicklung anhand einiger Schlaglichter beleuchtet habe, möchte ich eine Analyseperspektive anlegen, die aus der Soziologie der Kritik stammt. Die dort herausgearbeitete Bedeutung von Kritik als sozialer Praxis soll für die Analyse der ethischen Legitimation von autonomen Waffensystemen (AWS) fruchtbar gemacht werden. Ich möchte zeigen, dass es als ein nicht-intendiertes Ergebnis bestimmter Kritiken an AWS angesehen werden kann, dass in militärischen Entwicklungskontexten das Leitbild einer *vertrauenswürdigen, menschenzentrierten und verantwortungsvollen KI* an Bedeutung gewonnen hat und Konzepte wie *ethics by design* und *Ethical Legal & Social Aspects* (ELSA) Einzug gehalten haben. Dass ein Teil der Kritik an AWS in deren Legitimationsstrategien eingeflossen ist, droht wiederum die Kraft der Forderung nach einem Verbot von AWS zu schwächen, da die Einführung ‚softer' Regulierungsinstrumente die Setzung von rechtlichen Regulierungen obsolet machen könnte. Institutionalisierte Ethik würde die internationale Gesetzgebung ersetzen. Angesichts dieses Szenarios möchte ich abschließend die Frage stellen, welche neuen Wege eine Kritik an AWS einschlagen könnte, um ihre Ziele zu erreichen.

Jens Hälterlein

Dr. **Jens Hälterlein** ist Wissenschafts- und Technikforscher und beschäftigt sich seit mehreren Jahren mit den gesellschaftlichen Dimensionen von digitalen Sicherheitstechnologien. Seit Mai 2022 arbeitet er im Teilprojekt *Schwarmtechnologien. Kontrolle und Autonomie in komplexen Waffensystemen* des Forschungsverbunds *Meaningful Human Control. Autonome Waffensysteme zwischen Regulation und Reflexion* (MEHUCO) an der Universität Paderborn.

Dieter Engels, Jürgen Scheffran und Ekkehard Sieker

Krieg im Weltraum? Es ist mal wieder Fünf vor Zwölf

Der Vortrag soll die Entwicklung der Weltraumrüstung in den letzten 40 Jahren nachzeichnen, die heutigen Herausforderungen für die Rüstungskontrolle aufzeigen und die Möglichkeiten der Naturwissenschaftler:innen beleuchten, Einfluss gegen eine Bewaffnung und für eine zivile Nutzung des Weltraums zu nehmen.

1983 hat US-Präsident Ronald Reagan das Raketenabwehr-Programm SDI auf den Weg gebracht mit der Begründung, die Bevölkerung der USA vor einer atomaren Vernichtung schützen zu wollen. Naturwissenschaftler:innen aus aller Welt haben davor gewarnt, dass ein solches Ziel unter den gegebenen technischen Rahmenbedingungen nicht erreichbar ist. In Deutschland hat die Auseinandersetzung mit dem SDI-Programm dazu geführt, dass sich Naturwissenschaftler:innen verstärkt als Teil der Friedensbewegung organisiert haben. Die Gründung der Zeitschrift *Wissenschaft und Frieden* ist dadurch maßgeblich beeinflusst worden. Heute, 40 Jahre danach, ist die Aufrüstung im Weltraum als unverzichtbar gesehener Teil der Unterstützung der Kriege auf der Erde mit Hilfe von Satellitensystemen nahe an der Grenze, auch Waffen im Weltraum zu stationieren, die gegen gegnerische Satelliten oder Ziele auf der Erde eingesetzt werden können. Auch vor dieser Entwicklung wurde bereits vor 40 Jahren gewarnt.

Dieter Engels, Jürgen Scheffran und Ekkehard Sieker

Dieter Engels, **Jürgen Scheffran** und **Ekkehard Sieker** haben vor bald 40 Jahren die Bücher *Die Front im All* (1984) und *SDI – Falle für Westeuropa* (1986) herausgegeben. Die Bücher behandelten die Militarisierung des Weltraums im Licht des damals vorgeschlagenen Raketenabwehrsystems SDI (Strategic Defense Initiative). Die Autoren sind Mitglieder der Naturwissenschatfler:innen-Initiative *Verantwortung für Frieden und Zukunftsfähigkeit*.

Lukas Rademacher

Wie verifiziert man nukleare Abrüstung?

Weltweit gibt es etwa 12.500 Kernwaffen unter der Kontrolle von neun Staaten. Da diese Waffen eine existentielle, weltweite Bedrohung darstellen, ist globale nukleare Abrüstung ein Ziel im allgemeinen Interesse. Wie genau dieses Ziel erreicht werden kann, ist ein umstrittenes Thema und hängt stark von den internationalen politischen Gegebenheiten ab. Dennoch ist klar, dass nachhaltige nukleare Abrüstung verifizierbar, das heißt überprüfbar, sein muss. Wie diese Verifikation gestaltet werden sollte, wirft jedoch noch Fragen auf. Dieser Beitrag beleuchtet einige der Kernherausforderungen und stellt dazugehörige Lösungsansätze aus der naturwissenschaftlichen Friedensforschung vor.

Ein essentieller Schritt für nukleare Abrüstung ist die Zerlegung der Sprengköpfe. Dabei ist es von entscheidender Bedeutung, sicherzustellen, dass tatsächlich Sprengköpfe zerlegt werden und keine Imitate vorliegen. Ein aktueller Forschungsgegenstand ist es, die Messmethoden so weiterzuentwickeln, dass selbst aufwändig produzierte Imitate als solche erkannt werden können. Ein weiterer wichtiger Aspekt ist die Kontrolle des verbleibenden Spaltmaterials, um die Herstellung neuer Waffen zu verhindern. Hierbei müssen nicht nur die Materialien aus den zerlegten Waffen sichergestellt werden, sondern auch das restliche Material aus dem nuklearen Brennstoffkreislauf. In diesem Zusammenhang entwickelt die nukleare Archäologie wissenschaftliche Methoden, um die Produktionshistorie von Nuklearanlagen mithilfe technischer Messungen und Simulationen zu rekonstruieren. Somit könnten Deklarationen zu nuklearen Materialbeständen eines Staates überprüft werden.

Die Gruppe *Nukleare Verifikation und Abrüstung* (NVD) an der RWTH Aachen widmet sich intensiv der Forschung zu diesen Themen, um die technischen Rahmenbedingungen für verifizierbare nukleare Abrüstung zu schaffen und das Ziel einer kernwaffenfreien Welt zu unterstützen.

Lukas Rademacher

Lukas Rademacher ist Doktorand in der Forschungsgruppe *Nuclear Verification and Disarmament (NVD)* im Fachbereich Physik an der RWTH Aachen. Dort forscht er zur Rekonstruktion der Betriebshistorien von Nuklearreaktoren mittels nuklearer Archäologie. Er ist außerdem Mitarbeiter im interdisziplinären Forschungsprojekt *VeSPoTec – Verifikation in einer komplexen und unvorhersehbaren Welt: Soziale, politische und technische Prozesse*.

Rebecca Froese, Daniela Pastoors, Jürgen Scheffran und Melanie Hussak

„Frieden verbessert das Klima“ – Wie Konflikttransformation zur Bewältigung der Klimakrise beitragen kann

In diesem Fishbowl-Gespräch geht es neben der kritischen Bestandsaufnahme der Klima-Konflikt-Risiken vor allem um positive Synergien, wie Friedens- und Konfliktarbeit zur Bewältigung der Klimakrise beitragen kann und wie sich Klimapolitik für den nachhaltigen Frieden nutzen lässt. Das Fishbowl-Format ist als offener Austauschraum gedacht, bei dem einzelne Impulse einführen und dann weitere Personen in den inneren Kreis dazustoßen und die Diskussion bereichern.

Die Zusammenhänge zwischen Klimawandel und Konflikten sind vielfältig und werden in Politik und Forschung zunehmend thematisiert. Dazu gehört auch der bislang vernachlässigte negative Einfluss von Rüstung, Militär und Krieg auf Umwelt und Klima bis zu Ökoziden, was auch im Russland-Ukraine-Krieg deutlich wird. Die positiven Verbindungen von nachhaltigem Frieden und Klimagerechtigkeit sind hingegen noch wenig erforscht, auch hinsichtlich innergesellschaftlicher Risiken und Konflikte. Es ist eine große Herausforderung, von einem negativen Nexus aus Klimawandel, Konflikten und Risiken zu einem positiven Nexus aus Frieden, Umwelt und Entwicklung zu kommen, der Synergien zwischen politischen Handlungsfeldern nutzt. Da die anstehenden Transformationen unausweichlich mit Konflikten einhergehen, müssen wir uns den Herausforderungen stellen und diese konstruktiv angehen. Um sozial-ökologische Transformationen nachhaltig und friedlich zu gestalten, müssen klimapolitische Strategien mit Konfliktbearbeitung und -transformation zusammengedacht werden.

Das Fishbowl-Gespräch wird an Hand mehrerer Teilfragen strukturiert sein, unter Berücksichtigung der genannten Aspekte, so dass nach und nach mehrere Themen gemeinsam beleuchtet werden können. Folgende Fragen stehen im Vordergrund:

- Wie GESTALTEN wir die „erhaltende Entfaltung“ (nachhaltige Entwicklung) des Lebens auf diesem Planeten?
- Auf welche Weise können wir Frieden zwischen Menschen und Frieden mit der Natur zusammenbringen und fördern?

Von den vier Autor:innen erscheint im thematischen Zusammenhang ein Artikel zum Thema *Erhalten, Entfalten, Gestalten: Mittel der Konflikttransformation für Wege aus der Klimakrise einsetzen* in der Jubiläumsausgabe von Wissenschaft und Frieden 4/2023.

Rebecca Froese, Daniela Pastoors, Jürgen Scheffran und Melanie Hussak

Dr. **Rebecca Froese** forscht am Zentrum für Interdisziplinäre Nachhaltigkeitsforschung und am Institut für Politikwissenschaften an der Universität Münster. Dort beschäftigt sie sich mit dem Aufbau von Reallaboren zur Erforschung der Rolle von Organisationen, insb. Hochschulen, in der Transformation zur Nachhaltigkeit. Die Forschung ist eingebettet in größere Fragen nach Konflikttransformation, Friedensbildung und Gerechtigkeit in der sozial-ökologischen Transformation. Rebecca ist Associate Fellow der Friedensakademie Rheinland-Pfalz/RPTU Kaiserslautern-Landau und Host des Podcast Fokus Frieden.

Daniela Pastoors ist Sprecherin im AK *Curriculum & Didaktik* der Arbeitsgemeinschaft *Friedens- und Konfliktforschung* und Vorstandsmitglied in der Stiftung *Kraft der Gewaltfreiheit* und …
… ist fasziniert davon, wie Erhalten, Entfalten und Gestalten in sozial-ökologischen Transformationsprozessen zusammenspielen.
… will dazu beitragen, planetares, kollektives und persönliches Wohlergehen gemeinsam zu fördern und zu stärken.
… erkundet, wie Friedens- und Konfliktarbeit ein gutes Leben für alle unterstützen kann.

Jürgen Scheffran ist Professor für Integrative Geographie an der Universität Hamburg und leitet die Forschungsgruppe Klimawandel und Sicherheit am Centrum für Erdsystemforschung und Nachhaltigkeit und im Klima-Cluster CLICCS. Er ist Mitglied der Redaktion von Wissenschaft und Frieden.

Melanie Hussak ist wissenschaftliche Mitarbeiterin (Postdoc) am Friedensinstitut Freiburg an der Evangelischen Hochschule Freiburg. Dort ist sie im Bereich Forschung, Lehre und Transfer aktiv, wie aktuell in dem Transferprojekt *Entwicklung-Frieden-Nachhaltigkeit: Konfliktbearbeitung und Bildungsarbeit im Kontext der Klimakrise*. Zudem ist sie Co-Leiterin des Netzwerkprojekts *Developing the Next-Generation of Shared Society Theory and Practices* der Universität Haifa sowie Associate Fellow der Friedensakademie Rheinland-Pfalz/RPTU Kaiserslautern-Landau.

Themenstellungen und Zielsetzungen der Zeitschrift *Wissenschaft und Frieden* und des Symposiums *Wissenschaft für den Frieden* anlässlich des 40-jährigen Bestehens der Zeitschrift

Der folgende Text ist ein leicht abgeänderter Auszug aus einem Förderantrag an die Deutsche Stiftung für Friedensforschung, der von Hans-Jörg Kreowski, Regine Mehl und Conrad Schetter gestellt und von Klaus Harnack, Paul Schäfer, Jürgen Scheffran und David Scheuing mitverfasst wurde.

Obwohl die Nationen der Welt mit ihrer Unterschrift unter die UN-Charta versprochen haben, „die Menschheit von der Geißel des Krieges zu befreien" und ihre Konflikte friedlich zu lösen, ist dies weiterhin für viele Teile der Welt keine Realität. Viele Konflikte drohen zu neuen Kriegen zu eskalieren, innergesellschaftlich herrscht vielerorts Unfrieden und Militarisierung, Aufrüstung und Abschreckung treiben kollektive Friedlosigkeit voran. In Europa ist dieser allgemeine Befund vielen Menschen erst durch den völkerrechtswidrigen Angriffskrieg Russlands gegen die Ukraine wieder bewusst geworden – und hat Menschen auch jenseits von Politik und Wissenschaft aufgeschreckt. Er verdeutlicht nicht nur das Versagen der bisherigen Europäischen Friedensordnung, auch die erneut eskalierte Frontstellung zwischen mehreren Atommächten wirkt wie ein Rückfall in längst für überkommen gehaltene Zeiten. Globale Herausforderungen wie die Gestaltung des Klimawandels müssen nun im Kontext hochgradiger Konfrontation angegangen werden, auch wenn im internationalen System die Chancen für eine wirksame Transformation schlecht stehen.

Gewaltarme bzw. gewaltfreie Konflikttransformation mit ihren Kernelementen Ausgleich, Kooperation und gegenseitige Verständigung hat in Kriegszeiten keine Konjunktur. Umso mehr sind die Friedenswissenschaften gefordert, sich für ausdifferenzierte Interpretationen des Friedensbegriffs einzusetzen. Aus ihrem jahrzehntelang entwickelten Fundus sind die Friedenswissenschaften in der Lage, konkrete Konzepte und Vorhaben in den Diskurs einzubringen. Die Datenlage ist valide und viele ihrer Erkenntnisse können empirisch untermauert werden. Doch gibt es ein kommunikatives „Loch" zwischen diesen Polen wissenschaftlich begründeter umfassender Wege zum Frieden und öffentlicher Diskussion um Kriegsziele und Friedenserreichung. Das hier beantragte Symposium *Wissenschaft für den Frieden* soll dazu beitragen, diese Lücke zu schließen und aktuelle Debatten auf eine fundierte Basis zu stellen. Dies wird als Ausgangsbedingung gesehen, um auch den Blick auf eine zukünftige Friedenspolitik zu lenken: Welche Schritte müssen jetzt unternommen werden, um in einer Zukunft gewaltfreie Konflikttransformation und -lösung Wirklichkeit werden zu lassen?

Das Symposium will das interdisziplinäre Feld der Friedens- und Konfliktforschung im Rückblick und vor allem aber im Ausblick beleuchten. Es soll dazu dienen, neue und alte Herausforderungen zu benennen und Leitlinien für eine Friedens- und Konfliktforschung der Zukunft zu skizzieren. Das Symposium lädt dazu ein, sich Themen, Konzepten und Theorien aus den vergangenen 40 Jahren der Zeitschrift *Wissenschaft und Frieden* (W&F) zu nähern und diese auf ihre Tragfähigkeit und ihren Mehrwert für die Zukunft abzuklopfen. Das Symposium ermutigt alle Beitragenden, Zustand und Zukunft dieses Forschungs- und Aktionsfelds zu beschreiben und zu analysieren, neue Impulse zu setzen und Friedenswissenschaft für die Zukunft zu skizzieren.

Deshalb ergeben sich folgende zentrale Fragen:

1. Wo steht die Friedens- und Konfliktforschung heute?
2. Welche Friedensforschung wird in Zukunft gebraucht?
3. Welche Rolle können und werden dabei W&F und die Wissenschaftskommunikation einnehmen?

Als Ergebnisse des Symposiums werden neue Einsichten mindestens in folgenden Themenbereichen erwartet, die der Friedens-

und Konfliktforschung im weitesten Sinne neue Impulse verleihen und auch in die Friedensbewegung sowie in Politik und Medien ausstrahlen sollen:

- Naturwissenschaftliche und technische Aspekte von Rüstung, Rüstungskontrolle und Abrüstung insbesondere hinsichtlich Atomwaffen und Atomkriegsgefahr sowie neuer Waffensysteme,

- Gefährdung des Friedens durch Klimawandel und Umweltzerstörung gegenüber der Notwendigkeit einer nachhaltigen Weltwirtschafts- und Friedensordnung,

- Gefährdung des Friedens durch Nord-Süd-Konflikt, verfehlte Entwicklungspolitik, postkoloniale und (neo-)imperiale Verhältnisse und wirtschaftliche Abhängigkeiten,

- Zivile Alternativen gewaltfreier Konfliktbearbeitung und -transformation,

- Der Krieg Russlands gegen die Ukraine als aktuelle Herausforderung und Beispiel für eine krisenorientierte multiperspektivische Friedensforschung und Friedensbearbeitung (Geschichte, Politik, Jura, Soziologie, Anthropologie, Psychologie),

- Rolle der Ideologisierung, Religion und der religiösen Systeme beim Missbrauch von Glauben für Konflikte einerseits und religiöse Friedensgestaltung andererseits,

- Nationale Sicherheitsstrategie und die deutsche Rolle in der europäischen Friedensordnung: Friedens- versus Kriegslogik und die Rolle der Friedens- und Konfliktforschung,

- Rolle der Medien einschließlich sozialer Medien zwischen „Mediendemokratie" und Friedensjournalismus.

Um globale, zwischenstaatliche wie innergesellschaftliche Konflikte zu bewältigen, werden Analysen von Gewaltursachen und Konfliktdynamiken ebenso benötigt wie gewaltfreie Lösungen und Ansätze ziviler Konfliktbearbeitung. Wer die Krisen unserer Zeit meistern will, muss Grenzen zwischen Fachdisziplinen, Politik und Gesellschaft überwinden. Diese Transferleistung erbringt die Vierteljahreszeitschrift W&F seit 1983 für Friedensforschung, Friedenspolitik und Friedensbewegung wie keine andere Zeitschrift im deutschsprachigen Raum. Das 40-jährige Jubiläum ist somit Anlass, mit dem Blick nach vorn zu fragen, mit welchen Inhalten und Formen und in welchen Strukturen Wissenstransfer zu Konflikten, Gewalt und Frieden in die Gesellschaft im 21. Jahrhundert leistbar ist.

W&F ist deutschsprachig und deckt thematisch das gesamte interdisziplinäre Spektrum von Friedens- und Konfliktforschung insbesondere auch im internationalen und globalen Kontext ab. Allein schon die Titel der letzten drei Jahre weisen daraufhin: *Gewalt/Ökonomie, Krieg gegen die Ukraine, Kriegerische Verhältnisse, Täter:innen, Chinas Welt, Frieden lernen, aber wie?, Völkerrecht in Bewegung, „Friedensmacht" EU?, Umwelt-Klima-Konflikt, Der kranke Planet*. Das Symposium will diese thematische Breite aufnehmen und sich dabei vor allem auf die Zukunftsperspektiven fokussieren.

Das Symposium folgt damit der Ausrichtung von W&F, die auf Wissenstransfer aus dem gesamten fachlichen Spektrum angelegt ist. Auch der Veranstaltungsort Bonn wurde zur Unterstützung des Transfergedankens bewusst gewählt. Symbolisiert er nicht nur die Anfangstage der breiten öffentlichen Nachrüstungsdebatte der 1980er Jahre (und damit auch die Geburtszusammenhänge der Zeitschrift), so ist Bonn heute als deutscher Standort von UN-Organisationen, als Sitz des BMZ und anderer entwicklungs- und umweltpolitischer Organisationen sowie friedenswissenschaftlicher Einrichtungen wie dem BICC prädestiniert dafür, ein solches Kolloquium zu beherbergen. Durch die Verankerung der Veranstaltung im Rahmen der Veranstaltungstage *Bonner Friedenstage 2023* und auch durch eine explizit stadtöffentliche Bewerbung soll die Hemmschwelle für die Teilnahme an der Konferenz für Bürger:innen gesenkt und der „Elfenbeinturm" geöffnet werden.

Nicht zuletzt soll das Symposium auch der Stärkung von W&F als neben der Friedenswarte letzte verbliebene friedenswissenschaftliche Zeitschrift im deutschsprachigen Raum dienen. Ihr Spezifikum des Transfers wissenschaftlicher Expertise, ihre interdisziplinäre Organisationsstruktur und ihre intergenerative redaktionelle Zusammenarbeit sollen in Vor- und Nachbereitung das Symposium sichtbar prägen.

Das Symposium ist als Forum gedacht für die Auseinandersetzung mit aktuellen und zukünftigen Entwicklungen in der Friedens- und Konfliktforschung, der Friedensbewegung und ihrer Rezeption in Medien und Politik. Erwartet und gesucht werden daher entsprechend zukunftsorientierte Beiträge zu folgenden und verwandten Themen:

- (Selbst-)Kritische Retrospektiven auf 40 Jahre W&F,

- aktuelle friedenspolitische, militärstrategische und rüstungstechnische Fragen,

- Analyse und Kritik von Gewaltursachen und -verhältnissen,

- neue Herausforderungen für die Konfliktforschung und Konflikttransformation im 21. Jahrhundert: konzeptionelle, empirische und methodische Beiträge,

- Wege und Möglichkeiten zur zivilen und gewaltfreien Konfliktlösung, Sicherung des Friedens, zur Wahrung der Menschenrechte und zur Zukunftssicherung nach der „Zeitenwende",

- zunehmende Ideologisierung und Missbrauch religiöser Sichtweisen zur Begründung politischer, gewaltbehafteter Handlungen,

- Positionen zur Verantwortung der Wissenschaft sowie der Rolle der Natur- und Ingenieurswissenschaften in Krieg und Frieden; ebenso die Rolle der Sozial- und Geisteswissenschaften in Krieg und Frieden,

- Rolle der global agierenden Medien und ihrer Einflusssphären,

- Kurzdarstellung von erfolgreichen oder gelungenen Praxisbeispielen.

40 Jahre Wissenschaft und Friedenspolitik

Grußwort

40 Jahre Wissenschaft & Frieden sind ein schöner und ein trauriger Grund zu feiern. Schön, weil es zeigt, dass eine Publikation, die friedenswissenschaftliche und friedenspolitische Themen in den Vordergrund rückt, über diesen langen Zeitraum erfolgreich Bestand haben kann. Traurig, weil das Thema gerade in diesen Tagen wieder eine erschütternde Aktualität erlangt – und eine Zeitschrift, die friedenspolitische Alternativen der Militärlogik entgegensetzt, gerade dadurch so wichtig ist.

Blicken wir zurück auf das Jahr, in dem *Wissenschaft und Frieden* gegründet wurde: 1983 – der kalte Krieg war auf seinem Höhepunkt, die gerade ins Amt gekommene Bundesregierung setzte gegen große gesellschaftliche Widerstände den *NATO-Doppelbeschluss* um, Ereignisse wie der Abschuss eines südkoreanischen Passagierflugzeugs durch sowjetrussisches Militär rückte die Welt an den Rand eines Atomkriegs und der Spielfilm *The day after* führte die Folgen dieses Atomkriegs plastisch vor Augen (und löste angeblich bei dem US-amerikanischen Präsidenten Ronald Reagan ein Umdenken aus). Wie wir heute wissen, verdankt die Menschheit womöglich ihr Überleben nur dem sowjetrussischen Offizier Stanislaw Petrow, der angesichts eines gemeldeten US-amerikanischen Angriffs einen Fehlalarm vermutete und – entgegen seinen Anweisungen – keinen Gegenschlag einleitete.

Was wäre wohl geschehen, wenn nicht der Mensch Stanislaw Petrow, sondern ein automatisiertes System einen Angriff „erkannt" und programmgemäß einen Vergeltungsprozess in Gang gesetzt hätte? Die Gefahr eines *Atomkriegs aus Versehen* war ein Gründungsimpuls für das Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FIfF) ein Jahr später – dies zeigt die enge Überschneidung der Ziele des FIfF und von *Wissenschaft und Frieden*, die uns bis heute eint. Nicht zuletzt durch die Entwicklungen der Künstlichen Intelligenz bekommt das Thema eine bedrückende Aktualität, wie der Beitrag von Bläsius und Siekmann in der W&F-Ausgabe 3/2023 zeigt. Fast könnte man meinen, die Menschheit habe in den letzten 40 Jahren kaum Fortschritte gemacht.

Ohnehin wird Pazifismus und Friedenspolitik in diesen Tagen auf eine harte Probe gestellt. Ist es bereits die Belohnung eines Aggressors für seinen Angriff, wenn wir eine friedliche Lösung auf diplomatischen Weg anstreben? Friedenspolitik muss lange vor dem Ausbruch eines Krieges beginnen. Die Zeitschrift *Wissenschaft und Frieden* legt dafür seit 40 Jahren die Grundlagen, indem sie gesellschaftliche Konflikte benennt, militärische Entwicklungen darstellt und friedliche Lösungen aufzeigt.

Das Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung gratuliert der Zeitschrift *Wissenschaft und Frieden* zum 40-jährigen Bestehen. Wir wünschen uns mindestens weitere 40 Jahre friedenswissenschaftliche und friedenspolitische Analysen, Stellungnahmen und Lösungsansätze für eine friedliche Welt und werden auch weiterhin sehr gern unseren Teil dazu beitragen.

Stefan Hügel
Vorsitzender des FIfF

Wissenschaft & Frieden 4/23
40 Jahre W&F – Wissenschaft für den Frieden

Seit vierzig Jahren erscheint W&F durchgehend und liefert maßgebliche Beiträge zur friedenswissenschaftlichen Diskussion – auch in der breiteren Öffentlichkeit. Heft 4/2023 feiert diese vierzig Jahre und blickt mit einer Reihe von Beiträgen auch auf das eigene Archiv und vor allem aber voran. Ein Angebot mit klarem Kompass zu schaffen, das ist seit jeher der Anspruch von W&F. Mehr noch als bei der Gründung der Zeitschrift in Marburg vor 40 Jahren, als mit Fragen der Raketenstationierung, Weltraumrüstung und der Friedensbewegung die Themen des Blattes klar schienen, geht es heute viel stärker um die Herausforderung der gezielten und hilfreichen Auswahl von friedenspolitisch relevanten Beiträgen, um diese anschließend möglichst auch allgemeinverständlich zu präsentieren. Dieses Heft ist ein Potpourri, das die Vielfalt an Perspektiven und Zugängen zeigt, für die W&F seit vierzig Jahren steht.

Dem Heft liegt als Dossier der *RüstungsExportKontrollAtlas* bei, der die Tradition der Rüstungsatlanten erweitert um die Perspektive auf die nationale Rüstungsindustrie, die Exportpolitik und die Exportkontrolle. Zum ersten Mal sind hier Finanzierung, maßgebliche Standorte und aktuelle rüstungspolitische Herausforderungen, exportpolitische Kritik und Konsequenzen deutscher Rüstungsexporte sowie Möglichkeiten für eine Neuregelung der Rüstungsexportkontrolle in einem vielfach illustrierten Atlas versammelt.

W&F 4/23 | November/Dezember | 68 Seiten | 12 € (druck) / 9 € (ePUB+PDF)

Vincent Först

Dein Bild als Beute

18. November 2023 – *Das heimliche Filmen in der Öffentlichkeit hat sich zu einem eigenen Genre in den sozialen Medien entwickelt. Auf der Jagd nach authentischen Inhalten setzen sich Content-Creator:innen über die Privatsphäre und Rechte ihrer Mitmenschen hinweg. Eine neue Form der Überwachung entsteht.*

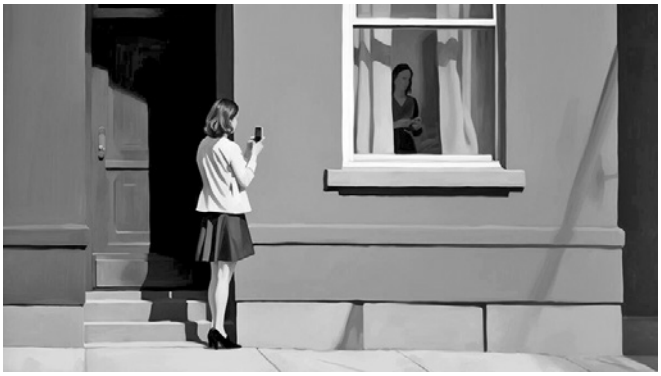
Lennart[1] sitzt im Fenster seiner Berliner Wohnung und raucht. Aus einem Café auf der gegenüberliegenden Straßenseite richtet ein unbekannter Musiker die Linse seiner Handykamera auf den jungen Mann und beginnt zu filmen. Lennart nimmt einen Zug von seiner Zigarette, schüttelt sachte den Kopf, als würde er einen unliebsamen Gedanken verscheuchen und blickt verträumt die Straße hinunter. Der Musiker stoppt die Aufnahme. Er legt seine neue Single unter das Video und veröffentlicht den Clip auf Instagram. Lennarts Cousine ist die erste einer ganzen Reihe von Freund:innen, Familienmitgliedern und Bekannten, die sich mit einem Link zum Video meldet: „Hier, deine fünfzehn Sekunden Fame." Lennart antwortet schockiert: „What the fuck." Innerhalb weniger Tage sehen über zwei Millionen Menschen das Video.

In den Kommentaren häufen sich derweil kritische Stimmen: „New Fear unlocked: those days while you're sitting on window smoking as you wanna process your shitty day someone else is also watching and recording like leave ppl alone". Knapp zwanzigtausend Likes gibt es für den Beitrag „A ***** can't even sit out his window and have a cig in peace anymore smh (Abk. für **s***hake* **m***y* **h***ead*)." Eine empörte Nutzerin fragt: „Can we stop normalising filming strangers?"

Mitmenschen als Content

Das Verhalten ist bekannt: Bei jeder sich bietenden Gelegenheit zücken Menschen ihr Smartphone und beginnen zu filmen oder zu fotografieren. Sie verarbeiten ihre Umwelt dabei zu *Content*[2], sprich in digitale Inhalte, die sie online teilen. Da sich beinahe die gesamte westliche Gesellschaft auf sozialen Medien bewegt[3], konsumieren und produzieren ihre Akteur:innen unablässig Content. Das gilt sowohl für Unternehmen von *Duolingo*[4] bis *Ryanair*[5], Personen des öffentlichen Lebens wie Leichtathlet:innen[6] oder Pornostars[7], Prominente, Künstler:innen[8] und die, die es werden wollen, sowie für eine stetig wachsende Zahl an Privatpersonen.

Aus dem Kreieren und Teilen von digitalen Inhalten respektive Content durch Nutzer:innen sozialer Medien ist unlängst ein eigener Wirtschaftszweig entstanden. Die sogenannte *Creator Economy*[9] setzt um die 250 Milliarden Dollar pro Jahr um, Tendenz steigend[10]. Es soll inzwischen über 200 Millionen[11] Content-Creator:innen auf dem Planeten geben. Davon verdienen nur etwa ein Prozent ihren Lebensunterhalt mit der Erstellung von Content.



Content-Creator:innen verwandeln die zufällige Schönheit eines besonderen Moments in visuelles Kapital. Quelle der Bilder: generiert von Vincent Först mit Midjourney, PD

Die Architektur und Funktionsweise sozialer Medien gibt eine auf Masse ausgelegte Postingfrequenz[12] vor, die Creator:innen dazu ermuntert, so viel Content wie möglich zu produzieren. Um für das Publikum eine Projektionsfläche zu bieten, versuchen Creator:innen ihren Content „relatable" zu gestalten. Relatable sind Inhalte dann, wenn sich die Nutzer:innen sozialer Medien mit ihnen identifizieren können[13].

Dazu gilt auf Instagram, TikTok und Co. vermeintliche „Echtheit" als wertvolles Gut[14]. Heimliche Aufnahmen versprechen Authentizität, da sich die Gefilmten „natürlich" verhalten. Symbolträchtige und stark emotionalisierte Handlungen und Bilder sind besonders beliebt. Diese Art von Content funktioniert so gut auf den sozialen Medien, dass sich um das heimliche Filmen in der Öffentlichkeit inzwischen eigene virale Genres gebildet haben.

Aufnahmen ohne Zustimmung der Gefilmten

Für den Trend „what people are wearing"[15] filmen Beobachter:innen aus Cafés, Restaurants, Bars oder den eigenen vier Wänden vorübergehende Passant:innen mit ausgefallenen Outfits – ohne vorher zu fragen. Die Zusammenschnitte der Aufnahmen werden später auf TikTok und Instagram geladen.

Unter den Suchbegriffen *NPC encounter*[16] oder *NPC conversations* finden sich Aufnahmen von Menschen, die sich in der Öffentlichkeit „abnormal" bis aggressiv verhalten. Der Begriff *NPC* stammt aus der Videospielkultur[17]. Die Abkürzung steht für

Non-Player-Character und beschreibt computergesteuerte Charaktere, die sich nach einem von den Spiele-Entwickler:innen vorgegebenen Handlungsmuster verhalten. Ihre repetitiven Monologe und Bewegungen sind oft unfreiwillig komisch.

Die Filmenden machen die Interaktionen mit den menschlichen *NPCs* zu einem Spektakel, das ein Millionenpublikum über den Smartphone-Bildschirm begafft und verhöhnt. Wer „NPC berlin" in der Suchleiste von Tiktok eingibt, bekommt unter anderem Videos von schlafenden Obdachlosen und Junkies beim Drogenkonsum in Berliner U-Bahnstationen angezeigt.

Im Rahmen des „*random act of kindness*"[18] Trends filmen sich Menschen dabei, wie sie anderen Menschen vorgeblich etwas Gutes tun, um die Videos dann auf ihren Kanälen hochzuladen. Dafür überreichen Creator:innen vor laufender Kamera Blumensträuße an alte Damen oder schenken Bettler:innen ein Mittagessen.

Die Verlockung der Reichweite

Auch wenn TikToker:innen an Bahnhöfen und in Fußgänger:innenzonen[19] Trends nachtanzen, im Fitness-Studio schwitzen[20], oder sich als Tube-Girls[21] in U-Bahnen inszenieren, geraten dabei zwangsläufig andere Menschen ins Bild. Eine Verpixelung der Gesichter findet dabei nur in den seltensten Fällen statt. Im Gegenteil heben Creator wie *ducuri.gmv*[22] oder *babalagrande*[23] unfreiwillige Interaktionen sogar mit Texteinblendungen über den Köpfen der Gefilmten hervor.

Wenn sich die Nutzer:innen über Inhalte empören und wie in Lennarts Fall ihren Ärger in kritischen Kommentaren ausdrücken, dient das der algorithmischen Verbreitung des Videos. Denn negative Gefühle erregen mehr Aufmerksamkeit[24] – ein Katalysator für „streitbaren" Content. Der Bruch der Privatsphäre lohnt sich.

Obwohl ein virales Video noch keine Goldgrube[25] für Creator:innen darstellt, hat die Verlockung der Reichweite eine starke Anziehungskraft entwickelt, hinter der Bedenken über Privatsphäre und mögliche Folgen für Gefilmte[26] verblassen. Für die Betroffenen besteht die latente Gefahr, später als Meme[27] durch das Netz zu geistern.

Unternehmerischer Blick

Selbstverständlich sind Smartphone-Kameras auch Waffen gegen Machtmissbrauch und Empowerment-Instrumente[28]. Die Aufdeckung unzähliger Straftaten ist der Zivilcourage von Filmenden[29] zu verdanken. Durch ihre Reichweite verschaffen Influencer:innen und Filmende wichtigen Themen die nötige Aufmerksamkeit[30]. Wenn ein Künstler betrunkene Business-Männer[31] auf Tokios Straßen fotografiert, sagt das immerhin etwas über die Arbeitskultur der japanischen Gesellschaft aus. Der umstrittene Rüpelfotograf *Bruce Gilden*[32], der Fußgänger:innen auf den Straßen New Yorks regelrecht auflauert und Portraits ohne deren Einwilligung schießt, ist mit seinem tragbaren Blitz und der analogen Kamera wenigstens als Fotograf identifizier- und ansprechbar. Den Aufnahmen der unsichtbaren

Heimliche Aufnahmen versprechen Authentizität, da sich die Gefilmten „natürlich" verhalten.

Creator:innen liegt dagegen oft kein künstlerischer oder aktivistischer Impuls zugrunde.

Es geht um Masse, Clicks und Reichweite, die harte Währung der sozialen Medien. Linoya Friedman, die selbsternannte Erfinderin von „what people are wearing", hat um den Trend eine Marke aufgebaut. Fans können die Outfits der gefilmten Menschen schnell und unkompliziert über *Affiliate-Links*[33] nachshoppen. Friedman kassiert dafür Provisionsgebühren.

Eine französische Bloggerin filmt von ihrem Büro aus Pariser Bürger:innen, lädt die Videos unter dem Hashtag *#stylespy* auf Instagram hoch und geht damit regelmäßig viral. Die dadurch gewonnene Reichweite hilft beim Verkauf der selbstgeschriebenen Reiseführer.

Dazu folgen zahlreiche Amateur-Accounts dem Erfolgsrezept des heimlichen Filmens in der Hoffnung, einen viralen Hit zu landen. Sie geben sich als versteckte Dokumentarfilmer:innen aus, wenn sie beispielsweise mit „love at every age" betitelte Videos von einer zu Straßenmusik tanzenden Mutter mit Kind auf dem Arm hochladen oder einen gutaussehenden jungen Mann wie Lennart filmen, der verträumt in den Berliner Himmel blickt und bewerben gleichzeitig eigene Produkte. Die Creator:innen ignorieren dabei, dass in Deutschland ein Recht am eigenen Bild existiert.

Heimliches Filmen kann strafbar sein

Die heimliche Anfertigung von Videoaufnahmen von Privatpersonen sei häufig rechtswidrig, sagt Nima Valadkhani, Rechtsanwalt und Experte für Urheber-, Medien- und Kartellrecht der Kanzlei Advant Beiten. Im Fall von Lennart gelte dies umso mehr, da er sich innerhalb seiner Wohnung befinde. Eine Wohnung unterliegt, ähnlich wie ärztliche Behandlungszimmer oder Umkleidekabinen, einem besonderen rechtlichen Schutz gegen Einblicke. Hinzu komme, dass der Musiker Lennart im Rahmen einer geschäftlichen Handlung dazu „benutzt", um seine neue Single zu bewerben.

Betroffene hätten Anspruch auf Schadenersatz, die Löschung des Videos, Auskunft, wo das Video überall hochgeladen wurde und gegebenenfalls auch darüber, wie viel Gewinn der Urheber

mit dem Video erzielt hat. Dass gegen Täter:innen juristische Mittel[34] eingelegt werden, passiert in Deutschland aber nur selten, auch weil die gefilmten Personen keine Kenntnis über die Veröffentlichung erlangen.

Schlafende und hilflose Menschen zu filmen sei ein massiver Eingriff am Recht am eigenen Bild und kann sogar strafrechtlich relevant sein, da sich die Betroffenen in einem Zustand der Ohnmacht befinden. Ein Gericht könne hier Geld- oder gar Haftstrafen verhängen. Laut Valadkhani komme es für die zivil- und strafrechtliche Bewertung jedoch immer auf den Einzelfall und die Rechtsanwendung durch die zuständigen Richter:innen an.

Symbol der Oberflächlichkeit

Nach einem Treffen mit Lennart und den Hinweis auf die Rechtswidrigkeit der Veröffentlichung hat der Musiker das virale Video entfernt. Der Vorfall hat sich für ihn trotzdem gelohnt: Immerhin haben zwei Millionen Menschen, die das Video auf Instagram gesehen haben, seine Musik gehört – mehr als jemals zuvor. „Eine gewisse Genugtuung gab es da trotzdem", sagt Lennart im Gespräch mit netzpolitik.org. „Einige Leute haben in mir wohl Dinge gesehen, mit denen ich mich teilweise auch identifizieren kann. Das beweist ja, dass ich eine gewisse Ästhetik transportiert habe."

Die Interpretationen des Videos in der Kommentarspalte lieferten dafür ein falsches Bild von der Situation. Dabei stört Lennart ganz grundsätzlich das Täuschungspotential sozialer Medien: „Als virale Persona werde ich gewissermaßen zu einem Symbol für diese Oberflächlichkeit. Denn ehrlich gesagt hatte ich in diesem Moment keine besonders tiefen Gedanken. Ich saß einfach da und war ich selbst." Für Lennart bleibt das unbehagliche Gefühl, im Fenster der eigenen Wohnung jederzeit beobachtet werden zu können. Seitdem hat sich seine Selbstwahrnehmung verändert. „Eigentlich will ich mir keine Gedanken über mein Aussehen machen, wenn ich in meinem Fenster oder auf meinem Balkon sitze", sagt Lennart. Damit sei es seit der Veröffentlichung des Videos vorbei.

Eine neue Form der Überwachung

„Das Bedürfnis nach Bestätigung der Realität und Ausweitung des Erfahrungshorizonts durch Fotografien ist ein ästhetisches Konsumverhalten, dem heute jedermann verfallen ist. Die Industriegesellschaften verwandeln ihre Bürger in Bilder-Süchtige; dies ist die unwiderstehlichste Form von geistiger Verseuchung", hielt Susan Sontag in ihrem 1977 erschienenen Buch *Über Fotografie*[35] fest. Laut Sontag lehren uns Bilder einen neuen visuellen Code[36], der die Vorstellung davon verändert, was wir für anschauenswert und beobachtbar erachten. Was damals noch allzu dystopisch klang, bedarf durch die dauerhafte Präsenz von hochauflösenden Kameras und der Möglichkeit, jedes Bild in Echtzeit zu teilen, einer gedanklichen Revision.

Wenn virale Trends wie „what people are wearing", „NPC encounter" oder die „random acts of kindness" regelmäßig in den Feeds und For-You-Pages auftauchen, radikalisiert sich durch deren Rezeption der Blick der Nutzer:innen. Das heimliche Filmen von Menschen wird normalisiert, die aggressive Aneignung ihrer Erfahrungen zum Tagesgeschäft. Durch das Zusammenspiel der weitgehend zügellosen Empfehlungsmechanismen sozialer Medien wie TikTok[37] und der Allgegenwart von Smartphones kann ein Beobachtungsdruck für alle anderen entstehen, was wiederum das Verhalten im öffentlichen Raum beeinflusst[38].

Selbst diejenigen, die sich den sozialen Medien entziehen, laufen Gefahr, von Creator:innen zu Subjekten ebendieser[39] gemacht zu werden. So entsteht eine neue Form der privaten Überwachung. Dahinter stehen nicht wie üblich datenhungrige Konzerne oder Sicherheitsbehörden, sondern Menschen mit gezückten Smartphones, die jederzeit auf „Record" drücken können, um uns ihrem „Content Gaze" zu unterwerfen.

Doppelte Ausbeutung

Diese ästhetische Ausbeutung durch die Creator:innen wiegt dabei doppelt schwer: Sie nehmen uns das Recht am eigenen Bild, um es dann als digitale Ware zu vermarkten und ihrem jeweiligen Geschäftsmodell, sei es im Dienste von Clicks, Reichweite, Aufmerksamkeit oder Einnahmen durch Werbung, zuzuführen.

Sie verwandeln die zufällige Schönheit eines besonderen Moments, wie er sich beispielsweise in der innigen Umarmung zweier Verliebter auf einer Parkbank zeigt, in visuelles Kapital. Wer heute in der Öffentlichkeit die Kontrolle verliert, in der U-Bahn einschläft, sich außerordentlich gut oder wahlweise schlecht anzieht, stilvoll Zigaretten raucht, *Main Character Energy*[40] ausstrahlt, szenisch küsst, lacht, weint oder sich auf jede andere erdenkliche Weise erzählenswert verhält und so ein narratives Potential für soziale Medien schafft, muss damit rechnen, heimlich fotografiert oder gefilmt zu werden und später ungefragt im Netz zu landen.

Quelle: https://netzpolitik.org/2023/ohne-einverstaendnis-dein-bild-als-beute/

Vincent Först

Vincent Först arbeitet als freier Autor und Journalist. An der Universität der Künste lehrt er Texttheorie- und Textgestaltung. Wenn er nicht gerade an seinem Schreibtisch sitzt, organisiert er Kulturveranstaltungen in Berlin. Mehr von ihm gibt es auf *Twitter*[41].

Anmerkungen

1 *Name geändert*
2 *https://en.wikipedia.org/wiki/Content_creation*
3 *https://de.statista.com/themen/1842/soziale-netzwerke/#topicOverview*
4 *https://www.instagram.com/p/CvxRMoNMCEV/*
5 *https://www.instagram.com/reel/CpiUBiAoew4/*
6 *https://www.instagram.com/p/Cu7eIR0xs6V/*
7 *https://www.instagram.com/p/CyWCvLbgKNP/?hl=de*
8 *https://www.instagram.com/p/CsGbpkYMbvN/*
9 *https://en.wikipedia.org/wiki/Creator_economy*
10 *https://www.wpbeginner.com/research/creator-economy-statistics-that-will-blow-you-away/*
11 *https://www.theleap.co/blog/creator-economy-statistics/*
12 *https://medium.com/@sozialpr/perfekte-postingfrequenz-oder-relevanz-und-nutzen-vor-volumen-baa0269b526b*
13 *https://www.nytimes.com/2019/01/08/magazine/the-scourge-of-relatable-in-art-and-politics.html*
14 *https://www.forbes.com/sites/forbesbusinesscouncil/2023/05/31/how-authentic-content-creators-build-social-media-trust/*
15 *https://www.tiktok.com/search?q=what people are wearing&t=1699617798659*
16 *https://www.youtube.com/results?search_query=npc+encounter*
17 *https://knowyourmeme.com/memes/real-life-npc*
18 *https://netzpolitik.org/2022/random-acts-of-kindness-so-monetarisierst-du-omas-und-obdachlose-auf-tiktok/*
19 *https://www.tiktok.com/@ducuri.gmv/video/7265696681943977217?lang=de-DE*
20 *https://www.tiktok.com/@thejoeyswoll/video/7284431006814211370*
21 *https://www.tiktok.com/@sabrinabahsoon/video/7278386962401152264?lang=de-DE&q=original tube girl&t=1699280023031*
22 *https://www.tiktok.com/@ducuri.gmv?lang=de-DE*
23 *https://www.tiktok.com/@babalagrande?lang=de-DE*
24 *https://www.leadersnet.de/news/64674,social-media-negative-posts-erregen-mehr-aufmerksamkeit.html*
25 *https://www.instagram.com/p/CtKEtyPOI2t/?hl=de*
26 *https://www.theguardian.com/technology/2022/jul/14/melbourne-woman-dehumanised-by-viral-tiktok-filmed-without-her-consent*
27 *https://medium.com/@MyMediaCritique/being-a-meme-ruined-my-life-2d7eadbb0c7d*
28 *https://www.spiegel.de/psychologie/heimwegtelefon-ich-habe-frauen-schon-zu-strassen-navigiert-die-kameraueberwacht-sind-a-5a7d2fb0-cd35-4322-a823-35b4ae1a0fc5*
29 *https://www.nbcnews.com/think/opinion/what-giving-pulitzer-prize-filming-george-floyd-s-murder-darnella-ncna1270778*
30 *https://www.tiktok.com/@charlidamelio/video/6832733422453116166*
31 *https://paweljaszczuk.com/High-Fashion-1*
32 *https://youtu.be/kkIWW6vwrvM?si=lsqU7Uhvas_OXPeV*
33 *https://de.wikipedia.org/wiki/Affiliate-Marketing#Funktionsweise*
34 *https://www.zdf.de/nachrichten/panorama/fotos-datenschutz-strafrecht-100.html*
35 *https://www.fischerverlage.de/buch/susan-sontag-ueber-fotografie-9783596230228*
36 *https://www.deutschlandfunkkultur.de/soziale-medien-koerper-wahrnehmung-junge-menschen-100.html*
37 *https://taz.de/Main-Characters-auf-Social-Media/!5833699/*
38 *https://youtu.be/L6IXkLb9FVs?si=_Q8tNKlxW4N5A0IT&t=4585*
39 *https://www.instagram.com/p/CXL03wsJHG_/*
40 *https://www.wikihow.com/Main-Character-Energy*
41 *https://twitter.com/vincentfoerst*

Daniel Leisegang

IT-Sicherheit gerät zur Randnotiz

17. November 2023 – *Die Bundesregierung will das Gesundheitswesen digitaler machen. Expert:innen begrüßten zwei Gesetzesvorhaben am Mittwoch im Gesundheitsausschuss. Die Themen Datenschutz und die Informationssicherheit kamen dabei allerdings nur am Rande vor.*

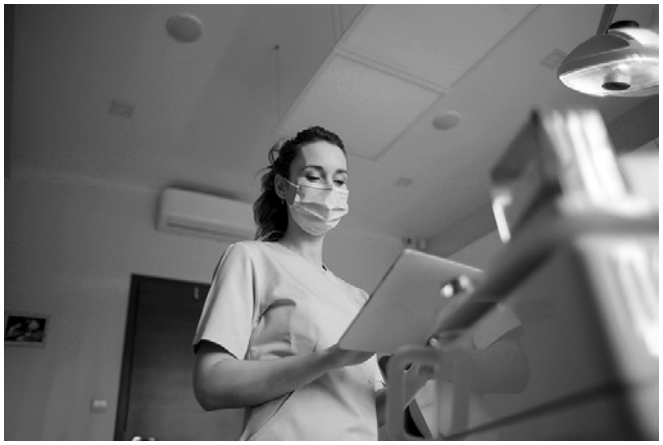
Digitale Gesundheitsdaten und IT-Sicherheit sollten stets Hand in Hand gehen. Denn diese Daten gelten als sensibel[1] und sind daher besonders sorgfältig zu schützen. Umso erstaunlicher ist es, dass das Thema Sicherheit am vergangenen Mittwoch im Gesundheitsausschuss nur eine marginale Rolle spielte. Dabei will die Bundesregierung in naher Zukunft die Gesundheitsdaten *von* rund 73 Millionen Bundesbürger:innen digitalisieren[2].

Zwei Gesetzesvorlagen standen an diesem Tag im Fokus: Das erste, das Digital-Gesetz (*DigiG*[3]), nimmt vor allem die digitale Gesundheitsversorgung in den Blick: die Elektronische Patientenakte (ePA) und das E-Rezept. Das Gesundheitsdatennutzungsgesetz (*GDNG*[4]) soll dagegen regeln, wie Gesundheitsdaten für die Forschung erschlossen werden. Der Bundestag hat beide Gesetzesvorlagen der Regierung in der vergangenen Woche zur weiteren Beratung an den Gesundheitsausschuss überwiesen[5].

Dieser hatte dann am Mittwoch gerade einmal zweieinhalb Stunden Zeit, um die Einschätzungen der geladenen Expert:innen anzuhören. Rund 60 schriftliche Stellungnahmen lagen den Ausschussmitgliedern für beide Gesetzesentwürfe vor.

Breite Zustimmung zum Opt-out bei der Patientenakte

Im ersten Teil[6] ging es um das Digital-Gesetz. Die geladenen Expert:innen[7] zeigten sich bereits in den zuvor eingereichten schriftlichen Statements weitgehend einig: Die elektronische Patientenakte (ePA), die ab Januar 2025 für alle Bundesbürger:innen kommen soll, bedeute einen „Wendepunkt“ und werde „einen Paradigmenwechsel einleiten“. Ferdinand Gerlach vom Institut für Allgemeinmedizin Frankfurt wertete die ePA gar als „Empowerment der Datensouveränität der Versicherten“.



Digitale Gesundheitsdaten und IT-Sicherheit?

Große Übereinstimmung herrschte ebenfalls beim geplanten Opt-out-Verfahren[8]. Auch der Verbraucherzentrale Bundesverband (vzbv) spricht sich „ausnahmsweise" für das nachträgliche Widerspruchsverfahren[9] aus, weil die ePA „eine große Bedeutung für den Erfolg der digitalen Transformation des Gesundheits- und Pflegewesens" habe. Umso wichtiger aber sei es, sagte Thomas Moormann vom vzbv, dass die Versicherten künftig „so einfach und feingranular wie möglich" der Nutzung widersprechen können. „Das ist das A und O", betonte Murmann.

Dem widersprach der Präsident der Bundesärztekammer (BÄK), Klaus Reinhardt. Er hält schon die geplante Regelung für wenig brauchbar. Demnach sei „die Ausgestaltung der Zugriffsverwaltung […] an manchen Stellen so kleinteilig gestaltet, dass die notwendige Praktikabilität und Überschaubarkeit für die Versicherten nicht gewährleistet ist", schreibt die BÄK in ihrer Stellungnahme[10]. Sie fordert, „das Widerspruchsrecht niedrigschwelliger zu gestalten". Patient:innen sollen nur darüber entscheiden können, ob sie einzelne Dokumente einer Ärzt:in vorgelegen oder nicht. Eine „Beschränkung des lesenden Zugriffs oder ein Verschatten von Informationen" sei nicht geboten, so die BÄK.

Umstrittene Beratung durch Krankenkassen

Besonders kritisch sahen viele Expert:innen den Vorschlag der Regierung, wonach Krankenkassen künftig auf Grundlage der ihnen vorliegenden Abrechnungsdaten eigenständig Risikoeinschätzungen vornehmen und die Versicherten individuell beraten sollen.

Die Bundespsychotherapeutenkammer (BPtK), der Deutsche Caritasverband, die Apothekerverbände und der vzbv lehnen diese Pläne strikt ab. Es sei „eine Kernaufgabe von Behandelnden", gesundheitliche Risiken zu beurteilen, betonte Nikolaus Melcop von der BPtK am Mittwoch. Der vzbv spricht sich für ein Opt-in-Verfahren bei dieser Form der Beratung aus[11]: „Eine Datenverarbeitung durch Kranken- und Pflegekassen und eine individuelle Ansprache dürfen nur nach vorheriger ausdrücklicher Einwilligung der Versicherten erfolgen."

Forschung nur fürs Gemeinwohl

Auch zum zweiten Vorhaben, dem Gesundheitsdatennutzungsgesetz, äußerten die Expert:innen Kritik. Es verfolgt das Ziel, Gesundheitsdaten zusammenzuführen und der Forschung zur Verfügung zu stellen. Auch die Datenfreigabe aus der Patientenakte zu Forschungszwecken soll über ein Widerspruchsverfahren geregelt werden. Die Daten würden dann aus der elektronischen Patientenakte weitergeleitet und mit weiteren Daten im Forschungsdatenzentrum (FDZ) verknüpft.

Jens Baas von der Techniker Krankenkasse sieht in dieser Zusammenführung eines der wichtigsten Gesetzesvorhaben der Legislaturperiode. „Das zu unterlassen, gefährdet Leben", so Baas. Bianca Kastl vom Innovationsverbund Öffentliche Gesundheit e. V. (InÖG) schrieb in ihrer Stellungnahme[12] hingegen, dass das Teilen von Daten, von dem vor allem privatwirtschaftliche Unternehmen statt die Gemeinschaft profitieren, Machtasymmetrien verstärken und sich daher aus Sicht der digitalen Zivilgesellschaft gegen das Gemeinwohl richten könnte.

Doris Pfeiffer vom GKV Spitzenverband plädierte dafür, dass der Einsatz von Gesundheitsdaten zu einer gemeinwohlorientierten „Forschungsrendite" beitragen müsse. Und auch der vzbv wirbt in seiner Stellungnahme[13] dafür, dass die Daten in der Forschung gemeinwohlorientierten Zwecken und damit – wie Alena Buyx, Vorsitzende des Deutschen Ethikrats, ergänzte – dem Wohle der Patient:innen dienen sollten.

Deutliche Kritik äußerten der Caritas Dachverband[14] und der vzbv auch an der fehlenden Unabhängigkeit des Forschungsdatenzentrums. Beide Verbände befürchten Interessenskonflikte, wenn sowohl die zentrale Datenzugangs- und Koordinierungsstelle als auch das Zentrum beim Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) angesiedelt sind, das wiederum das Bundesministerium für Gesundheit beaufsichtigt. Der vzbv fordert in seiner Stellungnahme[15], die zentrale Datenzugangs- und Koordinierungsstelle „schnellstmöglich" in eine eigenständige Institution zu überführen.

Hat hier jemand Sicherheitsbedenken?

Neben all diesen Themen kamen Sicherheitsbedenken in der Sitzung erheblich zu kurz. Als Einzige wies Bianca Kastl ausdrücklich darauf hin, dass zentrale Datensammlungen wie jene des FDZ das Risiko bergen, kompromittiert und angegriffen zu werden. Dieses Risiko sei mit Blick auf die Vertraulichkeit der Daten, ihre Verfügbarkeit und eine drohende Verfälschung gleichermaßen sehr hoch. „Eine Verlängerung der Speicherfrist im Forschungsdatenzentrum auf bis zu 100 Jahre […] führt zu einem Worst-Case-Szenario im Kontext der Informationssicherheit", warnt Kastl.

Die hohen Risiken könnten unter anderem durch eine stärkere Dezentralisierung der Daten, eine konsequente Pseudonymisierung sowie weitere „privatsphärenschonende Verfahren" minimiert werden. Das Vertrauen in die Datensysteme könnte außerdem dadurch gestärkt werden, dass die Systeme von unabhängiger Seite geprüft würden. „Es gilt hier in Anlehnung an die Petersberger Erklärung der [Datenschutzkonferenz][16]: Je sensibler die Daten, die verarbeitet werden, desto transparenter muss die Funktionsweise der verwendeten Systeme sein", schreibt Kastl.

„Höchst schwammig" definiert

Darüber hinaus bemerkt Kastl, dass in dem Entwurf „durchgängig nur von pseudonymisierten Daten gesprochen [wird]. Anonyme Daten werden nur in der Erläuterung erwähnt." Die Anwendung der Pseudonymisierung sei zudem „höchst schwammig" definiert. Entsprechende Verfahren müssten obendrein nur „im Benehmen" mit dem Bundesamt für Sicherheit in der Informationstechnik (BSI) und dem Bundesdatenschutzbeauftragten (BfdI) abgestimmt werden. Kastl plädiert dafür, dass diese Verfahren stattdessen „im Einvernehmen" mit BSI und BfDI beschlossen werden, die beiden Institutionen also nicht nur an der Verfahrensfindung mitwirken, sondern auch mitentscheiden.

Kastls Fazit fällt negativ aus. Die vom Bundesgesundheitsminister angestrebte Forschung sei „näher bei großen Forschungseinrichtungen denn bei den betroffenen Menschen". Ein besseres digitales Gesundheitswesen, das die Bürger:innen ins Zentrum stellt, sei möglich – „aber nicht mit diesem Entwurf".

Offenlegung: Bianca Kastl schreibt als Kolumnistin auf netzpolitik.org alle vier Wochen zur Digitalisierung der Verwaltung und des Gesundheitswesens.

Quelle: https://netzpolitik.org/2023/gesundheitsdigitalisierung-it-sicherheit-geraet-zur-randnotiz/

Anmerkungen

1 *https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_de*
2 *https://netzpolitik.org/2023/faq-zur-elektronischen-patientenakte-was-lauterbachs-plaene-fuer-aerztinnen-und-versicherte-bedeuten/*
3 *https://dserver.bundestag.de/btd/20/090/2009048.pdf*
4 *https://dserver.bundestag.de/btd/20/090/2009046.pdf*
5 *https://netzpolitik.org/2023/debatte-im-bundestag-abgeordnete-schwaermen-von-fliessenden-gesundheitsdaten/*
6 *https://www.bundestag.de/ausschuesse/a14_gesundheit/oeffentliche_anhoerungen/974674-974674*
7 *https://www.bundestag.de/resource/blob/978052/c8734ee2e5b-be434e441dacac4cc959f/Sachverstaendigenliste-data.pdf*
8 *https://netzpolitik.org/2023/gesundheitsdaten-opt-out-digitalisierung-ohne-ruecksicht-auf-versicherte/*
9 *https://www.bundestag.de/resource/blob/977252/2719441565def0078bf18cece5678a02/20_14_0163-14-_Verbraucherzentrale-Bundesverband_DigitalG-data.pdf*

Wer darf wann zugreifen?, Quelle: Mit KI erstellt

10 *https://www.bundestag.de/resource/blob/976840/3bc331567a109193ec455a5840772251/20_14_0163-11-_Bundesaerztekammer_DigitalG-data.pdf*
11 *https://www.bundestag.de/resource/blob/977262/fce3cea2371a57b05cd75f0d6c8be96b/20_14_0165-9-_Verbraucherzentrale-Bundesverband_GDNG-data.pdf*
12 *https://www.bundestag.de/resource/blob/977258/acdef80b55a7909fcc3cef32c1eac92b/20_14_0165-11-_Innovationsverbund-oeffentliche-Gesundheit-Bianca-Kastl_GDNG-data.pdf*
13 *https://www.bundestag.de/resource/blob/977262/fce3cea2371a57b05cd75f0d6c8be96b/20_14_0165-9-_Verbraucherzentrale-Bundesverband_GDNG-data.pdf*
14 *https://www.bundestag.de/resource/blob/976286/d0cc7dec299b614a651b59948c6aa3e7/20_14_0165-3-_Deutscher-Caritasverband-e-V-_GDNG-data.pdf*
15 *https://www.bundestag.de/resource/blob/977262/fce3cea2371a57b05cd75f0d6c8be96b/20_14_0165-9-_Verbraucherzentrale-Bundesverband_GDNG-data.pdf*
16 *https://www.bfdi.bund.de/SharedDocs/Downloads/DE/DSK/DSKEntschliessungen/104DSK-Petersberger-Erklaerung.pdf?__blob=publicationFile&v=2*
17 *https://www.blaetter.de/*
18 *http://www.schmetterling-verlag.de/page-5_isbn-3-89657-068-4.htm*
19 *https://www.alternativer-medienpreis.de/preistraeger-2016/daniel-leisegang/*
20 *https://www.eurozine.com/*
21 *mailto:daniel@netzpolitik.org*
22 *https://keys.openpgp.org/search?q=daniel@netzpolitik.org*
23 *https://mastodon.social/@dleisegang*
24 *https://twitter.com/dleisegang*

Daniel Leisegang

Daniel Leisegang ist Politikwissenschaftler und seit August 2022 Co-Chefredakteur bei *netzpolitik.org*. Zuvor war er Redakteur bei den *Blättern für deutsche und internationale Politik*[17]. 2014 erschien von ihm das Buch *Amazon – Das Buch als Beute*[18]; 2016 erhielt er für seinen Beitrag *Facebook rettet die Welt* den Alternativen Medienpreis[19] in der Rubrik *Medienkritik*. Daniel gehört dem Board of Trustees von Eurozine[20] an.
Kontakt: *E-Mail*[21] (OpenPGP[22]), *Mastodon*[23], *Twitter*[24], Telefon: +49-030-577148228 (Montag bis Freitag, jeweils 8 bis 18 Uhr).

Tomas Rudl

Es ging immer darum, Verschlüsselung zu umgehen

16. November 2023 – *Eine Gruppe von Expert:innen hat für die EU-Kommission Vorschläge erarbeitet, wie sich eine Chatkontrolle technisch umsetzen ließe. Dabei setzen die Vorschläge vor allem auf das so genannte Client-Side-Scanning, aber auch andere Formen der Überwachung verschlüsselter Kommunikation werden angedacht.*

Die sogenannte Chatkontrolle könnte womöglich nur ein erster Schritt sein, mit der die EU-Kommission stückweise verschlüsselte Kommunikation anzugreifen versucht. Langfristig könnten etwa zwischengeschaltete Server in die Inhalte von Nachrichten hineinschauen, um Darstellungen von Kindesmissbrauch oder sonstige illegale Inhalte aufzuspüren. Das geht aus Empfehlungen hervor, die eine Gruppe von Expert:innen im Rahmen des *EU Internet Forums* für die EU-Kommission erarbeitet hatte.

Quelle: Mit KI erstellt

Letzte Woche war bekannt geworden, dass die EU-Kommission sich sehr einseitig für ihre technische Folgenabschätzung zu dem umstrittenen Chatkotrolle-Gesetzentwurf[1] hat beraten lassen. Den Entwurf hat EU-Innenkommissarin Ylva Johansson vor über einem Jahr vorgelegt[2], er richtet sich unter anderem gegen die Verbreitung von Missbrauchsinhalten über das Internet, könnte aber einer anlasslosen Überwachung auch anderer Inhalte[3] Tür und Tor öffnen.

Besonders brisant ist an dem Vorschlag, gegebenenfalls auch private Nachrichten zu durchleuchten, die eigentlich mit Ende-zu-Ende-Verschlüsselung gesichert sind. Eine Technik dafür ist als *Client-Side-Scanning* (CSS) bekannt[4] und gleicht Inhalte mit Datenbanken ab, bevor sie verschlüsselt und versandt werden.

Das Vorhaben der EU-Kommission stand von Beginn an unter starker und bemerkenswert breiter Kritik[5], da damit eine neue Form anlassloser Massenüberwachung eingeführt und zugleich Verschlüsselung geschwächt würde. Kürzlich hat sich das EU-Parlament dagegen ausgesprochen[6], während sich die EU-Länder noch nicht auf eine gemeinsame Position[7] einigen konnten.

Expert:innen mit Schlagseite

Im Vorfeld hatte sich die EU-Kommission von über 30 Personen beraten lassen. Wie netzpolitik.org letzte Woche berichtete[8], hatte die Gruppe eine klare Schlagseite: Neben Vertreter:innen von Geheimdiensten und Polizeien hörte Johansson vor allem Expert:innen an, die mit Massenüberwachung nicht auf Kriegsfuß zu stehen scheinen.

Nach unserer Veröffentlichung wurde uns das gesamte Dokument zugespielt, das wir im Volltext veröffentlichen[9]. Darüber hatte bereits Politico im Jahr 2020 berichtet, dem damals veröffentlichten Dokument[10] fehlte jedoch die Liste der Expert:innen. Das undatierte Diskussionspapier gewährt auf 28 Seiten einen Einblick in die Denkweise der Kommission und welche Handlungsoptionen sie überhaupt in Betracht zieht.

EU-Kommission gibt Ziel klar vor

Von Anfang an ist klar: Es geht um die „proaktive Erkennung durch Unternehmen von Bildern, Videos und Text-basiertem Kindesmissbrauch wie Grooming oder Sextortion". Unter Grooming versteht man die Kontaktanbahnung Erwachsener zu Minderjährigen, Sextortion ist eine Form sexueller Erpressung. Das Papier beschränkt sich auf die Untersuchung von Messenger-Diensten und auf eine „spezifische Art illegaler Inhalte, auf Kindesmissbrauch".

Der Fokus auf derartiges Material erkläre sich unter anderem daraus, dass als solche erkannte Inhalte „unabhängig vom Kontext" seien, anders als etwa mutmaßlich terroristische Inhalte, heißt es in der Einleitung. Das ist eine bemerkenswerte Aussage, schließlich handelt es sich bei vielen Tatverdächtigen um Minderjährige[11] oder um sonstige Nutzer:innen, die aus völlig unverfänglichen Gründen[12] ins Visier von Online-Diensten und Polizeien geraten. Kontext, den automatisierte Werkzeuge nicht erfassen können, ist auch hier entscheidend.

Hauptproblem Verschlüsselung

Das Papier bewertet denkbare Ansätze nach fünf Kriterien: Effektivität, Machbarkeit, Privatsphäre, Sicherheit und Transparenz. Als Hauptproblem verortet es Ende-zu-Ende-verschlüsselte Kommunikation, die führende Messenger-Dienste wie Signal, WhatsApp oder iMessage seit geraumer Zeit standardmäßig einsetzen. „Gibt es irgendwelche technischen Lösungen, die die Erkennung von Missbrauchsmaterial erlauben, während sie die gleichen oder vergleichbare Vorteile von Verschlüsselung beibehalten?", fragen die Autor:innen.

Aus Sicht hunderter Wissenschaftler:innen und IT-Expert:innen[13] fällt die Antwort darauf nicht schwer: Techniken wie CSS würden

Ende-zu-Ende-Verschlüsselung und damit die Privatsphäre schwächen, der Abgleich mit Datenbanken voller digitaler Fingerabdrücke (*Hashes*) sei manipulierbar und nicht zuverlässig, und automatisierte Erkennung sowie damit verbundene Fehlerraten würden unnötig Ressourcen binden, die dann dem tatsächlichen Kampf gegen Kindesmissbrauch fehlen würden.

Solche Stimmen wollte die Kommission aber augenscheinlich nicht hören oder sie wurden im Bericht nicht merklich gewürdigt – so ist nicht auszuschließen, dass beispielsweise Facebooks Ex-Sicherheitschef oder manche Vertreter:innen von Google oder Microsoft, die mit am Tisch saßen, das Vorhaben kritisch sehen. Sie vertretende Wirtschaftsverbände wie CCIA[14] oder Eco[15] stellen sich nicht von ungefähr konsequent gegen die Aushöhlung von Verschlüsselung.

Bericht ermittelt „Top 3"-Ansätze

Der Bericht listet knapp ein Dutzend technisch denkbarer Ansätze auf, sie reichen vom Status Quo über Inhalteerkennung auf den Geräten beziehungsweise Servern bis hin zu relativ neuartigen Techniken wie der selektiven Überwindung sogenannter homomorpher Verschlüsselung[16]. Einige davon, etwa letztgenannte Technik, würden noch mehr Forschung benötigen, heißt es in der Zusammenfassung. Andere Ansätze wie uneingeschränkte Ende-zu-Ende-Verschlüsselung wie bisher seien für die Lösung des Problems nicht anwendbar.

Als „Top 3"-Ansätze gelten dem Papier zufolge: Inhalte werden vor ihrer Verschlüsselung auf den Geräten der Nutzer:innen in Hashes umgerechnet, während danach ein Server die Hashes mit einer Datenbank bereits gemeldeter Inhalte abgleicht. Das entspricht praktisch dem CSS-Ansatz. Zweitens in Frage käme eine ähnliche Technik, bei der die Hashes teils auf den Geräten selbst und teils auf Servern ermittelt würden. Und drittens könnten die Diensteanbieter oder Dritt-Anbieter speziell gesicherte Server, sogenannte „Secure Enclaves", in die Kommunikation zwischenschalten. Diese hätten dann vollen Zugriff auf den Klartext derartig „verschlüsselter" Inhalte.

In ihrem Gesetzentwurf[17] hat sich die EU-Kommission nicht auf eine genaue Technik festgelegt. Stattdessen könnten die Anbieter beliebige Technologien einsetzen, solange sie die Auflagen erfüllen und unter anderem „wirksam zur Erkennung der Verbreitung von bekannten oder neuen Darstellungen sexuellen Kindesmissbrauchs oder der Kontaktaufnahme zu Kindern" beitragen. Alternativ könnten sie kostenlos Techniken nutzen, die ihnen ein noch einzurichtendes EU-Zentrum zur Verfügung stellen würde.

Der Bericht spricht aber klare Empfehlungen aus: Sofort wäre CSS einsetzbar, notfalls könnte ein teilweises Erstellen von Hashes auf Servern stattfinden. Langfristig sollte die EU aber Geld in Forschung stecken, rät das Papier – das in der vorliegenden Fassung noch nicht vollständig finalisiert war. Erfolgsversprechend seien hierbei etwa „Secure Enclaves" bei den Anbietern oder der Einsatz sogenannter Klassifikatoren (*classifiers*). Diese könnten mit Hilfe von Machine Learning selbstständig verdächtige Inhalte aufspüren und an Behörden melden, so die Zukunftsvision.

Klar wird durch das Papier in jedem Fall eines: Von Anfang an stand die Einführung einer anlasslosen Massenüberwachung im Fokus aller Bemühungen.

Quelle: https://netzpolitik.org/2023/chatkontrolle-es-ging-immer-darum-verschluesselung-zu-umgehen/

Anmerkungen

1 *https://netzpolitik.org/2023/geheime-liste-wie-der-sicherheitsapparat-die-chatkontrolle-praegt/*
2 *https://netzpolitik.org/2022/gesetz-gegen-kindesmissbrauch-eu-kommission-will-private-nachrichten-durchleuchten/*
3 *https://netzpolitik.org/2023/ueberwachung-politiker-fordern-ausweitung-der-chatkontrolle-auf-andere-inhalte/*
4 *https://netzpolitik.org/2021/client-side-scanning-beruehmte-it-sicherheitsforscherinnen-warnen-vor-wanzen-in-unserer-hosentasche/*
5 *https://netzpolitik.org/2023/ungeahnte-allianzen-so-breit-ist-der-widerspruch-gegen-die-chatkontrolle/*
6 *https://netzpolitik.org/2023/ueberwachung-eu-innenausschuss-stimmt-fuer-die-ablehnung-der-anlasslosen-chatkontrolle/*
7 *https://netzpolitik.org/2023/rat-der-eu-chatkontrolle-abstimmung-zum-zweiten-mal-vertagt/*
8 *https://netzpolitik.org/2023/geheime-liste-wie-der-sicherheitsapparat-die-chatkontrolle-praegt/*
9 *https://cdn.netzpolitik.org/wp-upload/2023/11/2020-06-30_DG-HOME_CSA_E2EE_discussion-paper.pdf*
10 *https://www.politico.eu/wp-content/uploads/2020/09/SKM_C45820090717470-1_new.pdf*
11 *https://netzpolitik.org/2022/strafrecht-die-meisten-tatverdaechtigen-bei-kinderpornografie-sind-minderjaehrig/*
12 *https://netzpolitik.org/2022/falscher-verdacht-gegen-vater-ein-fall-aus-den-usa-zeigt-die-gefahr-der-geplanten-chatkontrolle/*
13 *https://netzpolitik.org/2023/offener-brief-der-wissenschaft-das-client-side-scanning-ist-zum-scheitern-verurteilt/*
14 *https://ccianet.org/news/2023/11/fighting-csa-european-parliament-position-marks-significant-improvement/*
15 *https://netzpolitik.org/2023/chatkontrolle-internetwirtschaft-fordert-schutz-von-verschluesselung/*
16 *https://de.wikipedia.org/wiki/Homomorphe_Verschlüsselung*
17 *https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52022PC0209*
18 *https://keys.openpgp.org/vks/v1/by-fingerprint/CA052285DC96CFC89E980514745121858AE13AED*

Tomas Rudl

Tomas Rudl ist in Wien aufgewachsen, hat dort für diverse Provider gearbeitet und daneben Politikwissenschaft studiert. Seine journalistische Ausbildung erhielt er im Heise-Verlag, wo er für die Mac & i, c't und Heise Online schrieb. Er ist unter+49 30 577148268 oder *tomas@netzpolitik.org* (PGP-Key[18]) erreichbar und twittert mal mehr, mal weniger unter *@tomas_np*.

Thomas Rudl

Hessen auf Hardliner-Kurs

13. November 2023 – *Die CDU will in Hessen künftig mit der SPD regieren. In einem Eckpunktepapier skizziert die Große Koalition ihre Prioritäten: Dazu zählen mehr Videoüberwachung mit Gesichtserkennung, die Vorratsdatenspeicherung von IP-Adressen, Staatstrojaner und mehr Daten für Palantirs HessenData.*

Eine „christlich-soziale Hessenkoalition“[1] wollen sie sein – und sie versprechen mehr Videoüberwachung, mehr Staatstrojaner und mehr Big-Data-Unterstützung für die Polizei. Was die CDU mit ihrem nunmehrigen Ex-Partner, den Grünen, nicht umsetzen konnte, fällt mit dem neuen Juniorpartner SPD offenbar leicht: Nach einem erst im Sommer noch unter schwarz-grün verabschiedeten Sicherheitspaket[2] zeichnet sich eine erneute Ausweitung der Befugnisse für Ermittlungsbehörden ab.

Ende letzter Woche verkündete die regierende CDU[3], Koalitionsgespräche mit den hessischen Sozialdemokraten aufzunehmen, die Eckpunkte[4] der Zusammenarbeit stehen bereits fest. Beginnen sollen die Gespräche morgen und bis Ende Dezember abgeschlossen sein, Mitte Januar nimmt der neue Landtag seine Arbeit auf. Damit dürfte die zehnjährige Regierungsbeteiligung der Grünen in Hessen zu Ende gehen. Ähnlich wie die SPD waren die Grünen bei der Landtagswahl im Oktober auf rund 15 Prozent abgesunken, während die Union auf knapp 35 Prozent zugelegt hat – und sich so bequem aussuchen kann, mit wem sich ihr Programm am besten umsetzen lässt.

Lauschangriff und mehr Videoüberwachung

„Wir schaffen verbesserte Rahmenbedingungen für Videoüberwachung“, heißt es in den Eckpunkten der designierten Großen Koalition. Im Blick hat sie dabei „vereinfachte Zulassung, Erweiterung um Akustik, Mustererkennung, zielgerichtete Fahndung via Gesichtserkennung“. Damit fasst sie gleich mehrere heiße Eisen an: Neben einem Lauschangriff und „intelligenter“ Videoüberwachung, die automatisch angeblich verdächtige Bewegungen erkennen[5] soll, will sie augenscheinlich auch die besonders umstrittene Gesichtserkennung[6] flächendeckend einführen.

Ferner sollen Fahndungsmöglichkeiten ausgeweitet und „in besonderen Fällen und auf richterlichen Beschluss de[r] Zugang zu audiovisuellen Systemen“ ermöglicht werden, schreiben die künftigen Koalitionäre. In Kombination mit mehr Videokameras im öffentlichen Raum könnte so ein engmaschiges Überwachungsnetz entstehen, dem sich kaum jemand entziehen kann. Außerdem wollen sie die Quellen-TKÜ (Staatstrojaner) für den Verfassungsschutz einführen. Der von Skandalen geplagte hessische Inlandsgeheimdienst[7] dürfte dann staatliche Spionage-Software auf den IT-Geräten von Verdächtigen aufspielen.

Erweiterter Blick in die Big-Data-Glaskugel

Erheblich mehr Material soll es für die Analyse-Software HessenData des US-Anbieters Palantir geben. Dazu soll der Straftatenkatalog erweitert sowie die „Nutzung von IP-/Maut-/ Verkehrsüberwachungsdaten“ ermöglicht werden, zudem soll sich der „Datenaustausch zwischen Sicherheitsbehörden und anderen Behörden“ einfacher gestalten. Die Big-Data-Anwendung verknüpft bereits heute umfangreiche Datenbestände[8] und soll menschlichen Polizist:innen dabei helfen, verborgene Zusammenhänge zu erkennen.

Allerdings steht der Ansatz rechtlich auf tönernen Füßen. Anfang des Jahres hatte das Bundesverfassungsgericht dem Einsatz solcher Techniken, die große Mengen an Daten mit Hilfe sogenannter Künstlicher Intelligenz (KI) durchforsten und auswerten, enge technische und rechtliche Grenzen gesetzt[9]. Die als Reaktion auf das Urteil in Karlsruhe eilig auf den Weg gebrachte Gesetzesänderung hatte kürzlich noch für Irritation bei der SPD gesorgt[10], das Gefühl scheint inzwischen verflogen: „Den Einsatz von KI zur automatisierten Auswertung großer Datenmengen und zur Erkennung von Hate-Speech im Netz werden wir ermöglichen.“

Neuer Anlauf für Vorratsdatenspeicherung

Den Rücken stärkt Schwarzrot ausgerechnet der Bundesinnenministerin Nancy Faeser, die als Spitzenkandidaten für die hessische SPD in den Wahlkampf gezogen und nun mangels Erfolg wieder auf die Bundesbühne zurückgekehrt ist. „Im Bundesrat werden wir einen Gesetzesentwurf zur IP-Adressspeicherung einbringen“, heißt es im Eckpunktepapier.

Damit geht die scheinbar endlose Debatte rund um die Vorratsdatenspeicherung[11] in die nächste Runde. Eigentlich hatte die Ampelregierung festgeschrieben, Daten nur „rechtssicher anlassbezogen und durch richterlichen Beschluss“ für polizeiliche Ermittlungen nutzen zu wollen. Passend dazu legte das Justizministerium einen Entwurf für eine Quick-Freeze-Lösung[12] vor. Die liegt allerdings seit über einem Jahr auf Eis – auch, weil sich Faesers Innenministerium weiterhin für eine anlasslose Vorratsdatenspeicherung für IP-Adressen stark[13] macht.

Auf Hardliner-Kurs will sich die Große Koalition auch beim Thema Migration begeben. Sie bekennt sich zur „Begrenzung der Migration und dem Schutz der europäischen und deutschen Außengrenzen“, will eine „Rückführungsoffensive“ starten und dabei „alle rechtsstaatlichen Möglichkeiten ausschöpfen“. Wer bleiben darf, muss sich einer „Wohnsitzauflage/Residenzpflicht“ beugen. Geflüchtete sollen auch keine monetäre Auszahlungen[14] mehr erhalten, stattdessen will Hessen „konsequent auf Bezahlkarten und Sachleistungen umstellen“.

Quelle: https://netzpolitik.org/2023/grosse-koalition-hessen-auf-hardliner-kurs/

Anmerkungen

1 *https://www.cduhessen.de/aktuelles/verhandlungsstart-fuer-christlich-soziale-hessenkoalition/*
2 *https://www.hessenschau.de/politik/landtag/cdu-und-gruene-setzen-umstrittenes-sicherheitspaket-fuer-hessen-durch-v1,landtag-sicherheitspaket-hessendata-100.html*
3 *https://www.hessenschau.de/politik/landtag/ein-neues-kapitel-cdu-will-hessen-mit-der-spd-regieren-v14,koalition-verhandlungen-hessen-100.html*
4 *https://www.cduhessen.de/data/documents/2023/11/10/2835-654e288a21fe2.pdf*
5 *https://netzpolitik.org/2023/intelligente-videoueberwachung-polizei-hamburg-will-ab-juli-verhalten-automatisch-scannen/*
6 *https://netzpolitik.org/2021/kuenstliche-intelligenz-eu-datenschuetzer-fordern-verbot-von-gesichtserkennung-im-oeffentlichen-raum/*
7 *https://www.fr.de/rhein-main/landespolitik/skandal-auf-skandal-in-hessen-91443925.html*
8 *https://netzpolitik.org/2019/big-data-bei-der-polizei-hessen-sucht-mit-palantir-software-nach-gefaehrdern/*
9 *https://netzpolitik.org/2023/urteil-des-bundesverfassungsgerichts-automatisierte-datenanalyse-fuer-die-vorbeugende-bekaempfung-von-straftaten-ist-verfassungswidrig/*
10 *https://www.fr.de/rhein-main/landespolitik/hessen-breite-kritik-an-reform-zur-hessendata-software-92360685.html*
11 *https://netzpolitik.org/tag/vorratsdatenspeicherung/*
12 *https://netzpolitik.org/2022/quick-freeze-buschmann-legt-alternative-zur-vorratsdatenspeicherung-vor/*
13 *https://netzpolitik.org/2023/polizeiliche-kriminalstatistik-faeser-wirbt-fuer-vorratsdatenspeicherung/*
14 *https://netzpolitik.org/2023/bezahlsysteme-fuer-gefluechtete-karten-der-abschreckung/*

Information zum Autor siehe Seite 60

Daniel Leisegang

Digitale Brieftasche mit Ausspähgarantie

9. November 2023 – *Jetzt steht es fest: Die europäische digitale Brieftasche kommt. Aus Sicht von Beobachtern bringt der im Trilog erzielte Kompromiss etliche Verbesserungen im Vergleich zum ursprünglichen Kommissionsentwurf. Bürgerrechtsgruppen und Datenschützer:innen warnen jedoch davor, dass Staaten durch die Wallet eine „panoptische Vogelperspektive" erhielten.*

„Wir haben es geschafft", jubilierte[1] EU-Kommissar Thierry Breton. Und Nadia Calviño, die Vize-Premierministerin Spaniens, deren Land derzeit den Vorsitz im Rat der EU innehat, versprach[2], dass die Einigung zur eIDAS-2.0-Verordnung „den Bürgerinnen und Bürgern in der neuen digitalen Welt eine Identität geben und unseren Binnenmarkt vertiefen wird – zum Nutzen der Innovation, der Privatsphäre, der Sicherheit und der Europäischen Union".

Den lang erhofften „Deal" schlossen EU-Kommission, der Ministerrat und das EU-Parlament am Mittwoch. In einem finalen politischen Trilog haben sie sich auf einen Kompromisstext geeinigt. Damit geht das größte digitalpolitische Projekt der Europäischen Union nun in die Phase der Umsetzung.

Die eIDAS-Reform legt das rechtliche Fundament für die sogenannte „European Digital Identity Wallet" (ID-Wallet). Demnach müssen bis zum Jahr 2026 alle 27 EU-Mitgliedstaaten ihren Bürger:innen eine digitale Brieftasche anbieten, mit der sie sich dann on- wie offline und in fast allen Lebensbereichen ausweisen können.

Zwei Jahre Verhandlungen

Clemens Schleupner, Referent für Vertrauensdienste & Digitale Identitäten beim Digitalverband Bitkom begrüßt die Einigung: „ID Wallets können sowohl Kosten sparen, indem Identifizierungsprozesse – zum Beispiel bei Banken – schneller und kostengünstiger durchgeführt werden können. Grundsätzlich bieten ID Wallets eine Vielzahl von Möglichkeiten, die heute noch nicht vollständig abzuschätzen sind, und zwar branchenübergreifend."

Den entsprechenden Verordnungsentwurf dafür hat die Kommission im Juni 2021 vorgelegt[3]. Er soll die eIDAS-Verordnung aus dem Jahr 2014 reformieren, die den sicheren Zugang zu öffentlichen Diensten sowie die Durchführung von Online-Transaktionen und grenzüberschreitenden Transaktionen in der EU regelt.

Datenschützer:innen und Bürgerrechtler:innen kritisierten das Vorhaben von Beginn an aus zwei zentralen Gründen. Zum einen drohe die Reform eine technische Infrastruktur zu schaffen, die es theoretisch ermöglicht, EU-Bürger:innen on- wie offline massenhaft zu identifizieren und zu überwachen. Zum anderen könnten nicht nur öffentliche, sondern auch private Stellen – also etwa Unternehmen – die Wallet einsetzen und ihre Kunden damit potentiell umfassend ausspähen.

Einige der Risiken, die der Ursprungsentwurf barg, wurden im Zuge der zurückliegenden Verhandlungen minimiert oder ausgeräumt. Andere bestehen aus Sicht von Bürgerrechtler:innen weiterhin – und sind so groß, dass Kritiker:innen eindringlich vor einer Nutzung der ID-Wallet warnen.

Der gelöschte Super-Cookie

Über die fortbestehenden Probleme können die erzielten Verbesserungen nicht hinwegtäuschen. Zu Letzteren zählt vor allem die Streichung einer eindeutigen, dauerhaften Personenkennziffer (Unique identifier). Dieser „Super-Cookie" hätte aus Sicht von Datenschützer:innen und Bürgerrechtler:innen geradezu zum Tracking und Profiling eingeladen. Eine solche Kennziffer soll nun nur noch optional bei grenzüberschreitenden Verwaltungsvorgängen zum Einsatz kommen.

Analoge Papiere am Scheideweg?, Quelle: Mit KI erstellt

Stattdessen sollen sich die Wallet-Nutzer:innen im Alltag allein mit ihren personenbezogenen Daten, einem Pseudonym oder einem sogenannten Zero Knowledge Proof (zu Deutsch: Null-Wissen-Beweis) identifizieren. Damit können die Nutzer:innen ihre Identität bestätigen, ohne persönliche Informationen über sich preiszugeben.

Allerdings hat die Identifizierung nach wie vor einen Haken: Denn das Recht auf Pseudonymität kann laut dem Kompromiss durch nationales oder durch EU-Recht eingeschränkt werden. Und der Zero Knowledge Proof findet sich im beschlossenen Kompromissentwurf nur als Forderung in den erläuternden Recitals (Erwägungsgründen) der Verordnung und stellt für die EU-Mitgliedstaaten damit keine Verpflichtung dar.

Schutz vor Diskriminierung und Verlinkung

Immerhin: Wer sich gegen den Einsatz der ID-Wallet entscheidet, soll keine Nachteile erleiden. Der Kompromissentwurf sieht – anders als der Kommissionsvorschlag – explizit einen Schutz vor Diskriminierung für Menschen vor, die sich gegen eine Nutzung entscheiden.

Jene, die die Wallet nutzen, sollen nachvollziehbar und transparent darüber bestimmen können, welche Daten sie etwa gegenüber sogenannten *relying parties* preisgeben und welche nicht. Diese vertrauenswürdigen Parteien, gegenüber denen Nutzer:innen ihre Identität bestätigen, müssen sich vorab in den jeweiligen EU-Mitgliedstaaten registrieren und darlegen, welche Daten sie zu welchem Zweck von den Nutzer:innen anfordern werden. Über ein sogenanntes Datenschutzcockpit können die Nutzer:innen einsehen, welche Daten von ihnen abgefragt und geteilt wurden – und gegebenenfalls Beschwerden einreichen.

Außerdem legt der Kompromiss fest, dass verschiedene Identifikationsvorgänge nicht miteinander verknüpft werden dürfen. Der Schutz der Transaktionsdaten war bis zum Tag der Einigung umkämpft. Transaktionsdaten zeigen an, wann, wie und wo Nutzer:innen die Wallet einsetzen, sie bilden also das konkrete Nutzungsverhalten ab.

Artikel 6a zufolge muss die ID-Wallet dafür „Techniken zur Wahrung der Privatsphäre ermöglichen, die die Unverknüpfbarkeit gewährleisten, wenn die Bescheinigung von Attributen keine Identifizierung des Nutzers erfordert." Konkret heißt das: Kauft eine Person also beispielsweise in dem gleichen Geschäft wiederholt Alkohol und belegt dabei mit ihrer ID-Wallet ihr Alter, dann kann das Unternehmen die unterschiedlichen Vorgänge nicht miteinander verknüpfen, um so das Kaufverhalten dieser Person über eine längere Zeitspanne zu tracken.

QWACS: Zertifizierte Unsicherheit

Mindestens ebenso umstritten waren auch die Vorgaben zu Zertifikaten. Bis zuletzt hatten die Trilog-Partner um Artikel 45 und damit um die Frage gerungen, ob die Verordnung Browseranbieter dazu verpflichten soll, bestimmte qualifizierte Zertifikate (QWACs) zu akzeptieren. Bereits Artikel 22 der bestehenden eIDAS-Verordnung verpflichtet Mitgliedstaaten dazu[4], vertrauenswürdige Listen von qualifizierten Vertrauensdiensteanbietern zu erstellen, zu führen und zu veröffentlichen[5]. Der Kompromisstext sieht darüber hinaus vor, dass Browser die in dem Zertifikat bescheinigten Identitätsdaten „in einer benutzerfreundlichen Weise" anzeigen, wenn diese eine bestimmte Webseite besuchen.

Zertifikate sollen im Internet die Authentizität, Integrität und Vertraulichkeit der Kommunikation sicherstellen. Sie werden in der Regel von sogenannten Trusted Root Certificate Authorities ausgegeben. Das sind Unternehmen oder Körperschaften, die Zertifikate nach strengen Prüfverfahren validieren, ausstellen und widerrufen. Diesen funktionierenden selbstregulierten Genehmigungsmechanismus umgeht die EU durch die staatlichen Listen[6] – mit womöglich dramatischen Folgen.

Staatliche Behörden könnten, so argumentieren Kritiker:innen, die Zertifikate dazu missbrauchen, um Webseiten zu kompromittieren und so die Internetkommunikation potentiell aller EU-Bürger:innen auszuspähen. Länder wie Kasachstan[7], China[8] und Russland[9] haben dies in der Vergangenheit getan. Entsprechend alarmiert hatten sich sowohl Bürgerrechtsorganisationen[10], IT-Sicherheitsexpert:innen[11] und Entwickler:innen[12] im Vorfeld der gestrigen Entscheidung geäußert.

Der Kompromiss sieht zwar vor, dass Browseranbieter selbst darüber entscheiden können, auf welche Weise sie Domains authentifizieren und die Internetkommunikation verschlüsseln. Wörtlich heißt es in einer Ergänzung zum Recital 32 bezüglich QWACs: „Die Verpflichtung zur Anerkennung, Interoperabilität und Unterstützung qualifizierter Zertifikate für die Website-Authentifizierung berührt nicht die Freiheit der Webbrowser, die Sicherheit des Internets, die Domänenauthentifizierung und die Verschlüsselung des Internetverkehrs in einer Weise und mit der Technologie zu gewährleisten, die diese für am besten geeignet halten."

Allerdings findet sich diese Ergänzung in den Erwägungsgründen der Verordnung, aus denen sich keine bindenden Rechtsfolgen ableiten lassen. Die Hoffnung ruht daher nun auf Browseranbietern wie Mozilla, die sich frühzeitig gegen QWACs ausgesprochen haben[13]. Auch das Sicherheitsteam von Chrome kritisiert die Pläne der EU[14]. Sie befürchten, dass ihnen mit den QWACs eine schwache Verschlüsselung aufgezwungen wird. Die Anbie-

ter könnten nun dazu übergehen[15], jeweils zwei verschiedene Varianten ihrer Browser anzubieten: eine unsichere für die EU und eine intakte für den Rest der Welt.

Blankoscheck zur Online-Überwachung

Die eIDAS-2.0-Reform würde es staatlichen Behörden aber nicht nur ermöglichen, den Internetverkehr auszuspähen, sondern sie könnten theoretisch auch die Wallets ihrer Bürger:innen einsehen.

Technisch wäre dies leicht zu verhindern gewesen. Das EU-Parlament hatte in seiner Position zum Verordnungsentwurf gefordert[16], die Wallet so zu gestalten, dass Transaktionsdaten nicht zentral erfasst werden. Der nun vereinbarte Kompromiss sieht nur eine logische Trennung dieser Daten vor. Mit den entsprechenden Befugnissen ist ein Zugriff und damit eine Zusammenführung der Daten theoretisch möglich.

Thomas Lohninger von der Bürgerrechtsorganisation epicenter.works zieht daher ein negatives Resümee der gestrigen Einigung: „Leider war der Zeitdruck der Verhandler am Ende stärker als ihre Sorgfalt. Bei diesem wichtigen Thema könnten wir das noch alle bereuen." Trotz der Verbesserungen, die in den vergangenen Monaten erzielt worden seien[17], warnt Lohninger vor dem neuen System: „Alles was man darüber tut, kann von staatlicher Seite eingesehen werden. Nachdem die Wallet in allen Lebensbereichen eingesetzt werden soll, hat der Staat damit die panoptische Vogelperspektive auf alles, was die Bevölkerung mit der Wallet macht", so Lohninger.

„Diese Verordnung ist ein Blankoscheck zur Online-Überwachung der Bürger und gefährdet unsere Privatsphäre und Sicherheit im Internet", kritisiert auch Patrick Breyer, Abgeordneter der Piratenpartei im Europäischen Parlament. „Dieser Deal opfert unverzichtbare Anforderungen im Verhandlungsmandat des Europäischen Parlaments, die die ID-Wallet datenschutzfreundlich und sicher gemacht hätten."

Eine Wallet für alle EU-Bürger:innen bis 2030

Diese Befürchtungen lassen sich im weiteren legislativen Prozess wohl kaum noch ausräumen. Nach der gestrigen Entscheidung wird es noch ein „technisches Treffen" geben, wo der Kompromisstext juristisch bereinigt wird. Wesentliche Änderungen sind dabei nicht mehr geplant.

Im Anschluss daran werden Rat und Parlament die Verordnung formell verabschieden. Der Rat tut dies planungsgemäß noch im Dezember dieses Jahres; das Parlament stimmt voraussichtlich am 28. November im ITRE-Ausschuss und im Februar 2024 im Plenum ab. Die Verordnung könnte dann frühestens im nächsten Frühjahr in Kraft treten. Bis zum Herbst 2026 müssten alle Mitgliedstaaten ihren Bürger:innen dann eine ID-Wallet anbieten. Geht es nach der Kommission, sollen alle EU-Bürger:innen bis 2030 über eine digitale Identität verfügen.

Parallel zum legislativen Entscheidungsprozess werden aber noch etliche technische Fragen geklärt. Hier lassen sich möglicherweise noch einige der offenen Schlupflöcher schließen. Fest steht bereits, dass die ID-Wallets der jeweiligen Mitgliedstaaten auf der gleichen technischen Architektur (Architecture Reference Framework) basieren sollen, um so EU-weit genutzt werden zu können. Die Details dafür erarbeitet eine technische Arbeitsgruppe.

Der Bundestagsabgeordnete Markus Reichel (CDU) blickt mit Spannung auf deren Tätigkeit in den kommenden Monaten und erwartet von der Bundesregierung ein starkes Bekenntnis zum Datenschutz: „Ohne die explizite Einwilligung des Nutzers sollten keine Daten geteilt werden können. Genau dafür muss sich die Bundesregierung in den Verhandlungen einsetzen. In der technischen Umsetzung muss die Unbeobachtbarkeit und die Unverknüpfbarkeit der Nutzungsdaten großgeschrieben werden", so Reichel. Der Erfolg der Wallet hänge von der Sicherheit und dem Datenschutz ab. „Umso wichtiger ist es, diesen für sichere digitale Identitäten auch in der Architektur zu gewährleisten", sagt der Abgeordnete.

Auch Clemens Schleupner vom Branchenverband bitkom erhofft sich noch Verbesserungen: „Es gibt noch zu viel Interpretationsspielraum für die Mitgliedsstaaten bei technischen Ausgestaltungen wie zum Beispiel bei der Frage, ob nur der Staat oder auch die Privatwirtschaft Wallets herausgeben darf." Langfristig könne die europäische digitale Brieftasche den Anstoß dafür geben, verschiedene Sektoren konsequent zu digitalisieren, vor allem die Verwaltung.

Quelle: https://netzpolitik.org/2023/eidas-reform-digitale-brieftasche-mit-ausspaehgarantie/

Anmerkungen

1 *https://twitter.com/ThierryBreton/status/1722291616654766485*
2 *https://x.com/eu2023es/status/1722285593869099443*
3 *https://netzpolitik.org/2021/digitale-brieftasche-eu-kommission-plant-europaweite-eid/*
4 *https://digital-strategy.ec.europa.eu/en/policies/eu-trusted-lists*
5 *https://eidas.ec.europa.eu/efda/home*
6 *https://scotthelme.co.uk/what-the-qwac/*
7 *https://www.golem.de/news/https-kasachstan-nutzt-erneut-ueberwachungszertifikat-2012-152617.html*
8 *https://www.golem.de/news/cnnic-google-wirft-chinesische-zertifizierungsstelle-raus-1504-113301.html*
9 *https://www.heise.de/news/Web-Verschluessleung-Russland-stellt-nun-eigene-TLS-Zertifikate-aus-6546463.html*
10 *https://www.eff.org/deeplinks/2023/11/article-45-will-roll-back-web-security-12-years*
11 *https://netzpolitik.org/2023/eidas-trilog-hunderte-wissenschaftlerinnen-und-dutzende-ngos-warnen-vor-massenueberwachung/*
12 *https://scotthelme.co.uk/what-the-qwac/*
13 *https://securityriskahead.eu/*
14 *https://security.googleblog.com/2023/11/qualified-certificates-with-qualified.html*
15 *https://www.eff.org/deeplinks/2023/11/article-45-will-roll-back-web-security-12-years*
16 *https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/ITRE/DV/2023/02-09/05_CA_eIDAS_EN.pdf*
17 *https://epicenter.works/en/content/eu-digital-identity-reform-the-good-bad-ugly-in-the-eidas-regulation*

Information zum Autor siehe Seite 58

Mitgliederversammlung (MV) des FIfF

Berlin, 5. November 2023, 12:05 – 15:15 Uhr

– Beschlussprotokoll –

Sitzungsleitung: Stefan Hügel als Vorsitzender des FIfF

1. **Begrüßung und Feststellung der Beschlussfähigkeit und der Protokollführung**
   Zur Versammlung wurde satzungsgemäß eingeladen. Sie ist damit beschlussfähig.
   Protokollführung: Werner Winzerling

2. **Beschlussfassung über Tages- und Geschäfts- und Wahlordnung**
   Der Tagesordnung wurde in der vorliegenden Form mit der Ergänzung zugestimmt.
   Geschäfts- und Wahlordnung werden von der MV in bekannter Form genehmigt.
   Ergänzung der Wahlordnung: Den Kandidaten wird die Möglichkeit gegeben, sich in 3 Minuten selbst vorzustellen

3. **Bericht des Vorstandes einschl. Kassenbericht**
   Stefan Hügel berichtet über die Arbeit des FIfF seit der letzten MV am 23.10.2022 sowie über den Haushalt mit Stand 31.10.2023.
   Außerdem berichteten einige Mitglieder über ihre Aktivitäten.
   Es gab keine Nachfragen zu dem Bericht des Vorstands.
   Es wurden keine Beschlüsse gefasst.

4. **Bericht der Kassenprüfer**
   Für die am 10.05.2023 in Bremen durchgeführte Kassenprüfung durch Margita Zallmann und Berthold Hoffmann berichtet Margita Zallmann der MV. Es wurden keine Beschlüsse gefasst. Aus dem Kassenprüfungsprotokoll:
   „Dem Vorstand wird eine dem Vereinszweck entsprechende, ordnungsgemäße Kassenführung bescheinigt. Einer Entlastung des Vorstandes steht nach unserer Auffassung nichts entgegen."

5. **Diskussion der Berichte**
   Es wurden keine Beschlüsse gefasst.

6. **Entlastung des Vorstandes**
   Die Kassenprüfer schlagen die Entlastung des Vorstandes vor.
   Die MV entlastet den Vorstand einmütig bei 5 Enthaltungen.

7. **Neuwahl des Vorstandes**
   Wahlleiter Eberhard Zehendner wird bei der Durchführung der Wahl unterstützt von Ingrid Schlagheck, Jens Rinne und Birgit Ahlmann.

   Es sind **26** stimmberechtigte Mitglieder anwesend.

   Die MV stimmt dem Vorschlag für den Wahlleiter zu.

   Wahlgang:
   Vorsitzender: Stefan Hügel (einziger Kandidat)
   Abgegebene gültige Stimmen: 26 (ja/nein/enth.): 26/0/0
   Wahl angenommen

   Stellvertretender Vorsitzender: Rainer Rehak (einziger Kandidat)
   Abgegebene gültige Stimmen: 26 (ja/nein/enth.): 25/0/1
   Wahl angenommen

   Wahl der Beisitzerinnen und Beisitzer

Kandidat	Gültige Stimmen	(ja/nein/enth.)	Wahl angenommen?
Michael Ahlmann	24	21/2/1	Ja
Gilbert Assaf	24	23/0/1	Ja
Alexander Heim	24	23/0/1	Ja
Sylvia Johnigk	24	23/1/0	?
Hans-Jörg Kreowski	24	22/0/2	Ja
Kai Nothdurft	24	24/0/0	?
Britta Schinzel	24	24/0/0	?
Friedrich Strauß	24	22/2/0	Ja
Werner Winzerling	24	15/5/4	Ja
Margita Zallmann	24	21/0/3	Ja

8. **Neuwahl der Kassenprüfer**
   Die MV wählt im Block einmütig zu den neuen Kassenprüfern des FIfF: Hui Shi (hat Zustimmung in Falle der Wahl vorher erklärt) und Berthold Hoffmann (hat Zustimmung in Falle der Wahl vorher erklärt)
   Abstimmung einstimmig dafür

9. **Diskussion über Ziele und Arbeit des FIfF, aktuelle Themen, Verabschiedung von Stellungnahmen, Berichte aus den Regionalgruppen**
   Über aktuelle Projekte des FIfF und aus Regionalgruppen wurden berichtet.

   Es wurden keine Beschlüsse gefasst.

10. **Anträge an die Mitgliederversammlung**
    Es lagen keine Anträge vor.

11. **Verschiedenes**
    Nachfragen von Mitgliedern wurden beantwortet.

12. **Genehmigung des Protokolls**
    Das Protokoll wird von der MV einstimmig genehmigt.

Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung

Im FIfF haben sich rund 700 engagierte Frauen und Männer aus Lehre, Forschung, Entwicklung und Anwendung der Informatik und Informationstechnik zusammengeschlossen, die sich nicht nur für die technischen Aspekte, sondern auch für die gesellschaftlichen Auswirkungen und Bezüge des Fachgebietes verantwortlich fühlen. Wir wollen, dass Informationstechnik im Dienst einer lebenswerten Welt steht. Das FIfF bietet ein Forum für eine kritische und lebendige Auseinandersetzung – offen für alle, die daran mitarbeiten wollen oder auch einfach nur informiert bleiben wollen.

Vierteljährlich erhalten Mitglieder die Fachzeitschrift FIfF-Kommunikation mit Artikeln zu aktuellen Themen, problematischen Entwicklungen und innovativen Konzepten für eine verträgliche Informationstechnik. In vielen Städten gibt es regionale AnsprechpartnerInnen oder Regionalgruppen, die dezentral Themen bearbeiten und Veranstaltungen durchführen. Jährlich findet an wechselndem Ort eine Fachtagung statt, zu der TeilnehmerInnen und ReferentInnen aus dem ganzen Bundesgebiet und darüber hinaus anreisen. Darüber hinaus beteiligt sich das FIfF regelmäßig an weiteren Veranstaltungen, Publikationen, vermittelt bei Presse- oder Vortragsanfragen ExpertInnen, führt Studien durch und gibt Stellungnahmen ab etc. Das FIfF kooperiert mit zahlreichen Initiativen und Organisationen im In- und Ausland.

FIfF online

Das ganze FIfF
www.fiff.de
Twitter FIfF e.V. – *@FIfF_de*

Cyberpeace
cyberpeace.fiff.de
Twitter Cyberpeace – *@FIfF_AK_RUIN*

Faire Computer
blog.faire-computer.de
Twitter Faire Computer – *@FaireComputer*

Mitglieder-Wiki
https://wiki.fiff.de

FIfF-Mailinglisten

FIfF-Mailingliste
An- und Abmeldungen an:
http://lists.fiff.de/mailman/listinfo/fiff-L
Beiträge an: *fiff-L@lists.fiff.de*

FIfF-Mitgliederliste
An- und Abmeldungen an:
http://lists.fiff.de/mailman/listinfo/mitglieder

FIfF-Beirat

Ute Bernhardt (Berlin); **Peter Bittner** (Bad Homburg); **Dagmar Boedicker** (München); Dr. **Phillip W. Brunst** (Köln); Prof. Dr. **Christina Claß** (Jena); Prof. Dr. **Wolfgang Coy** (Berlin); Prof. Dr. **Wolfgang Däubler** (Bremen); Prof. Dr. **Christiane Floyd** (Berlin); Prof. Dr. **Klaus Fuchs-Kittowski** (Berlin); Prof. Dr. **Michael Grütz** (München); Prof. Dr. **Thomas Herrmann** (Bochum); Prof. Dr. **Wolfgang Hesse** (München); Prof. Dr. **Wolfgang Hofkirchner** (Wien); Prof. Dr. **Eva Hornecker** (Weimar); **Werner Hülsmann** (München); **Ulrich Klotz** (Frankfurt am Main); Prof. Dr. **Klaus Köhler** (Mannheim); Prof. Dr. **Jochen Koubek** (Bayreuth); Dr. **Constanze Kurz** (Berlin); Prof. Dr**. Klaus-Peter Löhr** (Berlin); Prof. Dr. **Dietrich Meyer-Ebrecht** (Aachen); **Werner Mühlmann** (Calau); Prof. Dr. **Frieder Nake** (Bremen); Prof. Dr. **Rolf Oberliesen** (Paderborn); Prof. Dr. **Arno Rolf** (Hamburg); Prof. Dr. **Alexander Rossnagel** (Kassel); **Ingo Ruhmann** (Berlin); Prof. Dr. **Gerhard Sagerer** (Bielefeld); Prof. Dr. **Gabriele Schade** (Erfurt); **Ralf E. Streibl** (Bremen); Prof. Dr. **Marie-Theres Tinnefeld** (München); Dr. **Gerhard Wohland** (Mainz); Prof. Dr. **Eberhard Zehendner** (Jena)

FIfF-Vorstand

Stefan Hügel (Vorsitzender) – Frankfurt am Main
Rainer Rehak (stellv. Vorsitzender) – Berlin
Michael Ahlmann – Kiel / Blumenthal
Gilbert Assaf – Berlin
Alexander Heim – Berlin
Sylvia Johnigk – München
Prof. Dr. **Hans-Jörg Kreowski** – Bremen
Kai Nothdurft – München
Prof. Dr. **Britta Schinzel** – Freiburg im Breisgau
Dr. **Friedrich Strauß** – München
Prof. Dr. **Werner Winzerling** – Fulda
Margita Zallmann – Bremen

FIfF-Geschäftsstelle

Ingrid Schlagheck (Geschäftsführung) – Bremen
Anne Schnerrer – Berlin
Benjamin Kees – Berlin

Impressum

Herausgeber	Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung e. V. (FIfF)
Verlagsadresse	FIfF-Geschäftsstelle Goetheplatz 4 D-28203 Bremen Tel. (0421) 33 65 92 55 *fiff@fiff.de*
Erscheinungsweise	vierteljährlich
Erscheinungsort	Bremen
ISSN	0938-3476
Auflage	1 400 Stück
Heftpreis	7 Euro. Der Bezugspreis für die FIfF-Kommunikation ist für FIfF-Mitglieder im Mitgliedsbeitrag enthalten. Nichtmitglieder können die FIfF-Kommunikation für 28 Euro pro Jahr (inkl. Versand) abonnieren.
Hauptredaktion	Dagmar Boedicker, Stefan Hügel (Koordination), Sylvia Johnigk, Hans-Jörg Kreowski, Dietrich Meyer-Ebrecht, Ingrid Schlagheck
Schwerpunktredaktion	Hans-Jörg Kreowski
V.i.S.d.P.	Stefan Hügel
Retrospektive	Beiträge für diese Rubrik bitte per E-Mail an *redaktion@fiff.de*
Lesen, SchlussFIfF	Beiträge für diese Rubriken bitte per E-Mail an *redaktion@fiff.de*
Layout	Berthold Schroeder, München
Cover	Kölner Dom, Deutschland, 24. April 1945 Quelle: https://catalog.archives.gov/id/531287
Druck	Meiners Druck, Bremen Heftinhalt auf 100 % Altpapier gedruckt.

Die FIfF-Kommunikation ist die Zeitschrift des „Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung e. V." (FIfF). Die Beiträge sollen die Diskussionen unter Fachleuten anregen und die interessierte Öffentlichkeit informieren. Namentlich gekennzeichnete Artikel geben die jeweilige Autor:innen-Meinung wieder.

Die FIfF-Kommunikation ist das Organ des FIfF und den politischen Zielen und Werten des FIfF verpflichtet. Die Redaktion behält sich vor, in Ausnahmefällen Beiträge abzulehnen.

Nachdruckgenehmigung wird nach Rücksprache mit der Redaktion in der Regel gern erteilt. Voraussetzung hierfür sind die Quellenangabe und die Zusendung von zwei Belegexemplaren. Für unverlangt eingesandte Artikel übernimmt die Redaktion keine Haftung.

Wichtiger Hinweis: Wir bitten alle Mitglieder und Abonnent:innen, Adressänderungen dem FIfF-Büro möglichst umgehend mitzuteilen.

Aktuelle Ankündigungen

(mehr Termine unter *www.fiff.de*)

FIfF-Kommunikation

1/2024 „FIfF-Konferenz 2023"
Hans-Jörg Kreowski u. a.
Redaktionsschluss: 2. Februar 2024

2/2024 „40 Jahre FIfF"
Hans-Jörg Kreowski, Stefan Hügel u. a.
Redaktionsschluss: 3. Mai 2024

3/2024 „Datenschutz"
Jörg Pohle, Stefan Hügel
Redaktionsschluss: 2. August 2024

Zuletzt erschienen:

4/2022 100 Jahre Joseph Weizenbaum
1/2023 #FIfFKon22
2/2023 Mensch – Gesellschaft – Umwelt ... und Informatik
3/2023 IT-Gestaltung für Gute Arbeit

W&F – Wissenschaft & Frieden

1/23 Jenseits der Eskalation
2/23 Klimakrise
3/23 Gesellschaft in Konflikt
4/23 40 Jahre Wissenschaft & Frieden

vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik

#241 Demokratie und Rechtsstaat verteidigen
#242 Künstliche Intelligenz
#243 Kritische Kriminalpolitik
#244 Identitätspolitik

DANA – Datenschutz-Nachrichten

1/23 Europäische Entwicklungen
2/23 Europäische Entwicklungen, Teil 2
3/23 Whistleblowing
4/23 Internet der Dinge

Das FIfF-Büro

Geschäftsstelle FIfF e. V.
Ingrid Schlagheck (Geschäftsführung)
Goetheplatz 4, D-28203 Bremen
Tel.: (0421) 33 65 92 55, Fax: (0421) 33 65 92 56
E-Mail: fiff@fiff.de
Die Bürozeiten finden Sie unter *www.fiff.de*

Bankverbindung
Bank für Sozialwirtschaft (BFS) Köln
Spendenkonto:
IBAN: DE79 3702 0500 0001 3828 03
BIC: BFSWDE33XXX

Kontakt zur Redaktion der FIfF-Kommunikation:
redaktion@fiff.de

„A shy girl“ (2022) von Yú Xinhan

XinHan Yú schreibt mit seiner aufblasbaren Skulptur die Spielregeln der chinesischen Politik um und verursacht einen sichtbaren Kurzschluss in der realen Welt. Panzer sind ein spezielles Symbol in China, einerseits für Macht und Gewalt, andererseits für verbotene Wörter und unaussprechliche Ereignisse. Über das Massaker auf dem Platz des Himmlischen Friedens ist im Netz keine Information zu finden, es wird umschrieben als die Ereignisse vom 4. Juni. Am 4. Juni dieses Jahres verkaufte eine prominente Internet-Persönlichkeit mit 61,8 Millionen Followern Eiscreme in Form eines Panzers vom Typ 59, die chinesische Regierung reagierte darauf mit einer dauerhaften Sperrung.

Deshalb hat Xinhan Yú für sein Werk den Panzer gewählt. Die plüschige Skulptur sieht groß und hart aus, ist aber auch leicht zu zerstören. Die Farbe Rosa, auch in China weiblich konnotiert, lässt sie weich und zerbrechlich wirken.

Foto von der Ausstellung Arte Laguna Prize 2023, Venedig

Geeignete Texte für den SchlussFIfF bitte mit Quellenangabe an *redaktion@fiff.de* senden.