

E..I..f..F.. Kommunikation

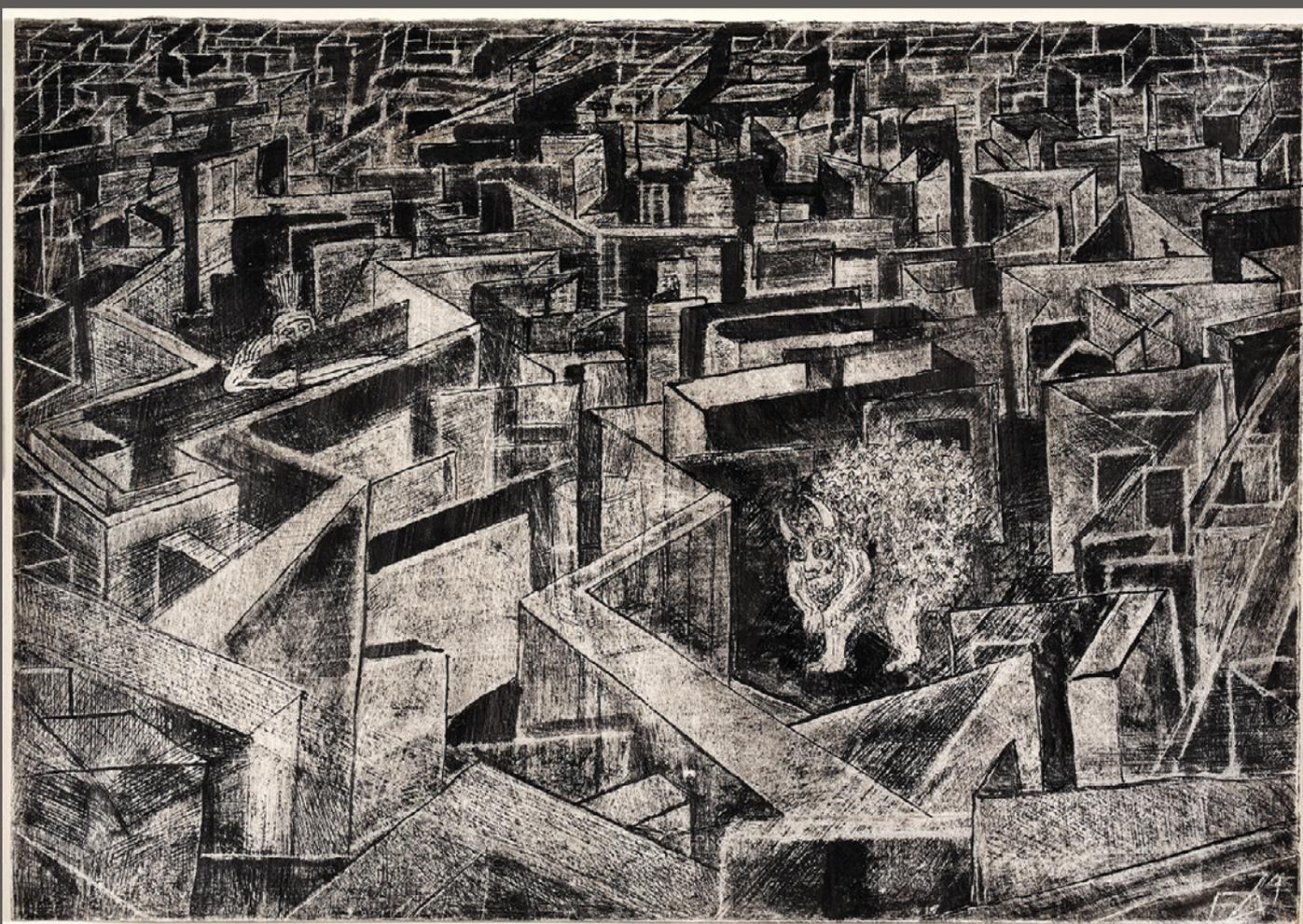
Zeitschrift für Informatik und Gesellschaft

39. Jahrgang 2022

Einzelpreis: 7 EUR

2/2022 – Juni 2022

Künstliche Intelligenz



ISSN 0938-3476

Inhalt

Ausgabe 2/2022

inhalt

- 03 Editorial
- *Stefan Hügel*

Forum

- 04 Der Brief: Pazifismus oder Militärlogik?
- *Stefan Hügel*
- 06 Betrifft: Faire Computer
- *Sebastian Jekutsch*
- 07 Schwerpunkt *Digitalisierung in Staat, Politik & Verwaltung* der FIF-Kommunikation 3/2022
- Call for Contributions
- 08 Cyberpeace – für Frieden, Freiheit und eine lebenswerte Welt
- *Hans-Jörg Kreowski und Aaron Lye*
- 09 Krieg mit Künstlicher Intelligenz
- *Christian Heck und Hans-Jörg Kreowski*
- 10 Quantenschlüsselverteilung: Von Glasfaser zu Satelliten
- *Aaron Lye*
- 13 Digitalisierung und Nachhaltigkeit zusammendenken – Zweite bundesweite *Bits & Bäume* Konferenz
- *Gemeinsame Pressemitteilung*

Weizenbaum-Medaille 2021

- 15 Verleihung der Weizenbaum-Medaille 2021
- *FIF e. V.*
- 15 Laudatio für Prof.in Dr.in Dr.in h.c. Christiane Floyd
- *Britta Schinzel und Hans-Jörg Kreowski*
- 18 Ansprache zur Verleihung der Weizenbaum-Medaille 2021
- *Christiane Floyd*

FIF e. V.

- 22 Ankündigung FIFKon 2022
make install peace – Impulse für den Frieden
- 23 Einladung zur Mitgliederversammlung 2022

Rubriken

- 75 Impressum/Aktuelle Ankündigungen
- 76 SchlussFIF

Künstliche Intelligenz

- 24 Künstliche Intelligenz: Zwischen hochgeschraubten Erwartungen und unkalkulierbaren Risiken
- *Hans-Jörg Kreowski und Stefan Hügel*
- 25 KI aus interdisziplinärer Perspektive
- *Andrea Reichenberger und Martin Ziegler*
- 29 Anna Strasser, Wolfgang Soest, Ralf Stapelfeld, Katja Stepec (Hg.): *Künstliche Intelligenz: Die große Verheißung*
- *Britta Schinzel*
- 32 Die Illusionsfabrik der »KI«-Narrative
- *Peter Brödner*
- 36 The Language Labyrinth: Constructive Critique on the Terminology Used in the AI Discourse
- *Rainer Rehak*
- 43 Human-Centered Data Science
- *Claudia Müller-Birn*
- 46 Je künstlicher die Intelligenz, desto weniger intelligent
- *Klaus Heß*
- 51 Stuart Russell: *Human Compatible* – Künstliche Intelligenz und wie der Mensch die Kontrolle über superintelligente Maschinen behält
- *Dorothea Kolossa*
- 53 Eine lebenswerte Welt – Kann Künstliche Intelligenz helfen, das Ziel zu erreichen?
- *Hans-Jörg Kreowski*

Joseph Weizenbaum

- 57 Joseph Weizenbaum: *On the Impact of the Computer on Society*
- *Andrea Knaut und Stefan Ullrich*
- 57 *On the Impact of the Computer on Society*
- *Joseph Weizenbaum*

BigBrotherAwards 2022

- 64 BigBrotherAwards 2022 – Einleitung
- *Stefan Hügel*
- 67 Kategorie *Behörden und Verwaltung* – Laudatio
- *Thilo Weichert*
- 70 Kategorie *Lebenswerk* – Laudatio
- *Rena Tangens*

Editorial

Künstliche Intelligenz bildete in der jüngeren Vergangenheit schon häufiger den Schwerpunkt der *Fiff-Kommunikation*. Nicht verwunderlich – gilt „KI“ doch manchen als „Künftige Informatik“, so dem KI-Experten Professor Dr. Wolfgang Wahlster.

So geht es also auch in dieser Ausgabe um *Künstliche Intelligenz – Zwischen hochgeschraubten Erwartungen und unkalulierbaren Risiken*. Eindrucksvollen Erfolgen, beispielsweise in der Bild- und Sprachverarbeitung, stehen teilweise gravierende Fehlschläge gegenüber, wie bei massiven Falschbewertungen bei der Erkennung von Kindergeldbetrug in den Niederlanden, die sogar politische Konsequenzen auf höchster Ebene hatten. Fehler bei der Auswahl von Trainingsdaten führen auch sonst zu falschen Klassifikationen bis hin zum *programmierten Rassismus*. Und nicht zuletzt übersteigerte Erwartungen an eine *starke KI* führen zu einer falschen Einschätzung der tatsächlichen Leistungsfähigkeit Künstlicher Intelligenz. Werden wir eines Tages einer neuen, intelligenten Lebensform gegenüber stehen, oder haben wir es letztlich doch „nur“ mit einer Methode der Informatik zur Herstellung leistungsfähiger Anwendungen zu tun? Welche Konsequenzen hat das Paradigma des *Maschinellen Lernens* für die Zuverlässigkeit und Nachvollziehbarkeit von Systemen, die ohnehin schon sehr komplex sind?

Anlässlich unserer Fiff-Konferenz 2021 verliehen wir die Weizenbaum-Medaille an Professorin Christiane Floyd. Sie hat bei der Gründung des Fiff als erste Vorsitzende eine wesentliche Rolle gespielt, in ihrer fachlichen Arbeit insbesondere die Softwaretechnik wesentlich beeinflusst und sie unter Einbeziehung sozialwissenschaftlicher, ethischer und philosophischer Gesichtspunkte beispielhaft ausgestaltet – so Britta Schinzel in ihrer Laudatio, in der sie ihre wissenschaftlichen Leistungen und ihr großes Engagement für eine gesellschaftlich und ethisch fundierte Informatik nachzeichnet. In ihrer Dankesrede erinnerte Christiane Floyd an die Anfangstage des Fiff und die friedenspolitische Situation im kalten Krieg:

In den 80er-Jahren war die Arbeit des Fiff vor allem durch den Kalten Krieg bestimmt, durch die beidseitige Aufrüstung und die berechtigte Furcht vor einer nuklearen Auseinandersetzung auf deutsch-deutschem Boden, die beide Teile des heute wiedervereinigten Landes ernsthaft gefährdet hätte. Dieses Bedrohungsszenario war sehr real. Es galt also zu warnen, sich eine Stimme zu verschaffen, sich mit anderen Engagierten zu vernetzen und in der Öffentlichkeit Präsenz zu zeigen.

In einem Nachsatz aus aktuellem Anlass betonte sie die Bedeutung der Debatte im Fiff angesichts der neuen Herausforderungen, die sich durch den Krieg in der Ukraine ergeben. Die Verhältnisse haben sich seit 1989 geändert, und so müsse eine neue Generation neue Antworten auf die militärische Bedrohung finden.

Im Januar 2023 jährt sich der Geburtstag von Joseph Weizenbaum zum hundertsten Mal. Neben dem geplanten Schwerpunkt in der Ausgabe 4/2022 enthält auch diese Ausgabe einen Text, den er 1972 veröffentlicht hat: *On the Impact of the Computer on Society – How does one insult a machine?*

Am 29. April 2022 wurden in Bielefeld die diesjährigen Big-BrotherAwards vergeben. Das ist selbstverständlich auch für uns ein Thema – diese Ausgabe enthält die Laudationes auf die deutsche Polizei in der Kategorie *Behörden und Verwaltung* und auf die Datenschutzbehörde der Republik Irland für ihr *Lebenswerk*.

Dazu kommen unsere Rubriken *Betrifft Faire Computer* und *Cyberpeace*. In ersterer stellt Sebastian Jekutsch den aktuellen Stand beim Lieferkettengesetz dar. Das Bewusstsein für dessen Bedeutung scheint nach dem Wechsel an der Spitze des Bundeswirtschaftsministeriums geschärft – inwieweit die notwendigen strukturellen Veränderungen tatsächlich in Angriff genommen werden, bleibt abzuwarten. Für uns gibt es ein Jubiläum zu feiern: Den Twitter-Account *@FaireComputer* gibt es nun seit zehn Jahren.

Die Rubrik *Cyberpeace* hat drei Abschnitte: Der erste stellt aktuelle Entwicklungen und Ereignisse im Umfeld des Cyberkriegs zusammen. Im zweiten Abschnitt berichten wir von einer Online-Anhörung zum *Krieg mit Künstlicher Intelligenz*, die wir gemeinsam mit dem Arbeitskreis *Gegen bewaffnete Drohnen* am 10. März 2022 durchgeführt haben und dessen Aufzeichnung auch auf unserer Homepage *fiff.de* zu finden ist. Der dritte Abschnitt befasst sich mit der Verschlüsselung auf Quantencomputern und in diesem Zusammenhang speziell der Schlüsselverteilung.

Auch uns beschäftigt der Krieg in der Ukraine. Neben einem Kommentar dazu im Brief dieser Ausgabe verweisen wir auf unsere Schwesterzeitschrift *Wissenschaft und Frieden*, die sich kurzfristig in ihrer aktuellen Ausgabe 2/2022 und ausführlicher in der nächsten Ausgabe 3/2022 mit dieser Thematik auseinandersetzt.

Zu zwei großen Konferenzen laden wir in diesem Jahr ein: Gemeinsam mit zwölf weiteren zivilgesellschaftlichen Organisationen richten wir vom 30. September bis 2. Oktober 2022 die Konferenz *Bits & Bäume* aus. Vom 21.-23. Oktober 2022 findet die diesjährige Fiff-Konferenz statt, bei der wir Impulse für den Frieden erzeugen wollen.

Zuletzt noch ein Hinweis in eigener Sache: Bekanntlich sind wir Mitherausgeber des jährlich erscheinenden *Grundrechte-Reports*, der in diesen Tagen veröffentlicht wurde. In 39 Beiträgen setzt er sich auch in diesem Jahr mit einer Reihe von Bürgerrechtsthemen auseinander – Schwerpunkte setzt er bei den Grundrechtsfragen, die durch den chaotischen Abzug der Bundeswehr aus Afghanistan entstanden sind, und bei den bürgerrechtlichen Aspekten des Klimaschutzes und der Proteste dagegen. Der Grundrechte-Report kann über unsere Geschäftsstelle bezogen werden; eine Aufzeichnung der Pressekonferenz zur Vorstellung des Grundrechte-Reports gibt es ebenfalls auf *fiff.de*.

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

Stefan Hügel
für die Redaktion



*Soldat, gingst Du gläubig und gern in den Tod?
Oder hast Du verzweifelt, verbittert, verroht,
Deinen wirklichen Feind nicht erkannt bis zum Schluss ...?*

Hannes Wader, Es ist an der Zeit



Pazifismus oder Militärlogik?

Liebe Freundinnen und Freunde, liebe Mitglieder des FIFF,

„Der Pazifismus der 30er-Jahre – der sich in seiner gesinnungsethischen Begründung nur wenig von dem unterscheidet, was wir in der Begründung des heutigen Pazifismus zur Kenntnis zu nehmen haben – dieser Pazifismus der 30er-Jahre hat Auschwitz erst möglich gemacht.“¹

Dieser Ausspruch von Heiner Geißler hat 1983 im Deutschen Bundestag und weit darüber hinaus große Empörung ausgelöst – nicht nur in der Friedensbewegung, sondern sogar in Teilen der eigenen Regierungskoalition. Vielen Menschen war damals noch in frischer Erinnerung, dass Pazifisten wie Carl von Ossietzky in den Konzentrationslagern des Nationalsozialismus festgehalten wurden, misshandelt und gefoltert wurden, starben.² Otto Schily – damals noch bei den Grünen – wies erbost auf die politische Vorgeschichte und den Weg zur Herrschaft des Nationalsozialismus hin, insbesondere das Ermächtigungsgesetz, das bei den damaligen konservativen und liberalen Parteien eine breite Mehrheit fand.

Doch heute scheint der Pazifismus erneut diskreditiert. Aber was kann Pazifismus angesichts militärischer Aggression bedeuten? Teile der Friedensbewegung der 1980er-Jahre beriefen sich auf das Friedensgebot der Bergpredigt³:

„Selig sind, die Frieden stiften; denn sie werden Gottes Kinder heißen.“ (Mt 5, 9)

Und:

„Ich aber sage euch, dass ihr nicht widerstreben sollt dem Bösen, sondern: Wenn dich jemand auf deine rechte Backe schlägt, dem biete die andere auch dar.“ (Mt 5, 39)

Das wirkt angesichts eines realen Krieges naiv. Pazifismus bedeutet aber nicht, sich angesichts heranrückender Panzer auf eine Wiese zu setzen und Gänseblümchen zu pflücken. Pazifismus bedeutet, durch vorausschauende Politik Kriege im Vorfeld zu vermeiden oder unmöglich zu machen – so, dass ein militärisches Eingreifen nicht mehr notwendig ist. Oder um ein anderes Bild zu verwenden: Pazifismus bedeutet, angesichts einer Geisterfahrer:in auf der Autobahn nicht noch zusätzlich Gas zu geben. Wenn es notwendig ist, alles zu tun, um weitere Eskalation zu verhindern und den Konflikt so schnell wie möglich zu beenden. Mit Parolen wie „Frieden ist erst möglich, wenn der Krieg gegen Russland gewonnen ist“, werden wir das aber nicht erreichen.

Bisher gab es einen klaren Grundsatz, der sich noch vor wenigen Monaten als Kernaussage im Wahlkampf der Grünen wiederfindet: „Keine deutschen Waffen in Kriegsgebiete und Diktaturen.“⁴ Dieser Grundsatz gilt nun angesichts des realen Krieges

offensichtlich nicht mehr – also nur ein naiver Irrtum, oder sind wir allzusehr bereit, unsere Grundsätze aufzugeben, wenn wir glauben, dass es ernst wird? Der Historiker Frank Biess kommentiert in *Spiegel Online*:

„... die heute zu beobachtende Kehrtwende ehemaliger friedensbewegter Aktivisten hin zu immer weiterreichenden und umfassenderen Forderungen nach Waffenlieferungen an die Ukraine, verbunden mit Denunziationen früherer pazifistischer Neigungen bei sich selbst und anderen, ist ... verwunderlich. ... [E]s ist auch eine Form des Geschichtsrevisionismus, wenn ehemalige friedensbewegte Aktivisten jetzt erklären, Helmut Schmidt habe 1979 mit dem Nato-Doppelbeschluss doch recht gehabt, und nur die Logik der Abschreckung habe Jahrzehnte des Friedens beschert.“⁵

Eine Verhandlung mit Präsident Putin sei nicht möglich, ein russischer Diktatfrieden niemals akzeptabel. Ja! Aber wie soll es dann weitergehen? Immer mehr „schwere“ Waffen an die Ukraine liefern? Bis zur nuklearen Eskalation? Welche realistischen Szenarien ergeben sich, wenn dieser Krieg noch Monate oder Jahre andauert? Auch die einflussreiche *New York Times* rät inzwischen zu realistischen Zielen.⁶

Keine Frage: Es steht uns nicht zu, den Menschen in der Ukraine Ratschläge zu erteilen, ob und wie sie sich gegen die Aggression verteidigen sollen. Doch der Blick lohnt sich, wer denn „die Menschen in der Ukraine“ sind. Kriege werden immer von Eliten geführt – und von der Bevölkerung verloren. Und von den tausenden gefallenen Soldat:innen auf beiden Seiten.

Es ist angesichts der Gräueltaten nicht verwunderlich, dass die derzeitige Debatte auch in Deutschland sehr emotional geführt wird. Emotionalität wird gezielt eingesetzt, um politische Ziele zu erreichen – auch das ist nichts Neues. Den Anfang machte Außenministerin Baerbock, die von ihren Anhänger:innen für ihre Rede von der UN gefeiert wurde, in der sie stark emotionalisierende Bilder nutzte: „Vor ein paar Tagen kam in einer U-Bahn-Station in Kiew ein kleines Mädchen zur Welt. Ich habe gehört, es heißt Mia.“⁷ Natürlich ist es gut und richtig, dass die Resolution⁸ zur Verurteilung des russischen Einmarschs in der Ukraine mit großer Mehrheit verabschiedet wurde.

In einem klugen Leitartikel des *Spiegel* zeigt Dirk Kurbjuweit das Dilemma zweier Ziele in diesem Konflikt auf:

„Nur Putins Niederlage zu wollen führte unweigerlich in einen großen Krieg. Nur die NATO raushalten zu wollen führte zur Auslöschung der Ukraine. Eine gute Balance

zu finden ist die große Aufgabe der Politik. Man wird sie nicht lösen können, ohne sich gegenüber den Menschen in der Ukraine schuldig zu machen, weil aus deren Sicht nur Ziel eins zählt. Mit dieser Schuld müssen wir leben.“⁹

In dieser Situation ist eine Bundesregierung wichtig, die eine besonnene und im Zweifel deeskalierende Politik betreibt und sich nicht von einer emotionalisierenden öffentlichen Debatte leiten lässt. Dem Bundeskanzler wurde deswegen wiederholt mangelnde Führungsstärke vorgeworfen. Doch geht es überhaupt noch um Führungsstärke und nicht eher darum, dass er endlich das tut, was seine Kritiker:innen von ihm erwarten? Wir beobachten eine Rhetorik, die militärisches Vorgehen ins Zentrum rückt – dazu gehört eine eskalierende Berichterstattung in den Leitmedien genauso wie die Entgleisungen¹⁰ des ukrainischen Botschafters Andrij Melnyk.¹¹ Die Wahrheit ist das erste Opfer des Krieges – dieser Satz gilt auch hier.

Als sich am 8. Mai 2022 das Ende des Zweiten Weltkriegs jährte, hieß es im *Spiegel*:

„Nie wieder Krieg‘ ist keine zeitgemäße Haltung mehr für Deutschland.“¹²

Ja, aber was denn sonst? Gerade die verstörenden Bilder, die uns täglich in allen Medien präsentiert werden, sollten deutlich machen, dass wir immer die Deeskalation über die militärische Logik stellen sollten. Gerade im 21. Jahrhundert sollte „nie wieder Krieg“ nicht nur zeitgemäß, sondern zwingend sein. Militärisches Eingreifen ist das allerletzte Mittel. Kluge Sicherheitspolitik sorgt im Vorfeld dafür, dass es nicht zur Eskalation kommt.

Wir müssen die Ukraine unterstützen und einen weiteren Vormarsch der russischen Armee verhindern. Wir dürfen uns aber nicht in eine Eskalation hineinsteigern, die zu weiterem Tod und zu weiterer Zerstörung führt – bis hin zum Einsatz von Atomwaffen. Ich setze darauf, dass viele Menschen hinter den politischen Kulissen an Lösungen arbeiten, ohne dies sofort in die Öffentlichkeit zu tragen. Das Ergebnis wird sicherlich Diskussionen auslösen und Widerspruch hervorrufen – auf der einen oder anderen Seite. Darüber wird dann zu reden sein, aber bitte ohne eskalierende Rhetorik.

Mit FlFFigen Grüßen
Stefan Hügel

Anmerkungen

- 1 Deutscher Bundestag, Plenarprotokoll 10/13, 15. Juni 1983, <https://dserver.bundestag.de/btp/10/10013.pdf>
- 2 Wikipedia, Stichwort Carl von Ossietzky, https://de.wikipedia.org/wiki/Carl_von_Ossietzky
- 3 Die Bibel, Matthäus-Evangelium, Kapitel 5, 1–7, 29
- 4 Deutschland. Alles ist drin. Bundestagswahlprogramm Bündnis 90/Die Grünen 2021, Seite 250, https://cms.gruene.de/uploads/documents/Wahlprogramm-DIE-GRUENEN-Bundestagswahl-2021_barrierefrei.pdf
- 5 Frank Biess: Die Rückkehr der German Angst. Spiegel Online, <https://www.spiegel.de/geschichte/deutsche-reaktionen-auf-den-ukrainekrieg-german-angst-a-1e4fe090-bfa8-46d1-ac51-f4313a1f3024>
- 6 New York Times, Editorial Board: „... as the war continues, Mr. Biden should also make clear to President Volodymyr Zelensky and his people that there is a limit to how far the United States and NATO will go to confront Russia, and limits to the arms, money and political support they can muster.“, New York Times, 19. Mai 2022, <https://www.nytimes.com/2022/05/19/opinion/america-ukraine-war-support.html>
- 7 Um einem scheinbar naheliegenden Einwand zuvorzukommen: Ich unterstelle Frau Baerbock nicht, emotional zu handeln. Im Gegenteil: Die Rede ist nicht emotional. Sie nutzt gezielt manipulativ emotionalisierende Bilder, um ein politisches Ziel zu erreichen.
- 8 Vereinte Nationen: Aggression gegen die Ukraine. Resolution ES-11/1 der VN-Generalversammlung, verabschiedet am 2. März 2022, <https://www.un.org/depts/german/gv-notsondert/a-es11-1.pdf>
- 9 Dirk Kurbjuweit: Die Abgründe der Politik. Der Spiegel 16/2022, Seite 6
- 10 Ob es sich dabei um echte Entgleisungen handelt, oder um gezielt eingesetzte Rhetorik, sei dahingestellt.
- 11 Dies wird kommentiert durch Otfried Höffe: Ukrainischer Botschafter Melnyk: Verbalattacken eines Hilfesuchenden, Frankfurter Rundschau, 27. April 2022, <https://www.fr.de/kultur/gesellschaft/melnyk-ukraine-botschafter-verbalattacken-eines-hilfesuchenden-91505390.html>
- 12 Susanne Beyer: Mut zur Selbstkritik. Der Spiegel 19/2022, Seite 6. Der Leitartikel wurde mit anderer Überschrift und anderer Einleitung auch auf Spiegel Online veröffentlicht: Nie wieder Krieg ist die falsche Lehre. Dort heißt es nun: „Die Strategie der Bundesregierung, die Ukraine klar zu unterstützen, auch mit schweren Waffen, ist angreifbar. Eine Lehre für den diesjährigen 8. Mai kann sein, dass es in Ordnung ist, sich angreifbar zu machen.“ Spiegel Online, <https://www.spiegel.de/politik/deutschland/gedenktag-8-mai-nie-wieder-krieg-ist-die-falsche-lehre-a-eff469a8-235d-4c8b-93c1-5bf8cb3058e6>



Das FlFF bittet um Eure Unterstützung

Viermal im Jahr geben wir die FlFF-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. Unsere jährlich stattfindende FlFF-Konferenz, der Weizenbaum-Preis, weitere Publikationen, Kommunikation im Web: Neben der tatkräftigen Mitwirkung engagierter Menschen sind wir bei unserer Arbeit auf finanzielle Unterstützung angewiesen.

FlFF

Bitte unterstützt das FlFF 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

Betrifft: Faire Computer



@FaireComputer feiert 10 Jahre Neuigkeiten per Twitter

„Wir ziehen mit unserem täglichen Leben eine Spur der Verwüstung durch die Erde, was immer wir tun hat Konsequenzen“, sagte Wirtschaftsminister Habeck in einer der vielen Talkrunden. Und als Beispiel fiel ihm in dieser spontanen Verteidigung seiner neuen Arbeit als Lieferkettenirgendwienordnungsbringer das Naheliegende ein: „Beim Abbau von Produkten für Handys, Computer usw. sind wir in einer Verantwortung ... und wir kümmern uns da niemals drüber. Aber punktuell entdecken wir dann immer unser moralisches Gewissen“, und weiter: „Was wir brauchen ist eine strukturelle Arbeit an der Veränderung, der Verbesserung.“

Das sind neue Töne aus diesem Ministerium. Sein Vorgänger hatte noch zusammen mit Industrievertretern erfolgreich gegen weitgehende Entwürfe des Arbeitsministeriums für ein deutsches Lieferkettengesetz gearbeitet. Lieferkettengesetze versuchen zu reparieren, was andere (fehlende) Gesetze an globalisierter Verantwortungslosigkeit ermöglichen. Das deutsche *Lieferkettensorgfaltspflichtengesetz* ist durch – wir berichteten in der letzten Ausgabe dieser Kolumne –, mal schauen, wie sich die neue Regierung nun bei der anstehenden EU-weiten Sorgfaltspflichtenregulierung verhält. Die EU-Kommission hat nämlich inzwischen einen Entwurf vorgelegt und zwar mit einer – in Deutschland noch verhinderten – zivilrechtlichen Haftung. Das heißt: Wer sich als Unternehmen nicht bemüht, kann verklagt und verurteilt werden. Allerdings haben Kläger die Beweislast, also die Betroffenen. Kümmerlich ist der Adressatenkreis von nur etwa 1 % aller Unternehmen, zudem ohne besondere Beachtung von Risikobranchen. Es soll auch nur für „etablierte Lieferbeziehungen“ gelten. Was auch immer das genau heißt, in unserer schnelllebigen IT-Branche ist das jedenfalls nicht unbedingt typisch. Damit zementiert der Entwurf die Machtungleichheit zum globalen Süden eher: Die Auswirkungen einer sich kurzfristig ändernden Einkaufspolitik der IT-Hersteller werden nicht angefasst. „Strukturelle Veränderung“ sieht anders aus.

Beim EU-Lieferkettengesetz ist nun wieder das EU-Parlament dran und danach die Mitgliedstaaten. Schon im danach folgenden Trilog befindet sich die EU-Batterieverordnung. Es geht um besseren Umweltschutz und die Einhaltung von Menschenrechten beim Bezug der Rohstoffe für Batterien, die in der EU hergestellt werden. Idealerweise sollen sie aus dem Recycling kommen. Die enormen Steigerungen erlauben das aber gar nicht, so dass es frisches Kobalt aus dem Kongo oder Lithium aus Südamerika (wir berichteten) betrifft, aber auch Nickel, von dem

übrigens noch viel aus Russland kommt. *Good Electronics* – das FIFF ist dort Mitglied – verlangt zusammen mit anderen NGOs, dass auch die Batterierohstoffe Bauxit (Aluminium), Eisen und Kupfer darin reguliert werden.

Für letzteres gibt es sogar in Deutschland eine Hütte mit Weltbedeutung, nämlich bei Aurubis, u. a. in Hamburg. Sie öffneten sich in der jüngsten Zeit ihrer Lieferkettenverantwortung ein wenig, wie mit der Entscheidung im vergangenen Jahr, kein Kupfererz aus indigenen Gebieten Norwegens zu beziehen. Das meiste Erz beziehen sie aktuell aus Peru. Von dort erreichen uns Nachrichten über gewaltsame Proteste der Anwohner, die sich auf ihr Recht zur Beteiligung an den Rohstoffvorkommen in ihren Regionen berufen und eine Ausweitung der Förderung ablehnen. Ob dies eine Quelle für Aurubis ist, ist allerdings nicht bekannt, so weit geht die Transparenz nicht. Von Protesten dieser Art hört man hingegen häufiger, vor allem aus der Andenregion.

Was gibt's Neues bei den Pionierprojekten?

- Fairphone hat ein neues Produkt, die Earbuds, also kabellose Kopfhörer. Das Faire an ihnen: Fair-Trade-Gold per Massenausgleich, wie auch schon in ihren Smartphones. Allerdings haben die Geräte keine austauschbaren Batterien und sind nicht reparierbar, das soll noch besser werden.
- Neu ist Syllucid mit einem USB-Kabel inklusive Steckersystem mit ebenfalls Fair-Trade-Gold (die Gründer waren früher bei Fairphone) und immerhin Recycling-Kupfer. Das Crowdfunding war im vergangenen Jahr erfolgreich, in Kürze wird geliefert. USB-Kabel sind beliebt: Auch Recable und GSN sind Kabelanbieter mit einem Fokus auf Fairness und Nachhaltigkeit.
- Shiftphone („die deutsche Alternative“) macht eher in Kreislaufwirtschaft denn in Fairness. Zusammen mit einem Institut der Uni Kassel soll „ein umfassendes Entwicklungskonzept für ein vollständig kreislaufwirtschaftsfähiges Smartphone erarbeitet und prototypisch umgesetzt“ werden. Zudem stammt nun das Stromkabel eines ihrer Ladegeräte aus dem Elektroschrott und bekommt auf diese Weise ein zweites Leben. Ähnliches gab es bislang noch nicht.
- Die Nager-IT-Maus wird nun optional mit kürzerem Kabel geliefert, das spart Rohstoffe. Wegen Lieferproblemen müs-

Sebastian Jekutsch

Sebastian Jekutsch ist Sprecher der AG Faire Computer des FIFF. Wer sich für die Quellen oder das Thema überhaupt interessiert, kann gerne Kontakt aufnehmen per sebastian.jekutsch@fiff.de.

sen fairere SMD-Widerstände inzwischen leider durch konventionelle aus Fernost ersetzt werden, es geht aber nur um wenige Milligramm. Und Nager-IT stellt ein und sucht Mitarbeitende für die öffentliche IT-Beschaffung. In Niedersachsen und Baden-Württemberg waren sie ja schon erfolgreich

und haben ihre bislang mit Abstand größten Aufträge dort aus öffentlicher Hand bekommen.

So wirken nicht nur die Mitglieder der EU, sondern auch die Kommunen in unserer Sache.



Call for Contributions

Schwerpunkt *Digitalisierung in Staat, Politik & Verwaltung* der FfF-Kommunikation 3/2022

Redaktion: Jörg Pohle & Stefan Hügel

Seit mehr als sechzig Jahren folgt in der (bundes-)deutschen Politik und Verwaltung eine Welle der Computerisierung, Informatisierung und Digitalisierung auf die andere. Die Corona-Pandemie und der Umgang damit haben gezeigt, wie vieles dabei noch im Argen liegt – trotz aller Erfolgsmeldungen aus der Vergangenheit. Zugleich waren sie aber auch Auslöser einer neuen Digitalisierungswelle. Wird jetzt endlich alles gut? Oder ist auch „4.0“ nur einfach „mehr vom Alten“?

Im Heftschwerpunkt *Digitalisierung in Staat, Politik & Verwaltung* der FfF-Kommunikation 3/2022 wollen wir einen Blick auf das breite Feld staatlicher Informationssysteme werfen, vor allem, aber durchaus nicht nur, in Deutschland. Es geht um eingesetzte und geplante, aber auch um längst überholte und bereits wieder abgeschaffte Systeme in Kommunen, Ländern und im Bund. Es geht um mit diesen Systemen erbrachte staatliche Leistungen, etwa der Verwaltung, um Meinungsbildung, Demokratie und politische Entscheidungsfindung, um die Organisation solcher Systeme und ihren praktischen Einsatz, um Transparenz und Rechtsstaatlichkeit staatlichen Handelns, sowie um die Auswirkungen dieser Systeme auf Individuen, Gruppen und die Gesellschaft insgesamt. Es geht aber auch um die Prozesse der Digitalisierung, wie sie ablaufen und ablaufen sollten, um Fragen der Mitbestimmung und Intervention, um Erfolge und um Scheitern.

Wir freuen uns über Arbeiten, die in alle Richtungen schauen: in die Vergangenheit, die Gegenwart und die Zukunft, auf die Sonnen- und die Schattenseiten, auf die Theorie und die Praxis, auf die Versprechungen und ihre Umsetzungen in Parlamenten und Amtsstuben, auf Gestaltungs-, Verwendungs- und Wirkungszusammenhänge. Wir freuen uns über Praxisberichte ebenso wie über politische Forderungen an die Digitalisierung und deren Begründungen, über lokale Erfahrungen ebenso wie über nationale oder supranationale Entwicklungen, über informatische Perspektiven ebenso wie über die anderer Disziplinen, über Analysen ebenso wie über Gestaltungsoptionen, über Fragen ebenso wie über mögliche Antworten.

Die möglichen Themen sind so breit gefächert wie der Einsatz staatlicher Informationssysteme und die Möglichkeiten von Bürger:innen, mit solchen Systemen in Kontakt zu kommen

oder konfrontiert zu werden. Aus Sicht von Bürger:innen stellt sich etwa die Frage, ob und wie der Einsatz technischer Systeme als Schnittstelle der Verwaltung nach außen das Verhältnis zwischen Bürger:innen und Staat, deren Interaktion und Kommunikation, aber auch deren jeweiliges Selbstverständnis ändert. Aus der Sicht der Informatik würde sich daran anschließen, welche Gestaltungsforderungen an diese Schnittstellen sich daraus ergeben. Der Blick ins Innere von Politik und Verwaltung wirft hingegen unter anderem die Frage auf, wie mit der zunehmenden „Verdatung“ (Herbert Fiedler 1975) aller Lebensbereiche Prozesse und Ergebnisse von Planungen und Entscheidungen in diesen Bereichen geändert werden oder werden sollten. Auch hier gibt es informatische Anschlussfragen, etwa zur Gestaltbarkeit und Gestaltung von Systemen, die Transparenz und Intervenierbarkeit ermöglichen oder erzwingen. Und auf der Ebene von Staat und Gesellschaft kann etwa gefragt werden, wie sich Gewaltenteilung und Dezentralität unter der Bedingung durchgreifender Digitalisierung aufrecht erhalten oder gar stärken lassen.

Wir freuen uns über Einreichungen von Beiträgen mit ca. 20.000 Zeichen (inklusive Leerzeichen) bis zum 8. Juli 2022 per E-Mail an Jörg Pohle (joerg.pohle@fiff.de) und Stefan Hügel (sh@fiff.de). Alle Beiträge zum Schwerpunkt werden peer-reviewed, und die Autor:innen erhalten bis zum 22. Juli Rückmeldungen zu ihren Beiträgen. Die finalen Fassungen der Beiträge sind bis zum 5. August einzureichen.

Wir freuen uns über die Nutzung der Open-Access-Lizenz *Creative Commons – Namensnennung / CC BY* für Ihren Text und die verwendeten Bilder.

Termine

Einreichungsfrist:	8. Juli 2022
Rückmeldung vom Review:	22. Juli 2022
Redaktionsschluss/ Einreichung der finalen Fassung:	5. August 2022

Hinweise für Autor:innen

vorläufig: <https://pads.fiff.de/p/klausurOkt2021-G3Leitfaden>

Cyberpeace – für Frieden, Freiheit und eine lebenswerte Welt

Seit der Ausgabe 2+3/2021 ist die Cyberpeace-Rubrik Teil der FIFF-Kommunikation. Die Rubrik ist gedacht für Ankündigungen, Berichte, kurze Texte und Stellungnahmen rund um das Thema Cyberpeace – für Frieden, Freiheit und eine lebenswerte Welt. Aber auch längere Beiträge sind willkommen. Alle Leser:innen sind aufgerufen, die Rubrik für eigene Beiträge zu nutzen. Sie können jederzeit an uns geschickt werden: kreo@fiff.de und lye@fiff.de.

In der Ausgabe 4/2021 füllte das Thema Künstliche Intelligenz (KI) und Kriegsführung einen ganzen Schwerpunkt. Auf dieser Grundlage fand am 10. März 2022 ein Online-Hearing des Arbeitskreises *Gegen bewaffnete Drohnen* zum Thema *Krieg mit Künstlicher Intelligenz* statt. Siehe dazu folgenden Bericht. Der Schwerpunkt, das Hearing und auch die kontinuierlichen Berichte zu KI zeigen, dass es wichtig bleibt, die militärischen und repressiven Anwendungen der KI aufzuzeigen und dem entgegenzuwirken.



In Ausgabe 1/2022 ist ein Artikel von Aaron Lye zu dem Thema abgedruckt, wie die NATO den Cyberkrieg probt. Der Artikel ist ein Nachdruck, der zuerst erschienen ist bei der *Informationsstelle Militarisierung (IMI)* Ausgabe 1/2022. Er basiert auf einem Vortrag beim IMI-Kongress im November 2021 zum Thema Manöver als Brandbeschleuniger: Kriegsspiele, Manöver und Konfrontation.

Ebenfalls im November 2021 fand zum sechsten Mal die jährliche Sicherheitskonferenz des *Stockholm International Peace Research Institute (SIPRI)* statt. Das Thema der Konferenz war *Schlachtfelder der Zukunft: Trends der Konflikt- und Kriegsführung im 21. Jahrhundert*. Als FIFF-Themen sind insbesondere die Sessions zu Kriegsführung im Cyber- und Weltraum, Kriegsführung mittels KI, Informationskrieg als auch Quantentechnologie zu nennen.

Der Weltraum und der Cyberraum sind zwar getrennte und unterschiedliche Domänen der Kriegsführung mit ihren eigenen (geo-)physikalischen Eigenschaften. Gemein haben sie, dass sie die Grundlage der globalen Kommunikations- und Informationsinfrastruktur sind. Das Funktionieren sowohl der Weltwirtschaft als auch militärischer Kommandostrukturen hängen von ihr ab. Des Weiteren ist die Abhängigkeit vom Funktionieren dieser Infrastruktur dabei stetig gewachsen. Operationen im Weltall wird ein enormes Eskalationspotenzial zugesprochen, denn auf Grund der globalen und strategischen Bedeutung von Weltraum-Ressourcen kann ein regionaler Konflikt durch Weltraum-Operationen schnell zu einem globalen Konflikt anwachsen. Dennoch ist es üblich, dass Staaten (und nicht-staatliche Akteure) Satelliten und ihre Computernetzwerke angreifen (oder dieses üben). Diese Entwicklungen müssen wir als FIFF beobachten. Erfreulich ist, dass die IMI sich mit dem Thema in der Ausgabe 2/2022 beschäftigt.

Im Jahr 2021 wurde auch das Weltraumkommando der Bundeswehr in den Dienst gestellt. Es ging aus dem erst am 21. September 2020 aufgestellten *Air and Space Operations Centre* hervor. Das Kommando soll alle mit dem Weltraum verbun-

denen Aktivitäten der Bundeswehr bündeln. Im Kommando erfolgt eine enge Zusammenarbeit mit dem Kommando Cyber- und Informationsraum sowie mit dem Deutschen Zentrum für Luft- und Raumfahrt. Offiziell heißt es, sollen künftig Defensivoperationen im All geplant und geführt werden. Vordringliche Aufgabe ist die Überwachung und der Schutz der sieben Satelliten der Bundeswehr zur Kommunikation und Aufklärung (Systeme SATCOMBw bzw. SAR-Lupe; SARah) vor Schäden durch Weltraumschrott und -waffen. Seit einigen Jahren nimmt die Bundeswehr auch am internationalen Weltraum-Manöver *Schrievers Wargame* teil. Aaron Lye hat für die IMI-Ausgabe 2/2022 einen Artikel verfasst, der dort zeitnah erscheint.

Satelliten sind von entscheidender Bedeutung für die militärische Kommunikation, Frühwarnsysteme, Aufklärung und Lagebild sowie globale Positionsbestimmung und Navigation via GPS in Echtzeit. Satelliten, Bodenstationen, Starteinrichtungen etc. sind entsprechend kritische Infrastrukturen, die resilient gegenüber Angriffen gestaltet werden. Ein aktuelles Beispiel ist der informationstechnische Angriff auf den Satellit KA-SAT 9A des amerikanischen Satelliteninternet-Betreibers Viasat am ersten Kriegstag der russischen Invasion in der Ukraine. Ukrainische Behörden bestätigten, dass der Angriff auf den Satelliten zu ernststen Problemen geführt hat.

Sichere Kommunikation ist im Zeitalter von Quantencomputern ein Problem. Quantencomputer nutzen quantenmechanische Eigenschaften und arbeiten deshalb anders als klassische Computer. Dies ist bemerkenswert, da mit diesen Rechnern bestimmte Probleme schneller berechnet werden können. Zu diesen Problemen gehören solche, auf die wir aktuell bei Verschlüsselung vertrauen. Wenn es gelingt, entsprechende Quantencomputer zu bauen, dann sind wesentliche und weltweit täglich viel genutzte Verschlüsselungsverfahren unsicher. Aus diesem Grund wird nach neuen Verschlüsselungsverfahren gesucht, die von Quantencomputern nicht effizient berechnet werden können. Quantenmechanische Eigenschaften ebenfalls für Verschlüsselung zu verwenden, ist eine Idee. Aaron Lye zeigt in seinem Artikel *Quantenschlüsselverteilung: Von Glasfaser zu Satelliten* die Entwicklungen an dieser Technik auf. Der Artikel ist ein Nachdruck eines Artikels, der zuerst in der IMI-Ausgabe 2/2022 zeitnah erscheint.

Auch die 100 Mrd. Euro *Sondervermögen* (Neusprech für Schulden) der Bundeswehr werden uns noch einige Zeit beschäftigen. Obwohl die Details noch nicht klar sind, bedeuten die Investitionen doch eine wesentliche Aufrüstung für Kriege der Zukunft.



Krieg mit Künstlicher Intelligenz

Bericht über das Online-Hearing am 10. März 2022

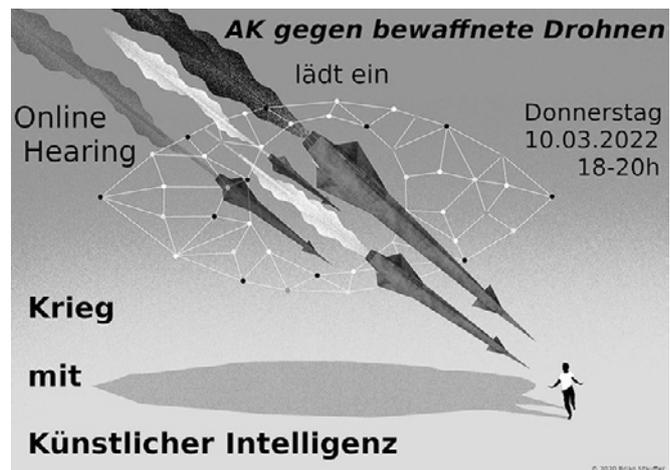
Das bisher letzte in einer Serie von Online-Hearings des Arbeitskreises Gegen bewaffnete Drohnen (<http://drohnen.frieden-und-zukunft.de/>) hat am 10. März 2022 zum Thema Krieg mit Künstlicher Intelligenz stattgefunden. Im Arbeitskreis arbeiten elf Organisationen der Friedensbewegung zusammen. Bewaffnete Drohnen, die schon seit Jahrzehnten tausendfach eingesetzt werden, sind ein Musterbeispiel für Waffensysteme, die ohne die Nutzung von Künstlicher Intelligenz (KI) undenkbar wären. Aber die Verflechtung von KI und Rüstung reicht weit darüber hinaus.

Viele Regierungen in aller Welt haben in den letzten Jahren nationale Strategien für Künstliche Intelligenz entwickelt, in denen KI zur Schlüsseltechnologie zukünftiger Wertschöpfung erklärt und mit immensen Finanzmitteln gefördert wird. In der Regel ist dabei der Einsatz im militärischen Bereich verschämt ausgeklammert. Tatsächlich aber werden seit Jahrzehnten KI-Methoden für Waffensysteme wie Killerdrohnen und für militärische Plattformen wie Battle-Management-Systeme entwickelt und eingesetzt – mit aktuell wachsender Tendenz. Nach den Plänen der Weltmächte und ihrer Militärs soll Krieg künstlich intelligent werden. Es gibt weltweit erhebliche Anstrengungen, um autonome Waffen und KI-gestützte militärische Planungs- und Entscheidungssysteme zum Einsatz bringen zu können. Es ist zu befürchten, dass die KI-Rüstung mit KI-Waffen (in Verbindung zu staatlicher bzw. universitärer Forschung und Start-Up-Unternehmen) die Gefahr von Kriegen noch einmal erheblich vergrößern wird. Denn durch neuartige *KI-basierte Waffen- und Kriegsführungssysteme* wird die *Rüstungsspirale weitergedreht, die Rüstungskontrolle erschwert und die Einsatzschwelle auf Grund vermeintlicher Überlegenheit gesenkt.*

Aus diesen Gründen hat ein breites Bündnis der Zivilgesellschaft gegen bewaffnete Drohnen zu einer öffentlichen Online-Debatte eingeladen. Das zweistündige Programm bestand aus acht Kurzvorträgen von je sieben Minuten, so dass etwa gleich viel Zeit für Fragen und Diskussion blieb. Die Vorträge waren in zwei Blöcke aufgeteilt: der erste zu der technologischen und der zweite zu der gesellschaftlichen Dimension des KI-Kriegs.

In den ersten vier Vorträgen wurde ein weiter Bogen vom Stand der Technik in die Zukunftsplanung des KI-Kriegs gespannt. Jakob Foerster (AIScientists4Peace) sprach zu *Der kurze Schritt von bewaffneten Drohnen zu autonomen tödlichen Waffen*, Christoph Marischka (IMI – Informationsstelle Militarisierung) zu *Das gläserne Gefechtsfeld*, Marius Pletsch (DFG-VK) zu *Menschen und Maschinen als Team im KI-Krieg* und Aaron Lye (FlfF – Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung) zu *Künstliche Intelligenz für den Luftkampf der Zukunft*.

Es sind wirklich nur ein paar kurze Schritte, wie es im Titel des ersten Vortrags heißt, von der Forschung an Drohnentechnologie, ihrer Beschaffung und ihrer Ausstattung mit Feuerwaffen hin zu letalen autonomen Waffensystemen (oft abgekürzt durch LAWS für *Lethal Autonomous Weapon Systems*), bei denen die letzte Entscheidung über Leben und Tod einem Computerprogramm übertragen ist. Bei den militärischen Zukunftsplanungen geht es jedoch nicht nur um die KI-gestützte Weiterentwicklung von Waffensystemen sondern auch von entsprechenden



Führungssystemen sowie der Personal- und Materialwirtschaft. Die Integration dieser Elemente führt zu „gläsernen Gefechtsfeldern“ mit der Implementierung ganz konkreter Kriegshandlungen. Eine der besonderen Schwierigkeiten bei dieser neuen Form der Kriegsführung besteht im Zusammenwirken der weitgehend eigenständig agierenden Militär- und Waffentechnik mit den beteiligten Soldat:innen. Die sogenannte Teambildung von Mensch und Maschine ist schon heute ein viel diskutiertes Problem im Militär. Verschärft wird es zutage treten bei der Entwicklung des neuen deutsch-französisch-spanischen Kampfflugzeugs (*Future Combat Air System, kurz FCAS*), das bis 2040 fertiggestellt sein soll und einschließlich Beschaffung Hunderte Milliarden Euro verschlingen wird. Bei FCAS soll ein KI-System dafür sorgen, dass ein Netz aus Kampfflugzeugen der sechsten Generation, Drohnenschwärmen, Satellitenanlagen und Cloud-Services reibungslos ineinandergreift. Vieles, was in diesem Bereich passiert und geplant ist, bleibt bisher für Außenstehende unsichtbar, eine öffentliche Debatte darüber ist aber dringend geboten.

Im zweiten Block trug Elke Schwarz (Queen Mary University of London) vor zu *Silicon Valley macht Krieg*, Jacqueline Andres (IMI – Informationsstelle Militarisierung) zu *No Cyber Valley und die Ökonomie der Angst*, Thomas Reinhold (PEASEC – Wissenschaft und Technik für Frieden und Sicherheit) zu *Die Militarisierung von KI und die Probleme für die Rüstungskontrolle* sowie Edwick Madzimore (WILPF – *Women's International League for Peace and Freedom Zimbabwe*) zu *Warum sich Frauen im globalen Süden gegen autonome Waffen engagieren*.

Ziel des zweiten Blocks war, einige tiefgreifende gesellschaftliche Dimensionen vorzustellen und zu versuchen, sie so zu konkretisieren, dass sie öffentlich debattiert und diskutiert werden

können. Dazu gehört, die Rolle der aktuellen KI-Forschung und des Militarismus für den Klimawandel und ihren Einfluss auf die Wirksamkeit von Grassroots-Bewegungen und den Lebensalltag für die Bevölkerung im globalen Süden zu untersuchen. Dazu gehört auch, in deutschen Städten die Folgen dieser Forschung auf die Straße zu bringen, wie es unter anderem das Bündnis gegen das *Cyber Valley* in Tübingen seit ein paar Jahren tut. Es geht um gesellschaftliche Konsequenzen, die durch den Bau von Komplexen wie dem Cyber Valley im Neckartal eintreten. Solche von öffentlichen Geldern geförderte Forschungscluster, die der Verschmelzung von freier Grundlagenforschung mit der Industrie und insbesondere mit Start-Up- sowie Rüstungsunternehmen dienen, entstehen seit einigen Jahren in weiten Teilen Europas, und sie stehen nicht nur in der Wahl ihrer Namensgebung dem Silicon Valley sehr nahe. Es geht konkret darum, differenziert zu zeigen, wie einzelne Instanzen des Silicon Valleys zu Kriegen unseres Jahrtausends beitragen, andere sich dem auch verweigern.

Das Hearing fand auf der datenschutzfreundlichen BigBlueButton-Instanz senfcall.de statt. Die Videokonferenzplattform ist eine Initiative von Studierenden vor allem aus Karlsruhe und Darmstadt, die eine datensparsame und sichere Alternative zu bekannten Konferenzsystemen bietet. Der Service ist DSGVO-konform und datensparsam, d. h. senfcall.de erhebt nur die Daten, die auch für den Service nötig sind. Alle Daten werden auf Servern in Deutschland verarbeitet. Das System hat mit den 90 Teilnehmerinnen und Teilnehmern einwandfrei funktioniert, so dass es für derartige Veranstaltungen sehr empfehlenswert ist.

Die Vorträge wurden von Ting Chun Liu mit der Open Source Software OBS-Studio aufgezeichnet und geschnitten. Die Video-Mitschnitte sind zu finden unter <https://media.ccc.de/c/kriegundki> und <https://vimeo.com/690465548>.

Die Begrüßung haben Christian Heck und Hans-Jörg Kreowski (FlfF – Forum InformatikerInnen für Frieden und gesellschaft-

liche Verantwortung) übernommen, moderiert haben Angelika Wilmen und Susanne Grabenhorst (IPPNW – Internationale Ärzt:innen für die Verhütung des Atomkrieges), Reiner Braun (International Peace Bureau) hatte das Schlusswort.

Dank des breiten Spektrums an interessanten Vorträgen ist die Veranstaltung gut angekommen. Die Diskussion hat gezeigt, dass der Informationsbedarf zum Thema Künstliche Intelligenz und Krieg groß ist und alle friedlich gesonnenen Menschen noch lange beschäftigen wird. Teilweise im direkten Zusammenhang mit den Vorträgen, aber insgesamt mit engem thematischen Bezug sind Ende 2021 von Hans-Jörg Kreowski und Aaron Lye zwei Publikationen unter dem Titel *Künstliche Intelligenz zieht in den Krieg* herausgegeben worden, erschienen als Dossier 93 der Zeitschrift *Wissenschaft und Frieden* und als Schwerpunkt der *FlfF-Kommunikation* 4/2021.

Das Hearing sollte dazu beitragen, ein Stück weit Einblicke in die gesellschaftlichen und in die technologischen Dimensionen von zukünftigen, aber auch von ganz aktuellen High-Tech-Elementen konventioneller Kriege zu gewinnen. High-Tech in Form von Killerdrohnen, Präzisionslenk Waffen, netzwerkzentrierten Kriegsführungstaktiken, Cyberattacken und Informationskrieg ist auch derzeit im Ukraine-Krieg wirksam, wenn auch für die Öffentlichkeit eher im Verborgenen. Einblicke in Kriegstechnologien zu gewinnen ist eine wichtige Basis der Friedensarbeit in Zeiten des Krieges – Kriege, in denen immer Menschen sterben: Soldat:innen und Zivilist:innen, Eltern und Kinder. Um Einblicke in die Rolle der „neuen“ Technologien und in ihre Funktion im Kontext der Grausamkeiten des Krieges zu erlangen, ist es wichtig, sich ein bisschen tiefer in diese kognitiven Systeme hineinzudenken und ein Stück weit sehen zu lernen, in welchem Maße diese technischen Objekte und künstlich intelligenten Systeme unterstützend wirken beim Treffen von menschlichen Entscheidungen.



Aaron Lye

Quantenschlüsselverteilung: Von Glasfaser zu Satelliten

Seit einigen Jahren wird wieder vermehrt über Quantencomputer berichtet. Die Entwicklung dieser Computertechnik wird aktuell weltweit massiv vorangetrieben und Milliarden Gelder fließen in die Erforschung und Entwicklung dieser Technologie – auch in Deutschland. Die Bundesregierung hat 2021 Fördergelder in Höhe von insgesamt zwei Milliarden Euro für die Entwicklung von Quantencomputern freigegeben¹, dazu kommen Fördergelder aus EU-Projekten. Öffentlich argumentiert wird das mit der industriellen Nutzung dieser Technologie, beispielsweise zur Bereitstellung sicherer Kommunikation, dem Lösen von schweren kombinatorischen Problemen (z. B. in der Logistik), sowie Materialforschung. Die Auswirkungen dieser Technologie auf die Kryptographie (Verschlüsselung) sind allerdings gravierend. Geschichtlich wie aktuell sind die Kryptographie und die Raumfahrt untrennbar mit Geheimdiensten, Militär und nationalistischer Politik verbunden. Die im Folgenden dargestellten Entwicklungen im Bereich der Quantenverschlüsselung haben daher immer auch geheimdienstliche und militärische Relevanz.

Quantencomputer nutzen quantenmechanische Eigenschaften und arbeiten deshalb anders als klassische Computer. Dies ist bemerkenswert, da mit diesen Rechnern bestimmte Probleme schneller berechnet werden können. Zu diesen Problemen gehören jene, auf die wir aktuell bei Verschlüsselung vertrauen. Das ist seit 1994 in Fachkreisen bekannt (und führte zu einem Schub dieser Technologie). Wenn es gelingt, entsprechende Quantencomputer zu bauen, dann sind wesentliche, weltweit täglich

und viel genutzte Verschlüsselungsverfahren unsicher. Und da Kommunikation abgehört und gespeichert werden kann, kann sie auch nachträglich entschlüsselt werden. Aus diesem Grund wird nach neuen Verschlüsselungsverfahren gesucht, die von Quantencomputern nicht effizient berechnet werden können. Ein Ansatz ist, quantenmechanische Eigenschaften ebenfalls für Verschlüsselung zu verwenden. Photonen (Lichtteilchen) haben diese Eigenschaften und lassen sich einfach kontrolliert

durch Laser erzeugen, übertragen und durch einen Sensor messen. So lassen sich kryptografische Schlüssel austauschen, um damit Nachrichten zu verschlüsseln. Die Idee der Verwendung von Quanteneffekten zum Austausch von kryptographischen Schlüsseln wurde bereits 1983 publiziert.² Sie besteht im Wesentlichen darin, eine Reihe von Photonen paarweise zu koppeln (genauer: in Superposition zu verschränken) und eines von jedem Paar zu übertragen. Eine Messung des übertragenen Photons bewirkt eine Zustandsänderung beider Photonen des Paares. Die dazu nötigen technischen Voraussetzungen existieren schon lange. Aber dadurch bestimmen die konkreten physikalischen Bedingungen des Netzwerks die Verschlüsselung.

Anbieter von Quantenverschlüsselung, Forscher:innen als auch die Medien stellen allerdings gelegentlich die kühne Behauptung auf, dass diese Technologie garantierte Sicherheit auf der Grundlage der physikalischen Gesetze biete. Die tatsächliche Sicherheit dieser Systeme ist aber nicht die theoretische Sicherheit, die sich aus den Gesetzen der Physik ergibt (wie modelliert und oft suggeriert wird), sondern die begrenzte Sicherheit, die durch Hardware- und Technikdesigns erreicht werden kann. Es gibt diverse sicherheitstechnische Probleme³, und es existieren auch andere Verfahren, welche wesentlich kostengünstiger sind und ein besser bekanntes Risikoprofil aufweisen. Trotzdem wird an der Entwicklung festgehalten.

1984 wurde bei IBM das erste quantenmechanische Protokoll zur Übertragung dieser Schlüssel entwickelt.⁴ 1991 konnte es erstmals erfolgreich demonstriert werden. Die Distanz zwischen Sender und Empfänger, welche durch eine Glasfaserleitung miteinander verbunden waren, betrug 32 cm. Seitdem hat sich durch kontinuierliche Forschung und Entwicklung weltweit einiges getan.

Ab 2004 entstanden die ersten größeren Glasfasernetze für diese neue Art des Schlüsselaustauschs. Das erste wurde von der DARPA entwickelt, bestand aus 10 Stationen und war drei Jahre in Massachusetts, USA, in Betrieb.⁵ In Europa sind insbesondere das 2008 von der EU finanzierte Glasfasernetz SECOQC (*Secure Communication Based on Quantum Cryptography*), welches sieben Standorte in Wien und Umgebung miteinander über Glasfaserkabel verband⁶ als auch das von Id Quantique von 2009 bis 2011 im Großraum Genf, Schweiz, installierte Glasfasernetz⁷ zu nennen. Ebenfalls im Jahr 2009 wurde in Wuhu, China, ein hierarchisches Netzwerk demonstriert, welches vier Teilnetze miteinander verband.⁸ 2010 wurde das Tokioter Netzwerk eingeweiht⁹, und auch in Russland gab es ab 2014 ein solches Netzwerk.¹⁰

Nachdem jahrzehntelang an Übertragungen von Schlüsseln per Glasfaser experimentiert wurde, entstanden weltweit kommerzielle Dienste. Bei Verwendung von Glasfasertechnologien ist die Entfernung zwischen Sender und Empfänger recht beschränkt. Durch Satellitenkommunikation kann diese Entfernung wesentlich vergrößert werden.

Im Juni 2017 haben chinesische Physiker:innen von der University of Science and Technology of China im Rahmen des Projekts *Quantum Experiments at Space Scale* zum ersten Mal verschränkte Photonen über eine Entfernung von 2400 km zwischen zwei Bodenstationen gemessen und damit die Grundlage für zukünftige interkontinentale Experimente zur Quantenschlüssel-



Nichtlokalität und Verschränkung – die unglaubliche Welt der Quantenteilchen, Foto: Sharon Apted, CC0 1.0

verteilung gelegt. Die Photonen wurden von einer Bodenstation zu dem 1200 km entfernten Satelliten (Micius genannt) und zurück zu einer anderen Bodenstation geschickt.¹¹ Das Experiment war Teil der im August 2016 gestarteten Weltraummission QUESS, welche wenig später einen internationalen Quantenschlüsselaustausch zwischen der University of Science and Technology of China und dem Institut für Quantenoptik und Quanteninformation in Wien, Österreich, ermöglichte.¹² Im Oktober 2017 wurde eine 2.000 km lange Glasfaserleitung zwischen Peking, Jinan, Hefei und Shanghai in Betrieb genommen.¹³ Zusammen bilden sie das weltweit erste Satelliten-gestützte Netzwerk zur Quantenschlüsselverteilung¹⁴. Bis zu 10 Micius/QUESS-Satelliten sollen zunächst bis 2020 ein europäisch-asiatisches Netzwerk und bis 2030 ein globales Netzwerk ermöglichen.¹⁵

In einem ähnlichen Zeitraum hat auch Japan mit dem *Small Optical Transponder (SOTA)* Laser-Kommunikationsterminal an Bord des Satelliten SOCRATES zunächst die Fähigkeiten mit Laser-basierter Datenübertragung vom Weltraum zum Boden demonstriert.¹⁶ Die Experimente beinhalteten keinen Quantenschlüsselaustausch. Entsprechende Experimente und Implementierung sind allerdings sehr wahrscheinlich, da das japanische Unternehmen Toshiba bereits seit über 20 Jahren ebenfalls an dieser Technologie forscht und entwickelt. Es ist auch Projektpartner beim aktuell laufenden EU-Projekt OPENQKD, bei dem es um die Infrastruktur für Quantenschlüsselverteilung geht.

Auch die Europäischen Weltraumorganisation (European Space Agency, ESA) hat früh angefangen, Satelliten für Quantenschlüsselverteilung zu entwickeln¹⁷. Die Aktivitäten sind eingebettet in der Errichtung einer Quantenkommunikationsinfrastruktur von 24 EU-Mitgliedstaaten, die innerhalb der nächsten 10 Jahre entstehen soll (EuroQCI bzw. QCI4EU). Diese soll aus weltraumgestützten und terrestrischen Systemen bestehen.

Das ESA-Programm ScyLight startete 2016 als spezielles Programm für optische Kommunikation, einschließlich Technologien der Quantenkryptographie und als Demonstration erster Dienste. 2019 startete die SAGA-Mission (*Security And cryptographic mission*), bei der die Entwicklung des Weltraumsegments des EuroQCI, die Satelliten und Bodenstationen wesentlich waren.

2018 wurde bekannt, dass die ESA zusammen mit einem europäisch-kanadischen Industriekonsortium geleitet von dem britischen Start-up *Arqit Ltd* einen low-orbit Satelliten für Quantenschlüsselaustausch bauen will (QKDSat)¹⁸. Arqit Ltd hat darüber hinaus noch eigene Pläne. 2023 will es zwei solcher Satelliten vom Weltraumbahnhof Cornwall (GB) aus an Bord des *LauncherOne* von Virgin Orbit starten. Sie sollen Teil des bereits existierenden regionalen, kommerziellen Netzwerks für Quantenschlüsselverteilung über Glasfaserkabel werden.¹⁹

Die kanadische Weltraumbehörde (CSA) arbeitet ebenfalls seit 2017 mit dem Institute for Quantum Computing (IQC) der University of Waterloo zusammen an dem *Quantum Encryption and Science Satellite* (QEYSSat) Projekt.²⁰ Während IQC die wissenschaftliche Expertise liefert, soll Honeywell zusammen mit Loft Orbital die Plattform für den Satelliten liefern.

Im Mai 2021 gab ein Team von Forschern aus Kanada und Großbritannien bekannt, dass sie ein gemeinsames System entwickeln, welches nach 2022 an Bord des QEYSSat getestet werden soll. Ziel der Forscher:innen ist es, Schlüssel zwischen Bodenstationen auf beiden Seiten des Atlantiks zu übertragen.²¹

Auch Russland arbeitete an einem entsprechenden Satellitenprogramm zur Quantenschlüsselverteilung. Ein Prototyp eines Satelliten wurde 2020 entwickelt. 2023 soll der erste Satellit gestartet werden. Die Aktivitäten sind Teil mehrerer Roskosmos-Programme als auch des Complex-SG-Projekts (2019-2023), welches Russland mit Belarus betreibt.²² Erklärtes Ziel ist die Entwicklung transkontinentaler Quantenschlüsselverteilung und die Zusammenführung von russischer, chinesischer und europäischer Infrastruktur. Aufgrund des Angriffskriegs auf die Ukraine ist unklar, ob der Satellit wirklich 2023 gestartet wird. Die Zusammenführung mit europäischer Infrastruktur ist äußerst unwahrscheinlich.

Vor kurzem hat die National Aeronautics and Space Agency (NASA) durch das National Space Quantum Laboratory (NSQL) begonnen, eine Technologie zu entwickeln, um satellitengestützten Quantenschlüsselaustausch zu ermöglichen und eine entsprechende Infrastruktur auf der Internationalen Raumstation zu schaffen.²³

Sowohl die indische Defence Research and Development Organisation als auch die Indian Space Research Organisation demonstrierten 2020/2021 Quantenkommunikation zwischen Laboren²⁴. Aktuell plant Indien die Entwicklung der satellitengestützten Quantenkommunikation.²⁵

Diese Aktivitäten belegen, dass die seit Jahrzehnten stattfindende Forschung und Entwicklung im Bereich Quantencomputing und Quanteninformation längst kein rein akademisches Thema mehr ist. Des Weiteren wird trotz dessen, dass es eine alternative, kostengünstigere und besser verstandene quantenre-

sistente Kryptographie gibt (die Post-Quantum-Kryptographie), ebenfalls an der Quantenschlüsselverteilung festgehalten. Mehr noch herrscht eine ähnliche Situation wie in den 1960er-Jahren, bei der Staaten ihre Fähigkeiten (auch im Weltraum) demonstrieren wollen und niemand zurückbleiben möchte.

Anmerkungen

- 1 Bundesregierung stellt zwei Milliarden Euro für Quantencomputer bereit. *Spiegel Online*. 11.05.2021
- 2 Stephen Wiesner, *Conjugate Coding*, *SIGACT News*, Vol. 15, No. 1, 1983, pp. 78-88. doi10.1145/1008908.1008920
- 3 National Security Agency/Central Security Service Search NSA 2020. *Quantum Key Distribution (QKD) and Quantum Cryptography (QC)*. nsa.gov; siehe auch e.g. Vakhitov, Makarov, and Hjelme, *Large pulse attack as a method of conventional optical eavesdropping in quantum cryptography*, *Journal of Modern Optics* 48, 2001; Makarov and Hjelme, *Faked states attack on quantum cryptosystems*, *Journal of Modern Optics*, vol. 52, 2005; Ferenczi, Grangier, Grosshans, *Calibration Attack and Defense in Continuous Variable Quantum Key Distribution*, *CLEO-IQEC*, 2007; Zhao, Fung, Qi, Chen, and Lo, *Experimental demonstration of time-shift attack against practical quantum key distribution systems*, *Physical Review A* vol. 78, 2008; Scarani and Kurtsiefer, *The black paper of quantum cryptography: Real implementation problems*, *Theoretical Computer Science* (560) 2014.
- 4 Charles H. Bennett and Gilles Brassard. *Quantum cryptography: Public key distribution and coin tossing*. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, volume 175, page 8. New York, 1984.
- 5 Knight, Will. *Quantum cryptography network gets wireless link*. 07.06.2005. Entwickelt wurde das DARPA-Netz von BBN Technologies, der Harvard University und der Boston University, in Zusammenarbeit mit IBM Research, dem National Institute for Standards and Technologies und QinetiQ.
- 6 Projektwebsite secoqc.network
- 7 Patrick Eraerds, et al. *Quantum key distribution and 1 Gbit/s data encryption over a single fibre*. arXiv:0912.1798 [quant-ph] 2009
- 8 Xu, FangXing; Chen, Wei; Wang, Shuang; Yin, ZhenQiang; Zhang, Yang; Liu, Yun; Zhou, Zheng; Zhao, YiBo; Li, HongWei; Liu, Dong. *Field experiment on a robust hierarchical metropolitan quantum cryptography network*, *Chinese Science Bulletin*, 54 (17): 2991–2997, arXiv:0906.3576. 2009
- 9 Projektwebsite www.uqcc2010.org/highlights/index.html. Das Tokio-ter QKD-Netzwerk entstand durch eine internationale Zusammenarbeit zwischen sieben Partnern: NEC, Mitsubishi Electric, NTT und NICT aus Japan sowie Toshiba Research Europe Ltd. (UK), Id Quantique (Schweiz) und All Vienna (Österreich).
- 10 Vladimir I. Egorov. *Quantum communication in Russia: status and perspective*. Präsentation beim ITU Workshop on Quantum Information Technology (QIT) for Networks. Shanghai, China, 5-7 June 2019. https://www.itu.int/en/ITU-T/Workshops-and-Seminars/2019060507/Documents/Vladimir%20Egorov_Presentation.pdf

Aaron Lye

Aaron Lye hat an der Universität Bremen Informatik studiert und dort auch Ende 2021 seine Promotion abgeschlossen. Er ist seit Jahren beim FIF aktiv.

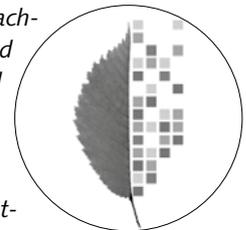
- 11 Juan Yin et al. *Satellite-based entanglement distribution over 1200 kilometers*. *Science*. 356 (6343): 1140–4. arXiv:1707.01339. doi:10.1126/science.aan3211. 2017
- 12 Lin Xing. *China launches world's first quantum science satellite*. *Physics World*. Institute of Physics. 16.08.2016
- 13 Wall, Mike. *China Launches Pioneering 'Hack-Proof' Quantum-Communications Satellite*. *Space.com*. *Purch*. 16.08.2016
- 14 Amy Nordrum. *China Demonstrates Quantum Encryption By Hosting a Video Call*. *IEEE*. 03.10.2017.
- 15 Jeffrey Lin; P.W. Singer; John Costello. *China's Quantum Satellite Could Change Cryptography Forever*. *Popular Science*. 03.03.2016
- 16 Dimitar R. Kolev and Morio Toyoshima, *Satellite-to-ground optical communications using small optical transponder (SOTA) – received-power fluctuations*, *Opt. Express* 25, 28319-28329. 2017
- 17 Eric Wille. *Space based QKD at ESA*. Präsentation beim ITUWebinar Quantum information technology – Episode 2: Joint Symposium on Quantum Transport Technology. 28.04.2021
- 18 ESA. *Secure communication via quantum cryptography*. *Esa.int*; Die ESA entwickelt QKDSat mit ArQit. ArQit leitet ein Industriekonsortium, dem folgende Unternehmen angehören: QinetiQ (Belgien), British Telecom und Teledyne e2v (Vereinigtes Königreich) sowie mehrere Akteure aus Deutschland, Österreich, Kanada, der Tschechischen Republik und der Schweiz.
- 19 Arqit space.com; Kürzlich gab ArQit die Zusammenarbeit mit dem US-Verteidigungsunternehmen Northrop Grumman und dem britischen Telekommunikationsbetreiber BT bekannt.
- 20 Projektwebsite <https://uwaterloo.ca/institute-for-quantum-computing/qeysat>
- 21 Siehe Arqit
- 22 Siehe Egorov
- 23 Joseph D. Touch, Lori W. Gordon. *Quantum Key Distribution in Space. Game Changer*. Center for Space, Policy and Strategy. Juli 2020. *space.org*; Die Nationale Quanteninitiative der USA (NQI) mit dem 1,2 Milliarden Dollar Jahresbudget ist wesentlicher Akteur bei der Finanzierung von Quantumcomputing und Quantuminformationsprojekten. 30 Millionen Dollar sind für die Quanten Kommunikation konzentriert, davon 3 Millionen US-Dollar für QKD. Allerdings sind derzeit zweistellige Milliardenbeträge an neuen Finanzmitteln für zivile Forschung und Entwicklung im Bereich der „Zukunftsindustrien“, einschließlich künstlicher Intelligenz und Quanten Informationswissenschaft geplant.
- 24 Ministry of Defence. *Quantum Communication between two DRDO Laboratories*. Press Information Bureau. 09.12.2020
- 25 ISRO makes breakthrough demonstration of free-space Quantum Key Distribution (QKD) over 300 m. *Indian Space Research Organisation*. 22.03.2021



Gemeinsame Pressemitteilung von 13 Organisationen aus Umwelt- und Digitalpolitik, Entwicklungszusammenarbeit und Wissenschaft

Digitalisierung und Nachhaltigkeit zusammendenken – Zweite bundesweite Bits & Bäume-Konferenz vom 30. September bis 2. Oktober 2022

7. Juni 2022, Berlin – In diesem Jahr findet die zweite Bits & Bäume-Konferenz für Digitalisierung und Nachhaltigkeit in Berlin statt. Mehrere Organisationen aus Umweltschutz, Digitalpolitik, Entwicklungspolitik und Wissenschaft laden dazu ein, Handlungsoptionen und politische Forderungen für ausreichend Klima- und Umweltschutz, soziale Gerechtigkeit und Demokratie im digitalen Zeitalter zu erarbeiten. Ein ganzes Konferenzwochenende dreht sich um die Frage, wie die Digitalisierung zu einer nachhaltigen und demokratischen Gesellschaft beitragen kann. Ziel der Veranstalter ist es, konkrete Beiträge zu diskutieren, wie eine global, wirtschaftlich, sozial und ökologisch gerechte Zukunft in der digitalisierten Welt aussehen kann. Auf der Vernetzungskonferenz werden rund 2.000 Interessierte erwartet.



Die Bewegung *Bits & Bäume* hatte sich nach der ersten Konferenz im Jahr 2018 gebildet und setzt sich seither dafür ein, Digitalisierung und Nachhaltigkeit in Einklang zu bringen. Denn aktuelle Digitalisierungstrends verschärfen gesellschaftliche Ungerechtigkeiten, Umweltzerstörung und Demokratieversagen: Tech-Monopole tragen durch ihre Technologien und Geschäftsmodelle stark zu Konsumsteigerungen sowie Energie- und Ressourcenverbrauch bei. Zudem verstärken sie die Polarisierung der Gesellschaft. Das Bits & Bäume-Bündnis eint die Überzeugung, dass eine andere Digitalisierung möglich und dringend notwendig ist. Doch viele politische Akteur:innen sind zu zögerlich bei der Umsetzung entsprechender Maßnahmen, so die Veranstalter. Daher müsse es nun verstärkt darum gehen, politische Gestaltungsmacht zurückzugewinnen und zu nutzen, damit die Digitalisierung dazu beitragen kann, den dringend notwendigen sozial-ökologischen Umbau von Gesellschaft und Wirtschaft zu unterstützen.

„Der Strom- und Ressourcenverbrauch digitaler Geräte und Infrastrukturen steigt ungebremst. Gleichzeitig gelingt die Energiewende nur mithilfe digitaler Technologien. Und das sind nur zwei Gründe, warum wir die Bits & Bäume so dringend brauchen“, sagt Hendrik Zimmermann, Senior Advisor für Digitalisierung, Demokratie und Nachhaltigkeit bei Germanwatch, das die Konferenz mitveranstaltet.

Auch für Rainer Rehak, Mitveranstalter und Ko-Vorsitzender des Forums InformatikerInnen für Frieden und gesellschaftliche Verantwortung (IfF), steht eine gemeinsame Lösung der drängenden gesellschaftlichen Fragen im Fokus: „Es ist wichtiger denn je, dass sich die kritische Tech-Szene, die Nachhaltigkeits- und die Umweltszene zusammenschließen und Allianzen mit zugewandten Akteur:innen in Politik und Wirtschaft schmieden, um endlich adäquat der Klimakatastrophe zu begegnen und eine lebenswerte digitale Gesellschaft für alle zu ermöglichen.“



Foto: Santiago Engelhardt, CC BY

Mit Blick auf den Umwelt- und Naturschutz fügt Antje von Brock, Geschäftsführerin beim Bund für Umwelt und Naturschutz Deutschland (BUND) hinzu: „Wir stehen mit der Klima- und Ressourcenkrise und dem Artensterben vor den größten Herausforderungen unserer Zeit und das Zeitfenster zum Handeln wird immer kleiner. Es braucht mehr Klima- und Umweltschutz, es braucht mehr Ehrgeiz, um das 1,5-Grad-Ziel zu erreichen. Damit Digitalisierung einen Beitrag für eine sozial-ökologisch gerechte Zukunft leisten kann, müssen wir sie so gestalten, dass sie eine faire digitale Gesellschaft für alle ermöglicht und im Einklang mit den planetaren Grenzen steht.“

Einreichungen

Noch bis zum 21. Juni 2022 werden Bewerbungen für Konferenzbeiträge angenommen. Angesprochen sind Menschen aus zivilgesellschaftlichen Organisationen, progressive Akteur:innen aus der Wissenschaft, Wirtschaft und Politik, die zu Themen an der Schnittstelle von Nachhaltigkeit und Digitalisierung arbeiten: Der Call for Participation reicht von vernetzter Mobilität über den Rohstoff- und Strombedarf digitaler Geräte, von Datensuffizienz und kooperativen Plattformökonomien bis hin zu „smarten“ Energienetzen oder der Analyse und Verhinderung von Monopolisierung in der digitalen Ökonomie. Zudem können nachhaltige Firmen, Start-Ups und Projekte ihre Ideen bei dem Wettbewerb *Pitch & Thrive for Sustainability* vorstellen.

Tickets

Der Ticketverkauf startet am 4. Juli 2022

Über Bits & Bäume

Bits & Bäume ist das größte deutschlandweite zivilgesellschaftliche Bündnis an der Schnittstelle zwischen Digitalisierung und Nachhaltigkeit. Es besteht aus 13 Organisationen aus den Bereichen Umweltschutz, Digitalpolitik, Entwicklungspolitik und Wissenschaft sowie einer wachsenden Community. Den Auftakt bestreift Bits & Bäume mit der ersten bundesweiten Konferenz 2018 und erarbeitete in diesem Rahmen elf politische Forderungen für eine ökologische, gerechte, friedliche und demokratische Digitalisierung. Seitdem hat sich eine Vielzahl von lokalen Ablegern

gegründet. Das Buch *Was Bits und Bäume verbindet* erschien 2019. So hat Bits & Bäume die Umwelt-, Digital- und Klimabewegungen zusammengebracht und einen weiteren Meilenstein für mehr Bewusstsein der Öffentlichkeit und Politik gelegt.

Die beteiligten Organisationen

Bund für Umwelt und Naturschutz Deutschland (BUND), Brot für die Welt, Chaos Computer Club (CCC), Deutscher Naturschutzring (DNR), Einstein Center Digital Future / Technische Universität Berlin, Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FIfF), Free Software Foundation Europe (FSFE), Germanwatch, Institut für ökologische Wirtschaftsforschung (IÖW), Konzeptwerk Neue Ökonomie, Open Knowledge Foundation Deutschland (OKF), Weizenbaum Institut.

Die Veranstaltung wird gefördert durch die Deutsche Bundesstiftung Umwelt (DBU). Die Arbeit von Germanwatch und des Forums InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FIfF) im Rahmen der Bits & Bäume-Konferenz 2022 wird gefördert durch die Stiftung Mercator.

Kontakt und weitere Informationen

Webseite und Konferenzprogramm:

<https://bits-und-baeume.org/konferenz>

Ticketverkauf ab 4. Juli 2022:

<https://events.fairetickets.de/gedlv>

Forderungen von 2018:

<https://bits-und-baeume.org/forderungen>

E-Mail: presse@bits-und-baeume.org

Twitter: [@bitsundbaeume](https://twitter.com/bitsundbaeume), Hashtag: [#bitsundbaeume22](https://twitter.com/bitsundbaeume22)

Mastodon:

https://mastodon.bits-und-baeume.org/@bits_und_baeume

Aufruf für Konferenzbeiträge:

https://bits-und-baeume.org/downloads/call_for_participation.pdf

Einreichung von Konferenzbeiträgen:

<https://pretalx.com/bitsundbaeume/cfp>



Foto: Santiago Engelhardt, CC BY

Verleihung der Weizenbaum-Medaille 2021

Dreimal hat das FIfF in seiner nun fast 35-jährigen Geschichte einen FIfF-Preis verliehen:

- 1998 an **Joseph Weizenbaum**, für seine Verdienste und seinen Einsatz für Verantwortung in der Informatik. Seine Auseinandersetzung mit ethischen Fragen des Computereinsatzes und der Technikentwicklung haben viele Menschen innerhalb und außerhalb der Informatik angeregt, sich selbst damit zu beschäftigen, die eigene Verantwortung zur Kenntnis zu nehmen und sich ihr zu stellen;
- 2001 an **David Lorge Parnas**, der 1985 den Beraterkreis des US-Präsidenten Ronald Reagan für die Strategic Defense Initiative (SDI) verlassen und damit ein vorbildliches Beispiel für die Wahrnehmung professioneller Verantwortung als Informatik-Experte gegeben hat;
- 2018 – erstmals als Weizenbaum-Medaille – an **Wolfgang Coy**, der das Fach Informatik und Gesellschaft in Forschung und Lehre einzigartig, beispielhaft und maßgeblich ausgestaltet hat, unter Einbeziehung wissenschaftstheoretischer, sozial- und kulturgeschichtlicher, medientheoretischer, fachdidaktischer, ethischer und philosophischer Gesichtspunkte.

2021 haben wir zum zweiten Mal die Weizenbaum-Medaille vergeben. Mit der Weizenbaum-Medaille zeichnet das FIfF Persönlichkeiten aus, die sich in besonderer Weise um das Themengebiet Informatik und Gesellschaft durch wissenschaftliche Leistungen, politisches Wirken und persönliches Handeln verdient gemacht haben oder durch ihr Handeln dazu beitragen, die Anwendung der Informatik am Nutzen der Gesellschaft und der Menschen auszurichten.

Ein kritischer Blick auf die Informatik und ihre Erzeugnisse ist wichtiger denn je. Nachdem wir nach einem Höhepunkt zu Beginn der 1990-er einen stetigen Abbau auf diesem Gebiet beobachten mussten, erhoffen wir uns von zwei Institutionen in Ber-

lin – dem Humboldt-Institut für Internet und Gesellschaft und dem Weizenbaum-Institut für die vernetzte Gesellschaft – wieder wichtige Impulse. Doch Berlin ist nicht genug: Alle Informatik-Fakultäten, so unsere Forderung, müssen Lehre und Forschung zu den gesellschaftlichen Auswirkungen der Informatik ermöglichen und fördern.

2021 haben wir die Weizenbaum-Medaille an Professorin **Christiane Floyd** verliehen.



Christiane Floyd hat bei der Gründung des FIfF 1984 das Amt der Gründungsvorsitzenden übernommen. Ihr programmatischer Text *Wo sind Grenzen des verantwortbaren Computereinsatzes?* ist fast 40 Jahre später immer noch aktuell. Der Text zeigt die Bedeutung ethischer Fragen für die Entwicklung und Nutzung informationstechnischer Systeme und für die Informatik als Disziplin. Auch in der Softwaretechnik hat sie die menschengerechte Gestaltung informationstechnischer Systeme ganz im Sinne des FIfF in den Vordergrund gerückt. Damit und weit darüber hinaus hat sie wesentliche Beiträge für das Themengebiet Informatik und Gesellschaft geleistet.

Die folgenden Seiten enthalten die Laudatio, die von Britta Schinzel gehalten und von ihr gemeinsam mit Hans-Jörg Kreowski verfasst wurde. Darauf folgt die Ansprache der Preisträgerin Christiane Floyd.

Britta Schinzel und Hans-Jörg Kreowski

Laudatio für Prof.in Dr.in Dr.in h.c. Christiane Floyd anlässlich der Verleihung der Weizenbaum-Medaille 2021

Christiane Floyd hat den in seiner Geschichte zweiten Weizenbaum-Preis des Forums InformatikerInnen mit der Weizenbaum-Medaille für ihre außerordentlichen Verdienste um die Informatik im gesellschaftlichen Kontext erhalten.

Das FIfF stiftet den Weizenbaum-Ehrenpreis in Erinnerung an den Wissenschaftler und Informatik-Pionier Prof. Dr. Joseph Weizenbaum, der auch zum Guru des gesellschaftlichen Engagements in der Informatik und der Künstlichen Intelligenz geworden ist. Mit der Vergabe des Preises wollen wir nebst der Betonung der Bedeutung der Informatik für die gesellschaftliche Entwicklung auf eine kritische öffentliche Auseinandersetzung mit den Erkenntnissen und Artefakten der Informatik dringen.

Christiane Floyd verdient unsere Weizenbaum-Medaille in vielfacher Hinsicht:

- Sie hat sich 1984 als Initiatorin und als treibende Kraft an der Gründung des FIfF beteiligt und wurde zu dessen erster Vorsitzenden gewählt.
- Sie war eine der ersten, die auf der Einbeziehung des Menschen und des Sozialen in die informatische Arbeit bestand

und ethische Gesichtspunkte und Verantwortungnahme für die Software-Produkte einforderte.

- Sie verdient den Preis als erste Professorin in der Informatik im deutschen Sprachraum,
- als Pionierin in vielen Bereichen des Software Engineering, der partizipativen Software-Gestaltung und als Vorbereiterin der Open-Source-Bewegung,
- mit ihrem Engagement für Frauen in der Informatik und
- dazu auch als mit Joseph Weizenbaum und Heinz von Förster befreundete Informatik/KI/Kybernetik-Kollegin.
- Schließlich und nicht zuletzt passte ihr beständiges Engagement ausgezeichnet in unser diesjähriges Tagungsmotto *Selbstbestimmung in digitalen Räumen*. Denn Entwurf und Modellierung, die in digitalen Räumen Selbstbestimmung ermöglichen, erfordern die Einbeziehung von Benutzenden und auch die von nicht benutzenden Betroffenen.

Zum letzten Punkt eine kleine Anekdote aus Brittas Umgebung:

Ihr 93-jähriger Freund besitzt ein Seniorentelefon, das nicht smart ist, er hat keinen Computer, keine E-Mail-Adresse. Mit seinem wachen Verstand, der alle Aktualitäten einschließlich der Buzz-Wörter der digitalen Gesellschaft kennt, aber nichts davon je genutzt hat, und mit seinem unglaublichen Gedächtnis schlägt er jede Google-Nachschlagende, ... – und trotzdem sagt er, dass er kein Mensch mehr ist. „Weiß ich was ein Mensch ist? Weiß ich, wer das weiß? Ich weiß nicht, was ein Mensch ist, ich kenne nur seinen Preis“ (Bert Brecht). Ein Mensch ist, wer sich beispielsweise für einen Corona-Test anmelden kann. Man geht zu einer Teststation, legt den Pass vor, nach aller Datenaufnahme wird man nach der Handy-Nummer gefragt: „Was, Sie haben kein Smartphone? Dann können wir Sie nicht registrieren. Haben Sie wenigstens eine E-Mail-Adresse?“ – „Nein, auch keinen Computer.“ – „Tut uns leid, wir können Sie nicht testen.“

Diese Entmenschlichung zeigt exemplarisch, worunter man in den kapitalistisch verseuchten digitalen Räumen leidet: unsichtbare, weit verzweigte Herrschaft, Wurzellosigkeit und Kälte, wie Horkheimer und Adorno in der Dialektik der Aufklärung feststellen. Sind wir so machtlos? Gerade wir als Informatiker:innen? Keineswegs, Christiane hat immer wieder gezeigt, was wir tun können, wo wir unsere Aufmerksamkeit hinlenken sollen, worin unsere Macht als Profis besteht.

Und damit zu unserer Preisträgerin:

Christiane Floyd ist in Wien geboren, studierte dort Mathematik und promovierte 1966 mit einem Thema aus der Algebra. Dann ging sie nach München, um bei Siemens an der Entwicklung eines ALGOL-Compilers mitzuarbeiten. Von 1968 bis 1973 arbeitete sie als Wissenschaftliche Mitarbeiterin an der Universität Stanford im KI-Projekt DENDRAL von Edward Feigenbaum. Danach ging sie als Leiterin des Bereichs Methodenentwicklung der Firma Softlab nach München zurück. Sie hat dort maßgeblich die weltweit erste Programmierungsumgebung *Maestro* mitentwickelt. Spätestens da fiel ihr auf, dass die Arbeit der Ent-

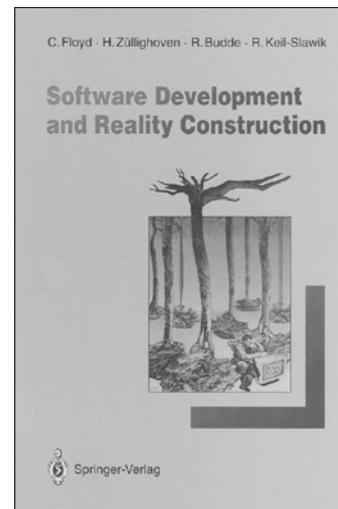


Christiane Floyd (*1943), österreichische Informatikerin (2015), Foto: WiseWoman, CC BY-SA 4.0

wickler – wir verwenden hier absichtlich nur die männliche Form – durch die Zentrierung auf Formales und Technik entfremdet war. Daher plädierte sie, dass Informatik keine reine Technikwissenschaft sein darf, dass Grundlagen aus den Sozial- und Geisteswissenschaften integriert werden und dass Entwickelnde und Anwendende in den Dialog treten müssen.

1978 erhielt sie einen Ruf an die Technische Universität Berlin, wo sie die Leitung der Gruppe Softwaretechnik übernahm. Mit STEPS, der *Softwaretechnik für evolutionäre, partizipative Systementwicklung* begründete sie die Einbeziehung der Benutzenden in den gesamten Software-Entwicklungsprozess. 1991 nahm sie einen weiteren Ruf an die Universität Hamburg an, wo sie ihre bahnbrechenden Arbeiten bis zu ihrer Pensionierung 2008 fortsetzte. Ihre wissenschaftliche Leistung ist u. a. in folgenden Publikationen dokumentiert:

(1) C. Floyd, F.-M. Reisin, G. Schmidt: *STEPS to Software Development with Users*. In: C. Ghezzi, J. A. McDermid (Eds.): *ESEC '89, Lecture Notes in Computer Science no. 387*. Springer, Berlin Heidelberg 1989, S. 48–64.



(2) C. Floyd, H. Züllighoven, R. Budde, R. Keil-Slawik (Hrsg.): *Software Development and Reality Construction*. Springer Verlag, Berlin, 1992.

(3) *Software Development Process – Some Reflections on the Cultural, Political and Ethical Aspects from a Constructivist Epistemology Point of View. In: Cybernetics & Human Knowing – A Journal of second-order cybernetics autopoiesis and cyber-semiotics. Vol. 6, No. 2, 1999, S. 5–18.*

Christiane Floyd hat ihr Fach Softwaretechnik in Forschung und Lehre einzigartig geprägt. Mit ihren genannten Anforderungen hat sie die Softwaretechnik maßgeblich beeinflusst und unter Einbeziehung sozialwissenschaftlicher, ethischer und philosophischer Gesichtspunkte beispielhaft ausgestaltet.

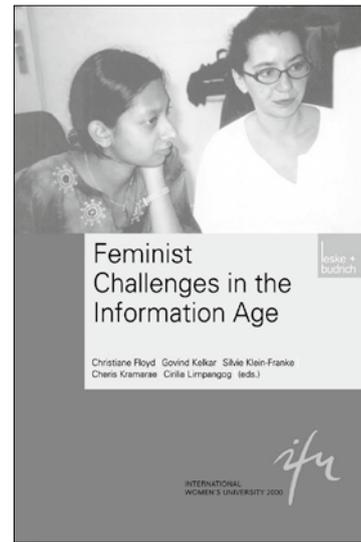
Deshalb wurde sie 2012 zur Honorarprofessorin der Technischen Universität Wien bestellt, 2017 erhielt sie die Ehrendoktorwürde der Fakultät für Elektrotechnik, Informatik und Mathematik der Universität Paderborn, 2020 erhielt sie von der Klaus-Tschira-Stiftung und der Gesellschaft für Informatik die Klaus-Tschira-Medaille.



Erste deutsche Informatikerin Prof. Christiane Floyd erhält Ehrendoktorwürde der Uni Paderborn, zu sehen: Prof. Dr. Birgit Riegraf, Vizepräsidentin für Lehre, Studium und Qualitätsmanagement, Prof. em. Dr. h. c. Christiane Floyd und Prof. Dr.-Ing. Reinhard Keil, Dekan
Foto: Universität Paderborn, Johannes Pauly

Sie hat mehrfach die jährlich stattfindende Informatica Feminale in Bremen als Dozentin unterstützt und sich maßgeblich an der Internationalen Frauenuniversität (ifu) *Technik und Kultur* beteiligt. Insbesondere entstand in diesem Zusammenhang der von Christiane und anderen 2002 herausgegebene Sammelband *Feminist Challenges in the Information Age – Information as a Social Resource*, der als Band 5 in der Schriftenreihe der internationalen Frauenuniversität *Technik und Kultur* bei Leske & Budrich, Opladen erschienen ist.

Aus Christianes ifu-Einsatz hat sich ein nachhaltiges Engagement für universitäre Informatik in Addis Abeba (Äthiopien) ergeben, dem sie sich bis heute widmet. Näheres findet man in Christianes Artikel *Informations- und Kommunikationstechnologien für Entwicklung – am Beispiel von drei Projekten in Äthiopien* erschienen in *Frank Fuchs-Kittowski, Werner Kriesel (Hg.): Informatik und Gesellschaft. Festschrift zum 80. Geburtstag von Klaus Fuchs-Kittowski. Peter Lang Ltd. International Academic Publishers, Frankfurt a. M. 2016.*



Christiane Floyds Ideen trafen anfangs auf viel Unverständnis, sind heute im Fach weitgehend akzeptiert, aber in der Praxis keineswegs immer durchgesetzt. Für den ersten *Fiff-Rundbrief* 1/1984 hat Christiane ein zehnteitiges Positionspapier verfasst mit dem Titel *Wo sind die Grenzen des verantwortbaren Computereinsatzes?* Während die Gründung des Fiff ganz überwiegend auf der im Rahmen der Friedensbewegung geführten Diskussion um die unheilvolle Verflechtung von Rüstung und Informatik beruhte, hat Christiane weitsichtig den gesamtgesellschaftlichen Kontext im Auge gehabt und als Aufgabenbereich des Fiff identifiziert. Sie nennt drei Kategorien:

„Grenzen des fachlich verantwortbaren Computereinsatzes sehe ich dort, wo Computer aufgrund eines verfehlten Vertrauens in die Leistungsfähigkeit von Programmen eingesetzt werden, ...

Grenzen des zwischenmenschlich verantwortbaren Computereinsatzes sehe ich dort, wo Computer aufgrund einer verfehlten Gleichsetzung von Menschen mit Maschinen eingesetzt werden, ...

Moralisch/politische Grenzen des Computereinsatzes schließlich sind dort zu ziehen, wo (im Sinne Weizenbaums) mit Computern versucht wird, was ohne Computer nicht gemacht werden darf, ...“

In ihrem Grußwort zum dreißigjährigen Bestehen des Fiff beschreibt sie das Fiff als Kind des Kalten Krieges mit den Herausforderungen der Zeit, die sich nun aktualisiert als globalisierte Herausforderungen erweisen: Informatik und Rüstung, Informatik und die Arbeitswelt, Informatik und Datenschutz. Insbesondere schreibt sie:

„Als erste Fiff-Vorsitzende habe ich ein Positionspapier über Grenzen des verantwortbaren Computereinsatzes geschrieben. Diese Prinzipien waren einfach und zeitlos. Doch die Erarbeitung von konkreten Gesichtspunkten dazu erfordert tiefgreifendes Spezialwissen, um anspruchsvolle Technologien kritisch zu würdigen, die Fähigkeit zur interdisziplinären Arbeit und Kommunikation, das Aufzeigen von rechtlichen Rahmenbedingungen und politischen Durchsetzungsstrategien.“ (aus

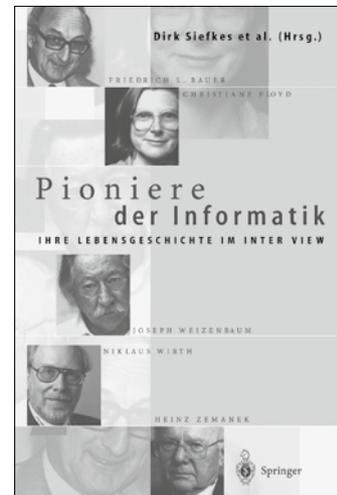
Grußwort in: Peter Bittner et al. (Hg.): *Gesellschaftliche Verantwortung in der digital vernetzten Welt*, Reihe: *Kritische Informatik*, LIT Verlag Münster 2014).

Wir konnten sie bei vielen Diskussionen und Tagungen zu Informatik und Gesellschaft erleben, wo sie ihre gewichtige Stimme erhoben hat. Erinnerung sind dabei beispielsweise intensive Diskussionen um die Ausarbeitung und Wirksamkeit von ethischen Anforderungen, wie sie auch Eingang in Lehrmaterialien fanden: *Christiane Floyd und Herbert Klaeren: Informatik: gestern, heute und morgen. Modul im Fernstudium Informatik und Gesellschaft (Erprobungsfassung)*, Universität Tübingen, 1998. Was die sozialen Anforderungen an die Modellierung betrifft, haben wir u. a. mit Fanny Michaela Reisin gefordert, dass Computer niemals Entscheidungen treffen dürfen, sie müssen dem Menschen überlassen bleiben. Nun sind leider, leider Entscheidungsunterstützung und verdeckte Entscheidungen innerhalb von Plattformlösungen ein Thema unserer jetzigen Tagung geworden.

Die Stimme des FIF ist zwar laut und wird immer lauter, aber immer noch in ihrer Reichweite zu beschränkt. Die viel größere Gesellschaft für Informatik unternimmt jetzt auch Anstrengungen, sich in Politik und Gesellschaft mehr Gehör zu verschaffen, und die Arbeitsgruppen zu Informatik und Gesellschaft und Informatik und Ethik bekommen innerhalb der GI mehr Aufmerksamkeit. Es ist höchste Zeit.

Es soll nicht unerwähnt bleiben, dass die ersten zwei Professorinnen für Informatik im deutschen Sprachraum Österreicherinnen sind oder waren, die dritte Professorin in Genf war Schweizerin und war von Hermann Maurer aus Graz betreut worden. In der Tat studierten in den 1960er-Jahren in Österreich zu etwa einem Drittel Frauen Mathematik, während zur gleichen Zeit an den westdeutschen Hochschulen Mathematikerinnen Einzelwe-

sen waren. In der DDR lag allerdings der Frauenanteil in der Mathematik bis 1988 bei ca. 50 % und darüber, 1989 fiel der Anteil auf 8 %. Aus diesem Fach geschah jedoch die größte Wanderung bzw. Ausgründung in die Informatik. Diese Generation von Mathematiker:innen und Ingenieur:innen hat das Fach begründet und definiert, nachdem mit der zunehmenden Bedeutung von Programmierung und Software in der Industrie eine gezieltere Ausbildung erforderlich geworden war.



Christiane Floyd hat ihr Fach im Sinne einer gesellschaftlichen Verantwortung für die Informatik aufgefasst, gestaltet und gelebt – eine außergewöhnliche Persönlichkeit und ein großes Vorbild. Zurecht wird sie neben vier „Pionieren der Informatik“ als einzige Pionierin gestellt: *Anette Braun, Dirk Siefkes, Klaus Städtler, Peter Eulenhöfer und Heike Stach (Hrsg.): Pioniere der Informatik: Ihre Lebensgeschichte im Interview*. Springer, Berlin Heidelberg, 1998.



Christiane Floyd

Ansprache zur Verleihung der Weizenbaum-Medaille 2021

Liebe FIFler und FIFlerinnen,

meinen ganz herzlichen Dank für die schöne Laudatio von Britta Schinzel, die mich, weit über ihren Auftrag hinausgehend, deutlich mehr als nötig gepriesen und ganz verlegen gemacht hat. Und danke vor allem für die Verleihung der Weizenbaum-Medaille. Dass mir das FIF diese Medaille verleiht, bedeutet für mich in zweifacher Hinsicht etwas sehr Persönliches.

Zum einen hat mich eine langjährige Freundschaft mit Joseph Weizenbaum verbunden. Wie ich an anderer Stelle schon deutlich gemacht habe¹, war unsere Freundschaft ungewöhnlich, weil ich ihn in seinen beiden Identitäten gekannt habe: zuerst in der amerikanischen, dann in der wiedergefundenen deutschen – und schließlich in der immer bewusster gelebten jüdischen. Er wiederum lernte mich zuerst als Professoren-Gattin in Stanford kennen und traf mich nicht allzu lange danach als Professorin an der TU Berlin wieder. Als er 1980 nach Jahrzehnten des Exils erstmals wieder längere Zeit in Berlin verbrachte und einen

intensiven Prozess der Suche nach seinen Wurzeln durchlebte, war ich an seiner Seite. Umgekehrt hatte er einen wichtigen Einfluss auf die Ausrichtung meiner Forschung. Denn in seinem Seminar über Transparente Systeme an der TU fanden Reinhard Keil und ich mit unserem gemeinsamen Anliegen zusammen, Verantwortung in der Informatik zu verwirklichen. Dass Weizenbaum bereits 1980 erkannte, wie wichtig das Thema Transparenz, das damals noch kaum jemanden interessierte, einmal werden würde, zeigt exemplarisch die Klarheit seiner Einsicht in die Gefahren einer bedenkenlosen Computerisierung.

Zum anderen ist meine Beziehung zum FIF fast vier Jahrzehnte lang. 1984 wurde ich nicht nur Gründungsmitglied, sondern erste Vorsitzende. Ich möchte die heutige Gelegenheit nutzen, um insbesondere den Jüngeren die Frühzeit des FIF in Erinnerung zu rufen, so wie sie sich mir aus meinem (West-) Berliner Blickwinkel darstellte. Wie allgemein bekannt, ist das FIF aus der Friedensbewegung der frühen 80er-Jahre entstanden, und zwar dezentral, in Gruppen, die über die ganze damalige Bun-

desrepublik verteilt und sowohl an Universitäten wie auch in der Praxis aktiv waren. Ich erinnere mich an Gruppen in München, Berlin, Darmstadt, Bonn, Dortmund, Hamburg, Bremen – und bitte alle um Entschuldigung, die ich jetzt nicht erwähnt habe. Besonders aktiv war die Gruppe in der Gesellschaft für Mathematik und Datenverarbeitung (GMD) in St. Augustin, die unter Federführung von Helga Genrich eine Schlüsselstellung bei der Gründung des FIFf innehatte.

Von alledem wusste ich zunächst nur wenig. Meine Verpflichtungen als deutschlandweit erste Professorin für Softwaretechnik, in Berlin allein lebend mit meinen beiden kleinen Kindern, machten es mir unmöglich, mich als Friedensaktivistin zu betätigen. Dagegen waren meine engsten Mitarbeiter:innen in der Forschungsgruppe Softwaretechnik, Michaela Reisin und Reinhard Keil, in der informatikbezogenen Friedensbewegung sehr engagiert. Eines Tages kamen sie zu mir und erzählten von der bevorstehenden Vereinsgründung. Ich sagte sofort, dass ich gern Mitglied werden würde. Gut, aber das reicht nicht, antworteten sie. Du wirst als Vorsitzende gebraucht. Ein ziemliches Risiko für mich, doch ließ ich mich darauf ein.

Die Gründungsversammlung 1984 fand bei der GMD in einem großen Hörsaal mit sehr hoher Beteiligung statt. Am Podium vorne stand Joseph Weizenbaum – damals Gastwissenschaftler bei der GMD – und hielt einen aufrüttelnden Impulsvortrag. Er wurde unter großem Applaus zum Ehrenvorsitzenden ernannt. Die anschließende Vorstandswahl verlief einfach und routinemäßig. Alle Kandidat:innen waren vorher angesprochen worden und ich wurde, wie geplant, zur Vorsitzenden gewählt. Ich kannte kaum jemanden – kaum jemand kannte mich. Auf der Heimfahrt begann ich zu überlegen, wie ich mich als Person in diesem Amt einbringen könnte.

Mein erster Beitrag war die Namensgebung. Ursprünglich sollte es ein „Forum InformatikerInnen für Frieden“ werden. Ich brachte die Erweiterung „und gesellschaftliche Verantwortung“ ins Gespräch. Mir war aus meiner Beschäftigung mit der Künstlichen Intelligenz und der Softwaretechnik schon sehr früh bewusst geworden, dass die gesellschaftliche Umsetzung der Informationstechnik eine Fülle von neuen Problemen aufwarf, auf die niemand vorbereitet sein konnte. Mein Vorschlag wurde wohl vor allem deshalb akzeptiert, weil Datenschutz im Zuge der Volkszählung 1983 zu einem hochbrisanten Thema geworden war. Das andere gesellschaftlich aktuelle Thema der 1980er-Jahre war Computer und Arbeit, ausgelöst durch den Einzug von interaktiven Systemen und Arbeitsplatzcomputern in die Praxis. Dieses Thema wurde interdisziplinär und in Zusammenarbeit mit gesellschaftlichen Akteuren wie den Gewerkschaften diskutiert, innerhalb der Informatik betraf es die Fachgebiete Informatik und Gesellschaft, Softwaretechnik und die neu gegründete Software-Ergonomie.

Das Kürzel FIFf, das sich in der Folge so bewährt hat, stammt von Peter Löhr, ebenfalls Mitglied im Gründungsvorstand. Er meinte, dass die Ähnlichkeit zu „Pfiff“ und zu „pfiffig“ ein positives Vorzeichen für die Arbeit des FIFf sein würde. Und so war es denn auch.

In meinem Berliner Büro angekommen, ging es mir vor allem darum, wie ich auf die Richtung der Arbeit im FIFf Einfluss nehmen



Christiane Floyd empfängt die Weizenbaum-Medaille von Rainer Rehak – Quelle: media.ccc.de, CC BY 4.0

könnte. So manches, was im Raum stand, wollte ich *nicht*. Ich wollte, dass das FIFf *in* der Informatik, *nicht gegen* die Informatik arbeitet. Ich wollte, dass sich das FIFf *keiner Ideologie* oder politischen Richtung verpflichtet fühlt. Und ich wollte *kein Forum für Anti-Amerikanismus* schaffen. Stattdessen wollte ich ein Umdenken in der Informatik² fördern.

Und so schrieb ich ein Positionspapier über „Grenzen des verantwortbaren Computereinsatzes“³, in dem ich drei Kategorien von Grenzen aufzeigte: Grenzen des *fachlich verantwortbaren* Computereinsatzes, wo Computer aufgrund eines verfehlten Vertrauens in die Leistungsfähigkeit von Programmen eingesetzt werden; Grenzen des *zwischenmenschlich verantwortbaren* Computereinsatzes, wo Computer aufgrund einer verfehlten Gleichsetzung von Menschen mit Maschinen eingesetzt werden; *moralisch/politische* Grenzen des Computereinsatzes, wo (im Sinne Weizenbaums) mit Computern versucht wird, was ohne Computer nicht gemacht werden darf.

Zu meiner Freude fand dieses kleine Papier eine große Resonanz, wurde viel diskutiert und schließlich als Stellungnahme im Informatik Spektrum der Gesellschaft für Informatik abgedruckt. Damit wurde es als legitime Position innerhalb der Informatik anerkannt.

In den 80er-Jahren war die Arbeit des FIFf vor allem durch den Kalten Krieg bestimmt, durch die beidseitige Aufrüstung und die berechnete Furcht vor einer nuklearen Auseinandersetzung auf deutsch-deutschem Boden, die beide Teile des heute wiedervereinigten Landes ernsthaft gefährdet hätte. Dieses Bedrohungsszenario war sehr real. Es galt also zu warnen, sich eine Stimme zu verschaffen, sich mit anderen Engagierten zu vernetzen und in der Öffentlichkeit Präsenz zu zeigen.

In Berlin gelang über die Mauer hinweg schon bald die Vernetzung zwischen Ost und West. Ich war befreundet mit Klaus Fuchs-Kittowski, damals Professor an der Humboldt-Universität. Seine Möglichkeiten waren zwar aus politischen Gründen eingeschränkt, er engagierte sich aber von Anfang an in überzeugender Weise und hat das seither lebenslang getan.

1985 veranstalteten mein Kollege Hartmut Ehrig und ich an der TU Berlin eine wissenschaftliche Konferenz TAPSOFT (Theory and Practice of Software Development), wo sich führende

Wissenschaftler:innen aus der Softwaretechnik und der Theoretischen Informatik trafen⁴. Im Rahmenprogramm konnte ich eine Podiumsdiskussion über Verantwortung platzieren, die vom FIF Berlin organisiert wurde, mit einem renommierten, internationalen Panel und circa 200 Mitdiskutierenden.

Mitten im Publikum saß David L. Parnas – hoch angesehen für seine wissenschaftlichen Leistungen, aber unter Kritik wegen seiner Beratungstätigkeit bei der US Navy. Er argumentierte mit Nachdruck, dass Wissenschaftler keine politische, sondern fachlich begründete Kritik äußern sollten. Das genügte den engagierten Berliner Intellektuellen nicht. Warum er nicht bei der Navy zurücktrete? Die Antwort von Parnas ist mir immer noch im Ohr: *I can only do that once.*

Niemand von uns ahnte, dass er schon dabei war, seinen Ausstieg vorzubereiten. Er war eingeladen worden, in leitender Stellung an der von Präsident Reagan initiierten *Strategic Defense Initiative* mitzuarbeiten, in der für die USA ein elektronischer Schutzschild zur Abwehr feindlicher Raketenangriffe entwickelt werden sollte. Anfliegende Raketen sollten automatisch erkannt und ohne Intervention von Menschen vernichtet werden. Es war ein Traum, ausgehend von einer Illusion von absoluter Sicherheit und einem irreführenden Glauben an die Leistungsfähigkeit von Wissenschaft und Technik. Mehrere Disziplinen sollten involviert sein – David Parnas als Vertreter des Software Engineering.

Nur wenige Wochen nach unserer Konferenz veröffentlichte er einen Offenen Brief an die US-Regierung, in dem er die Einladung höflich ablehnte und eine fachlich fundierte Begründung lieferte, warum aus informatischer Sicht das SDI-Projekt nicht gelingen konnte⁵. Zugleich trat er als Berater der Navy zurück.

Durch diese spektakuläre Aktion wurde David Parnas zum Vorbild, zum *role model*, für viele, die sich kritisch in der Informatik engagierten. Sein Offener Brief wurde mehrfach veröffentlicht und als Grundlage zahlreicher Lehrveranstaltungen verwendet. Begleitend zu seinem Vorgehen zog Parnas einschneidende persönliche Konsequenzen: Er gab seine amerikanische Staatsbürgerschaft auf, wurde Kanadier und lebte seither in Kanada, zeitweise in Irland. Vermutlich hätte er das nicht müssen. Seine amerikanischen Landsleute haben jedoch wenig Verständnis für seine Art von Kritik.

Bei der Vernetzung mit den Naturwissenschaftler:innen-Initiativen spielte das FIF Hamburg eine wichtige Rolle, das den Kontakt zu meinem hoch engagierten späteren Kollegen und Freund, dem Physiker Hartwig Spitzer herstellte. Unter seiner Regie fand 1986 ein eindrucksvoller Kongress an der Universität Hamburg statt, bei dem mehrere Nobelpreisträger sprachen, auch solche, die ursprünglich im Manhattan-Projekt bei der Entwicklung der ersten Atombombe beteiligt waren und später zu den prominenten Warnern vor dem Einsatz von Nuklearwaffen gehörten. Bei diesem Kongress leitete das FIF Hamburg eine Arbeitsgruppe.

1987 verbrachte ich mein Forschungssemester an der Universität Stanford und konnte die Vernetzung des FIF mit den Computer Professionals for Social Responsibility herstellen, einer im benachbarten Palo Alto zentrierten Initiative, in der die mit mir befreundeten KollegInnen Terry Winograd und Lucy Such-



Christiane Floyd bei ihrer Ansprache anlässlich der Verleihung der Weizenbaum-Medaille – Quelle: media.ccc.de, CC BY 4.0

man eine tragende Rolle spielten. Soweit ich weiß, ist es bei der freundschaftlichen Vernetzung geblieben, von einer inhaltlichen Zusammenarbeit weiß ich nichts.

Unmittelbar nach der Wende bereitete die Initiative Verantwortung der Naturwissenschaftler:innen in Zusammenarbeit mit Wissenschaftler:innen aus Ost und West einen eindrucksvollen Kongress *Challenges – Science and Peace in a Rapidly Changing Environment* an der TU Berlin vor, bei dem das FIF als Mitveranstalter auftrat, Wolfgang Hesse eine Arbeitsgruppe über Verantwortung leitete, und ich einen Vortrag über Wissenschaft und Ethik hielt⁶. Schon das Programmkomitee-Meeting in Leningrad – das gerade im Begriff war, wieder zu St. Petersburg zu werden – und der Austausch dort mit sowjetischen Kollegen und Kolleginnen in den Weißen Nächten 1991 (genau 50 Jahre nach dem Einmarsch der deutschen Wehrmacht) blieben unvergesslich. Der Kongress im September 1991 war dann gewissermaßen der Höhepunkt und zugleich der Abschluss der wissenschaftlichen Friedensarbeit im Kalten Krieg und gleichzeitig die bewusste Zuwendung zu neuen Herausforderungen.

Die ungeheuren Veränderungen, welche die damalige Wende auslöste, brachten auch für das FIF eine Re-Orientierung der Arbeit mit sich. Nun, da die akute Dringlichkeit der Rüstungsthemen allmählich verblasste, traten die zivilen Themen im Zuge der Weiterentwicklung der Informations- und Kommunikationstechnik umso mehr in den Vordergrund. Das FIF hat sich in den letzten Jahrzehnten professionalisiert und tritt mit fundierten Einschätzungen und Kritiken über aktuelle technologische Entwicklungen an die Öffentlichkeit.

Ich selbst habe mich aus der aktiven FIF-Arbeit im Zuge meines Wechsels an die Universität Hamburg mit Bedauern zurückgezogen. Umso mehr bin ich meinen früheren Berliner Kollegen Reinhard Keil und Hans-Jörg Kreowski dankbar, die in vorbildlicher Weise die Organisation, die Arbeit und die Außenwirkung des FIF weiter begleitet und geprägt haben.

Für mich war und blieb Verantwortung ein zentrales Anliegen meiner wissenschaftlichen Arbeit. Das war schon so bei der Entwicklung des Ansatzes STEPS (Softwaretechnik für evolutionäre, partizipative Systementwicklung) an der TU Berlin. Wir betrachteten Software im Kontext von Arbeits- und Kommunikations-

prozessen ihrer Benutzer und Benutzerinnen und stellten Menschen in den Mittelpunkt der Entwicklung und Nutzung von Technik. Partizipation der Beteiligten, Prototyping und versionsorientierte Entwicklung, Rückkopplung aus der Nutzung für Redesign und Weiterentwicklung – das alles waren damals kontroverse Ansätze, die in der formal-technischen Ausrichtung des Software Engineering Mainstreams der 80er-Jahre als unwissenschaftlich galten, mittlerweile aber weltweit anerkannt werden. Auch an der Uni Hamburg haben wir sie, unter Leitung von Ingrid Wetzel (verh. Schirmer) erfolgreich in Projekte umgesetzt.

Mein weiterer Lebensweg hat dazu geführt, dass ich mich bis heute gemäß denselben Prinzipien im Bereich ICT for Development in Äthiopien engagiere. In einem österreichisch-äthiopischen Projekt TEMACC (Technology Enabled Maternal and Child healthCare⁷) geht es darum, die Informations- und Kommunikationstechnologien zu nutzen, um Frauen und Kindern im ländlichen Raum einen besseren Zugang zum primären Gesundheitssystem zu ermöglichen. Information allein führt zwar noch nicht zu Gesundheit, sie hat jedoch eine Schlüsselstellung bei der Weiterbildung von Müttern und Gesundheitsarbeiter:innen, bei der Koordinierung der Arbeit und der Kommunikation zwischen gesundheitsrelevanten Organisationen, dem Aufbau einer kontinuierlichen Mutter-Kind-Betreuung und der Qualität von Diagnose und Therapie. Der Versuch, hochwertige ICT Dienstleistungen auf einfachster technologischer Basis in Partizipation von vorwiegend analphabetischen Müttern und nur gering ausgebildeten Gesundheitsmitarbeiter:innen bereitzustellen – in Gemeinschaften, die in jeder Hinsicht vom Mangel geprägt sind – ist fachlich und menschlich ungeheuer spannend. Auch halte ich die Entwicklungszusammenarbeit auf Augenhöhe, insbesondere mit afrikanischen Ländern, für eine der wichtigsten Aufgaben unserer Zeit.

Damit will ich schließen. Ich danke Ihnen für Ihre Aufmerksamkeit und freue mich sehr darüber, dass ich meine Freundschaft mit den FIFF bei diesem schönen Anlass erneuern konnte.

Nachsatz

Diese Ansprache habe ich im November 2021 gehalten und im Mai 2022 schriftlich rekonstruiert. Ich kann nicht schließen, ohne auf die ungeheuren Veränderungen seither Bezug zu nehmen. Wie selbstverständlich schien es den meisten von uns noch vor wenigen Monaten, dass Themen wie Krieg, Rüstung und militärische Bedrohung in Europa zur Vergangenheit gehören. Und nun sind sie auf brutale Weise in unsere Gegenwart eingedrungen.

Diese Umwälzung führt zu neuen Herausforderungen für alle, die sich für Frieden und Verantwortung engagieren, auch für das FIFF. Es tut gut zu wissen, dass das FIFF zusammen mit anderen frühzeitig Position bezogen hat. Im Anschluss an die erste Pressemitteilung hat sich per E-Mail eine heftige, generationsübergreifende Diskussion ergeben, aus der klar wurde, wie tief betroffen wir alle waren, wie manche Älteren versuchen, ihre Positionen aus der Vergangenheit weiter zu vertreten, und wie manche Jüngeren erst dabei sind, ihre eigene Position zu formulieren.

Aus meiner Sicht ist es sehr wichtig, diese Diskussion im FIFF weiter zu führen, und zwar unter wechselseitiger Achtung zwischen den Generationen. Hier ist ein Auszug aus einer E-Mail, die ich in diesem Zusammenhang geschrieben habe:

„ ... Damals waren wir jung. Wir hatten Ansichten und Überzeugungen, die damals aktuell waren. Damals gab es den Kalten Krieg. Es gab zwei einander feindlich gegenüberstehende Blöcke. Die Verhältnisse haben sich seit 1989 geändert. Die Welt ist polyzentrisch geworden und wir sind mit unseren Überzeugungen aus den 70er- und 80er-Jahren gealtert. Älter geworden, haben wir uns längst aus der aktiven Mitarbeit zurückgezogen.

Im November habe ich in München den Weizenbaum-Preis des FIFF erhalten. Bei dieser Gelegenheit – es war immerhin die Jahrestagung des FIFF – war niemand aus meiner Generation anwesend. Dafür hatte ich die Gelegenheit, viele junge Leute kennen zu lernen. Die ticken anders als wir damals.

Ich meine, dass es uns, als Altgedienten, gut zu Gesicht steht, die Jungen im FIFF zu unterstützen, aber ihnen die Verantwortung dafür zu überlassen, welche politische Position, die für sie heute aktuell ist, sie vertreten. Danke an Rainer Rehak und an alle anderen für Euer Engagement.“

Anmerkungen

- 1 C. Floyd: Die Bedeutung Joseph Weizenbaums für die deutsche Diskussion um Informatik und Gesellschaft. In: H. Oberquelle (Hrsg.): Ehrendoktor Joseph Weizenbaum, Dokumentation des Festkolloquiums, Mitteilung 322, Universität Hamburg, Fachbereich Informatik, 2003, S. 29-32
- 2 Das war auch der Titel der zweiten FIFF-Jahrestagung, die 1986 vom FIFF Berlin an der TU organisiert wurde.
- 3 C. Floyd: Wo sind die Grenzen des verantwortbaren Computereinsatzes? In: Informatik-Spektrum, Band 8, Heft 1, 1985, S. 3-6
- 4 H. Ehrig, C. Floyd, M. Nivat, J. Thatcher (Hrsg.): Formal Methods and Software Development. Proceedings of the International Joint Conference on Theory and Practice of Software Development (TAPSOFT), Vol. 1: Colloquium on Trees in Algebra and Programming, Vol. 2: Colloquium on Software Engineering, Springer Verlag, Berlin, Heidelberg 1985
- 5 David L. Parnas: Software Aspects of Strategic Defense Systems. Communications of the ACM 28, S.1326-1335
- 6 C. Floyd: Science and Ethics. In: R. Rilling, H. Spitzer, O. Greene, F. Hucho, G. Pati (Hrsg.): Challenges – Science and Peace in a Rapidly Changing Environment, Band 1, BdWi, Marburg 1992, S. 172-189
- 7 R. Bekele, T. Biru, J. Sametinger, I. Groher, G. Pomberger, C. Floyd: Adapting Ethnography for Design Research: Lessons Learnt from Design of Mobile Systems for Rural Health Care in Ethiopia. In: Proceedings of the International Conference on Information Systems, München 2019 R. Bekele, I. Groher, J. Sametinger, T. Biru, C. Floyd, G. Pomberger, P. Oppelt: User-Centered Design in Developing Countries: A Case Study of a Sustainable Intercultural Healthcare Platform in Ethiopia. In: Proceedings of the IEEE/ACM Symposium on Software Engineering in Africa (SeiA) 2019



Ankündigung der FiffKon 2022 mit der Mitgliederversammlung

FiffKon 2022 – make install peace

Impulse für den Frieden

21. – 23. Oktober 2022 – Archenholdsternwarte Berlin Treptow

Unser Ziel ist nichts weniger als Weltfrieden. Dieser Wunsch klingt so phantastisch und utopisch, weil die Vorstellung einer friedlichen Welt für viele Menschen in ungreifbarer Ferne steht.

Umso dringlicher stellt sich die Frage, wie ausgehend vom Hier und Jetzt der Weg hin zu einer friedlichen Welt gestaltet und gegangen werden kann. Welche Maßnahmen befördern dabei kurz-, mittel- und langfristig friedliche Gesellschaften?

Als Forum kritischer Informatiker:innen, das den Begriff „Frieden“ seit 1984 stolz im Namen trägt, wollen wir uns mit der FiffKon22 dieser Fragen annehmen. Dabei verstehen wir Frieden nicht lediglich als die Abwesenheit von Krieg und als bloßes Fehlen offener Gewalt. Unser Leitbegriff ist der positive Frieden. Bei diesem wirken gerechte Institutionen, eine faire Verteilung von Ressourcen und eine respektvolle Haltung zwischen Menschen und auch Gruppen zusammen. Er ist die Basis dauerhaft friedlicher Gesellschaften, die tatsächlich menschlichen Fortschritt mit sich bringen. Friedensarbeit ist dabei immer langfristig wirksam. Unser Ziel ist ein pluralistisches, ausdifferenziertes Zusammenleben, das auf die Bedürfnisse der Menschen eingeht, um ein zufriedenes, selbstbestimmtes und komfortables Leben aller zu ermöglichen. Dazu wollen wir mit der Konferenz einen wissenschaftlich fundierten Beitrag leisten.

Mit Speaker:innen aus verschiedenen Fachgebieten beleuchten wir, welche Rolle Informationstechnik bei Friedensfragen spielt und welche Verantwortung der ihr zugrundeliegenden Wissenschaft – der Informatik – zukommt. Wir betrachten dabei Technologien in konkreten gesellschaftlichen Kontexten, um daraus sowohl erforderliche Änderungen an strukturellen Rahmenbedingungen als auch Handlungsmöglichkeiten abzuleiten. Dabei werden wir vom lebensdienlichen – konvivialen – Gedanken

getragen, der Umsetzungsstrategien innerhalb der begrenzten Ressourcen des Planeten einfordert.

Wir stellen die Konferenz unter zwei Leitfragen:

1. Wie kann Technologie gestaltet werden, um Frieden zu fördern?

Zunächst wollen wir die gesellschaftlichen, politischen, rechtlichen und ethischen Dimensionen beleuchten, um die Rahmenbedingungen und Voraussetzungen für positiven Frieden im Grundsatz zu verstehen:

- Welche Verantwortung kommt der Politik zu, welche dem Individuum, welche dem Bildungswesen und welche einer aktiven Zivilgesellschaft?
- Wie können Interessen von Menschen und Organisationen erfasst werden, so dass angemessen auf Situationen und Veränderungen reagiert werden kann?
- Wie sieht ein Rechtssystem aus, das vor allem die Ursachen und Hintergründe von friedensgefährdenden Vorfällen verstehen und positiv beeinflussen will?
- Wie müssen logistische und organisatorische Abläufe gestaltet werden, sodass sie auch absichtlichen Störungen widerstehen?

Aus diesen und weiteren Fragestellungen heraus wollen wir ableiten, was sich im Konkreten für die Gestaltung soziotechnischer Systemen und den Einsatz von Technologien lernen lässt:

- Welche der Herausforderungen können mit technischen Mitteln angegangen werden?
- Welche digitalen Werkzeuge und Technologien fördern individuelle Einbeziehung, Empowerment, Emanzipation, Transparenz und soziale Vernetzung sowie gesellschaftliche Resilienz?
- Mit welchen Methoden kann man – sich in der Regel lange im Vorfeld abzeichnende – Konflikte erkennen und kriegsbegünstigenden Entwicklungen entgegen wirken?
- Welche IT- und Medienkompetenzen müssen Einzug in Mentalität und Kultur erhalten, um den Frieden auch in den Köpfen zu verankern?
- Welche Beziehungen bestehen zwischen politischer Macht und Technikgestaltung, Friedenserhalt und Interessenskonflikten?
- Welche Rolle spielen politische, militärische und wirtschaftliche Machtkonzentrationen?
- Wie kann Technologie eingeschränkt werden, die überbordende Überwachung und Kontrolle ermöglicht?
- Wie sieht eine digitale Abrüstung aus? Wie kann sie durchgesetzt werden?
- Können Softwarelizenzen die militärische Nutzung ziviler (Software-)Entwicklung ausschließen?
- Welche Rolle spielen Kryptographie, Sicherheitslücken, Open-Source-Software, Lizenzmodelle und Systemarchitekturen?
- Welche Schäden hat die bisherige friedensblinde Gestaltung von Informationssystemen verursacht und wie können wir dies in Zukunft vermeiden?

2. Wie kann Technologie eingeschränkt werden, die Frieden stört?

Bei einer Umsetzung der Vision vom Frieden dürfen wir die bestehende Situation nicht außer Acht lassen. Wir müssen verstehen, welche Vorgänge und Entwicklungen den Weg hin zur Friedensutopie verhindern und was diesen entgegengestellt werden kann:

Mit der FifFKon entsteht in Berlin für drei Tage ein inspirierender Ort des Austauschs mit Vorträgen, Workshops und lebendiger Atmosphäre für gemeinsames Lernen, Diskutieren und Kontakt – mit Kulturprogramm, Lounges, Getränken und Musik bis in die Nacht.

Lasst uns tatsächlich mutig sein und mit dieser Konferenz dazu beitragen, eine neue, bessere Welt für alle Menschen zu erdenken, um sie dann zu bauen.

Einladung zur Mitgliederversammlung 2022

des Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung (FifF e. V.)

Wir laden fristgerecht und satzungsgemäß zur ordentlichen Mitgliederversammlung 2022 ein.

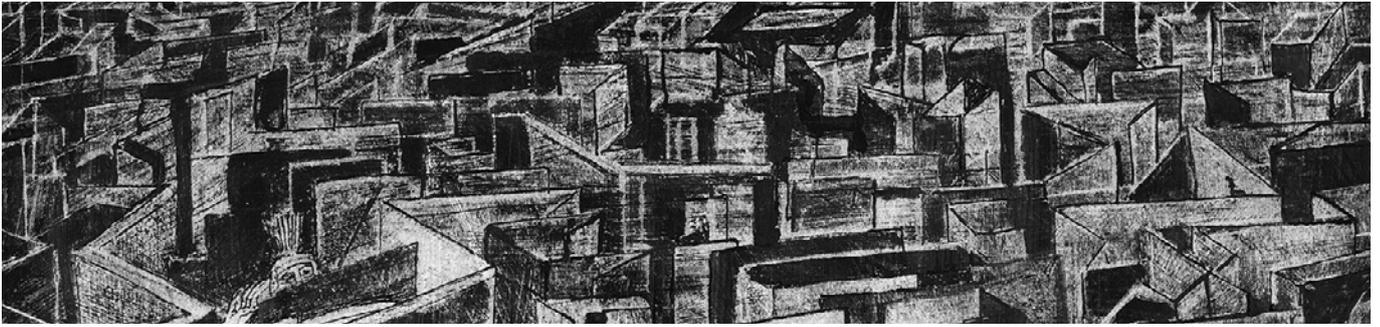
Sie findet am Sonntag, dem 23. Oktober 2022, von 12:00 bis 15:00 Uhr statt.

Adresse: Archenholdsternwarte Berlin Treptow

Vorläufige Tagesordnung

1. Begrüßung, Feststellung der Beschlussfähigkeit und Festlegung der Protokollführung
2. Beschlussfassung über die Tagesordnung, Geschäftsordnung und Wahlordnung
3. Bericht des Vorstands einschließlich Kassenbericht
4. Bericht der Kassenprüfer
5. Diskussion der Berichte
6. Entlastung des Vorstands
7. Neuwahl der Kassenprüfer
8. Diskussion über Ziele und Arbeit des FifF, aktuelle Themen, Verabschiedung von Stellungnahmen, Berichte aus den Regionalgruppen
9. Anträge an die Mitgliederversammlung
Anträge müssen schriftlich bis drei Wochen vor der Mitgliederversammlung bei der FifF-Geschäftsstelle eingegangen sein
10. Verschiedenes
11. Genehmigung des Protokolls

gez. Stefan Hügel
für den Vorstand und die Geschäftsstelle des FifF



Hans-Jörg Kreowski und Stefan Hügel

Künstliche Intelligenz: Zwischen hochgeschraubten Erwartungen und unkalkulierbaren Risiken

Editorial zum Schwerpunkt

Labyrinth – wie das auf der Titelseite – sind von alters her Sinnbilder menschlicher Intelligenz, die es bis in die griechische Mythologie geschafft haben. Perseus hat den Minotaurus, ein menschenfressendes Ungeheuer, in seinem Labyrinth auf Kreta erschlagen und konnte dann mit Hilfe des Ariadnefadens den Ausgang finden. Wir verwenden heute stattdessen Such-, Navigations- und Kürzeste-Wege-Algorithmen. Die Bibliotheken der Welt sind voller solcher Zeugnisse menschlicher Intelligenz, wie technische Artefakte die intelligente Gestaltung unserer Lebenswelt belegen. Es war deshalb ein werbetechnischer Geniestreich (und eine ungeheure Anmaßung), den Begriff „Intelligenz“ für die Benennung eines neuen Wissenschaftszweigs zu usurpieren. Zehn junge Wissenschaftler trafen sich für ein paar Wochen bei der Dartmouth Conference und begründeten die „Künstliche Intelligenz“. Im Förderantrag heißt es:

„The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.“

Das ist schon ganz dicht dran an dem, was auch heute unter Künstlicher Intelligenz (KI) verstanden wird. Und auch die damals angedachten Methoden wie regelbasierte Systeme und neuronale Netze sind bis heute zentral. Anfangs hat die Idee eines *General Problem Solver* nicht sehr weit geführt, erfolgreicher waren schon die anwendungsspezifischen Varianten der Expertensysteme. Einen gewaltigen Aufschwung hat die KI durch das japanische 5th-Generation-Programm und die *Strategic Computing Initiative* des Department of Defense der USA in den 1980er-Jahren genommen, um dann erst einmal wegen scheinbarer Erfolglosigkeit in einen KI-Winter zu versinken.

In den letzten 20 Jahren ist KI nach und nach wieder groß herausgekommen durch eindrucksvolle Erfolge bei Schach, Go und anderen Spielen sowie bei der Bild- und Sprachverarbeitung. Einige Zeit lang konnte man fast täglich in den Printmedien, dem Radio oder dem Fernsehen etwas über KI lesen, hören und sehen. Der große Hype um die KI scheint zumindest medial schon

wieder etwas abzunehmen. In Politik und Wirtschaft sind die Erwartungen aber weiterhin groß, wie die nationalen Strategien für Künstliche Intelligenz vieler Regierungen in aller Welt zeigen. KI wird als Schlüsseltechnologie zukünftiger Wertschöpfung auserkoren und mit gigantischen Finanzmitteln gefördert. Selbst eine geostrategische Bedeutung wird darin gesehen, bei KI die Nase vorn zu haben. Interessanterweise wird in den Strategiekonzepten auch häufig auf ethische Aspekte hingewiesen. KI soll zum Wohle der Menschen genutzt und problematische Anwendungen sollten vermieden werden. Tatsächlich gibt es auch viele warnende Stimmen. Die 26 Autor:innen des 2018 erschienenen 100-seitigen Reports *The Malicious Use of Artificial Intelligence: Forecasting, Prevention and Mitigation* beispielsweise warnen vor den Gefahren für die digitale, physische und politische Sicherheit durch die ungebremste Ausweitung von KI und maschinellem Lernen.

Die Folgen können gravierend sein. In den Niederlanden führten fehlerhafte Risikoindikatoren – beispielsweise der Besitz einer doppelten Staatsbürgerschaft – in einer Software, die Betrug beim Bezug von Kindergeld aufdecken sollte, zu massiven Falschbewertungen. In der Folge erhielten Kindergeldbezieher:innen Rückforderungen in teilweise sechsstelliger Höhe, gerieten dadurch in Armut; einzelne begingen Suizid. Die niederländische Regierung musste zeitweise zurücktreten (führte ihr Amt 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.“¹

Der Schwerpunkt dieser Ausgabe der *FfF-Kommunikation* bewegt sich in dem skizzierten Spannungsfeld. Auf unseren Aufruf hin sind sieben Beiträge eingegangen, die einige interessante Schlaglichter auf das Themenspektrum werfen. Die Beiträge sind voneinander unabhängig, so dass die gewählte Reihenfolge nicht signifikant ist.

Andrea Reichenberger und Martin Ziegler gehen davon aus, dass die gesellschaftliche Wahrnehmung der KI ambivalent zwischen Stein der Weisen und Büchse der Pandora schwankt. Ihr Diskurs versucht KI rund um diese beiden Pole zeitlich bis in die Zukunft, technisch versus gesellschaftlich und fundiert versus spekulativ auszuleuchten. Sie plädieren für eine weise Regulierung.

Britta Schinzel hat den Sammelband *Künstliche Intelligenz – Die große Verheißung* rezensiert, dessen vier Herausgeber:innen sich aus dem Berliner philosophischen Arbeitskreis MoMo Berlin und der DenkWerkstatt Berlin kennen. Dementsprechend sind die meisten der 26 Beiträge aus philosophischer Sicht verfasst, während die technische Seite der KI vielleicht etwas zu kurz kommt.

Peter Brödner geht auf die KI-Narrative ein, die sich im Laufe von 50 Jahren gewandelt haben, aber ein gemeinsames Charakteristikum besitzen: Sie produzieren Illusionen, wobei das sogenannte maschinelle Lernen in großen Datenbeständen neue Probleme mit sich bringt, was die Validität der Daten und die potenzierte Undurchschaubarkeit der Systeme betrifft.

Rainer Rehaks Beitrag ist ein Nachdruck aus Peter Verdegems Sammelband *AI for Everyone?*, der im vorigen Jahr bei University of Westminster Press erschienen ist, und deshalb von der Länge und Sprache her aus dem Rahmen fällt. Inhaltlich setzt sich der Autor mit dem Problem der irreführenden Begriffsbildung in der KI auseinander, wobei er seine Kritik durchaus konstruktiv wendet.

Claudia Müller-Birn setzt sich in ihrem Beitrag mit dem neuen Fachgebiet Data Science auseinander, in dem es um die Analyse großer Datenmengen geht und das mit KI große Überschneidungen aufweist. Dabei kommt es ihr auf die Verbindung mit qualitativen Methoden an. Sie diskutiert, wie eine solche kritisch-reflexive menschenzentrierte Datenpraxis in der akademischen Ausbildung verankert werden kann.

Klaus Heß unternimmt eine große Expedition durch die Gegenwart der KI, die in vielfältiger Weise längst unseren Alltag erreicht hat, und eine von einigen Protagonisten prognostizierten Zukunft ungeahnter technologischer Entfaltung. Der Microsoft-Chef hält KI für eine wirklich grundlegende Technologie und der Google-Chef findet sie bedeutender als Elektrizität und Feuer. Jürgen Schmidhuber versteigt sich gar zu der Annahme, dass KI die Milchstraße erobern wird.

Dorothea Kolossa hat eine Rezension über Stuart Russells Monographie *Human Compatible* verfasst. Der Autor beschwört ein existenzielles Risiko herauf, das von den rasanten Entwicklungen in der KI ausgeht, um dann Grundsätze herauszuarbeiten, wie Menschen die Kontrolle behalten können, selbst wenn die Maschinen eines Tages superintelligent werden sollten.

Hans-Jörg Kreowski konfrontiert den aktuellen Stand der KI als technisch-wissenschaftliche Disziplin einerseits und in ihrer Rezeption in den Medien, der Politik und der Wirtschaft andererseits mit den globalen Herausforderungen der heutigen Zeit und der Idee einer lebenswerten Welt für alle Menschen.

Wir hoffen, dass die Artikel auf großes Interesse stoßen und für alle Leserinnen und Leser neue Gesichtspunkte und erhellende Einsichten bieten.

Anmerkungen

- 1 Melissa Heikkilä (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/>



Andrea Reichenberger und Martin Ziegler

KI aus interdisziplinärer Perspektive

Der gesellschaftliche Ruf von KI-Technologien war und ist ambivalent. Er bewegt sich seit jeher zwischen Enthusiasmus und Enttäuschung, Glorifizierung und Diskreditierung. Mehr denn je reicht das Meinungsbild vom informationsverarbeitenden Stein der Weisen bis zum Sargnagel der Menschheit; gepriesen als Allheilmittel für Wirtschaft und Wissenschaft, verdammt als Pandorabüchse zur Versklavung der Menschheit; als willfähriger Assistent (Siri/Echo, Tesla Autopilot, Gecko Carebot) ebenso wie als Mastermind der kommenden Roboterrevolution (Boston Dynamics' BigDog/Atlas).

Dieser Essay versucht, den Diskurs zwischen den und jenseits dieser Pole zu erweitern: in sechs Abschnitten, von historisch bis zukünftig, von technisch bis gesellschaftlich, von fundiert bis spekulativ. Weder blauäugige Promotion noch blinde Panik propagierend, gehen wir aus von der Erfahrung (bspw. aus Kernenergie), dass technische Möglichkeiten sich in der Regel nicht dauerhaft verhindern¹ lassen – und stattdessen in ihren potenziellen Konsequenzen erst frühzeitig zu erkennen und dann weise zu regulieren sind.

Der erste Abschnitt rekapituliert und problematisiert gängige Vorstellungen und Verwendungen des Begriffs KI. Der zweite Abschnitt gibt einen kurzen Abriss der bisherigen und erwarteten Entwicklung von KI aus technologischer Sicht. Aktuelle und bevorstehende Regulierungsbemühungen fasst der dritte Abschnitt zusammen. Der vierte Abschnitt verschiebt den Fokus auf absehbare Auswirkungen im gesellschaftlichen und ökonomischen Bereich. Der fünfte schlägt unkonventionelle Blickwinkel auf KI vor jenseits menschenzentrierter Werkzeuge; und der letzte Abschnitt spekuliert über potenzielle langfristige soziale Konsequenzen.

Was ist (Künstliche) Intelligenz?

Seit den 1960er-Jahren hat das gesellschaftliche Ansehen von KI mehrfach gewechselt zwischen Enthusiasmus und Ernüchterung und steht aktuell mal wieder hoch im Kurs. Aber was ist KI?

Die Bund-Länder-Initiative zur Förderung der Künstlichen Intelligenz in der Hochschulbildung versteht darunter „Technologien zur maschinellen und algorithmen-basierten Verarbeitung großer Datenmengen“ (Bundesanzeiger vom 24.02.2021) – und übergeht dabei *zwei* zentrale Paradigmenwechsel: Training statt Programmierung und Heuristiken statt Algorithmen. In der Tat basiert klassische Datenverarbeitung auf den Antipoden Hardware/Software (vgl. Harvard/von Neumann-Architektur). Diese Unterscheidung spiegelt sich wider beispielsweise in den Fachbereichen Elektrotechnik/Informatik oder in Firmen wie Intel/Microsoft. Maschinelles Lernen fügt dieser Unterscheidung eine neue dritte Komponente hinzu: *Training*. Und weil ein solches Training naturgemäß nie vollständig abgeschlossen ist, sind durch KI generierte Lösungen inhärent *heuristisch* – im Gegensatz zu jenen von wohlspezifizierten und analysierten garantiert korrekten Algorithmen. Warnungen vor einer *Algokratie* gelten daher doppelt für eine *Heuristokratie*.²

Solche Verwirrungen und divergierende Einschätzungen mögen auch daher rühren, dass die technischen Hintergründe und genauen Funktionsweisen nicht nur Laien kaum verständlich sind: Der Online-Cartoon (Abbildung 1) bringt dies auf den Punkt, und aus gutem Grund bildet XAI (Explainable Artificial Intelligence) ein eigenes Forschungsgebiet innerhalb der KI.

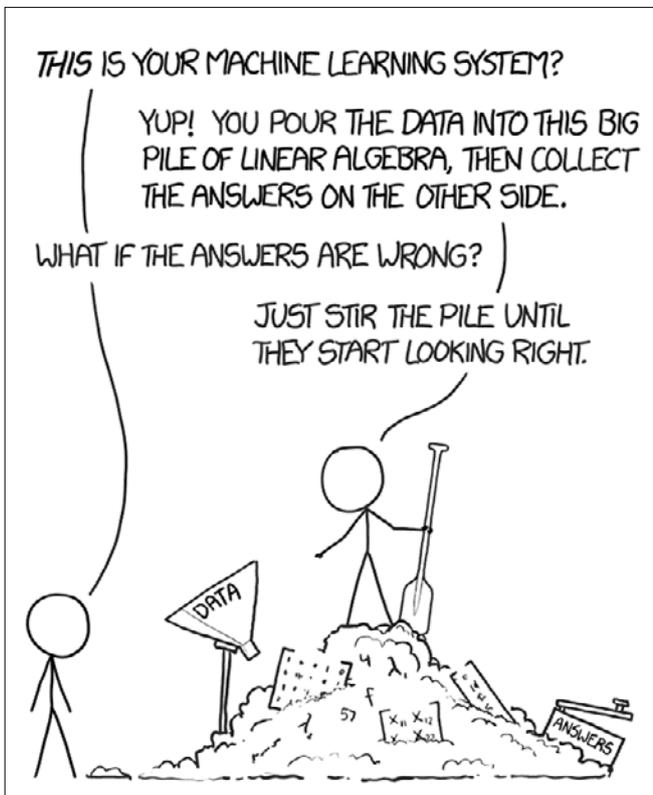


Abbildung 1. Machine Learning, Quelle: <https://xkcd.com/1838>, CC BY-NC 2.5

So sind Debatten oft spürbar von Science Fiction beeinflusst, sowohl hinsichtlich utopischer Visionen (wie *Der Zweihundertjährige* von Asimov oder *Commander Data* in *Star Trek*) wie hinsichtlich dystopischer Warnungen (*Blade Runner*, *Terminator*). Und gängige Vorstellungen von KI erweisen sich als anthropozentrisch einseitig dominiert von *menschlicher* Intelligenz und deren Imitation – obgleich selbst die bloße Definition menschlicher Intelligenz bekanntlich Gegenstand zahlreicher Kontroversen ist, beispielsweise in Soziologie und Erziehungswissenschaft.

Zusammengefasst stellt der Begriff *KI* also eher eine Hülse dar für verschiedenste Projektionen denn ein tatsächliches Fachwort. Dies hat geschichtliche Gründe:

Theorie und Technik

Alan Turing³ und Konrad Zuse⁴ gelten nicht nur als Mitbegründer der Theoretischen Informatik, sondern auch der KI.⁵ 1951 implementierte der Princeton-Student Marvin Minsky mit SNARC (*Stochastic Neural Analog Reinforcement Calculator*) das erste neuronale Netz, als Gründer und langjähriges Aushängeschild der KI. Bis zu seinem 1986er Buch *Society of Mind*⁶ stand KI noch der Psychologie und Physiologie nahe.

Seitdem ist der Zugang mathematisch-abstrakter⁷: Dieser Siegeszug der *Neats* über die *Scruffies* bildet die qualitative Grundlage der aktuellen Erfolge – zeigte Wirkung aber erst nach dem sogenannten *zweiten KI Winter* ab ca. 1993: Vorher reichte die Rechenleistung nicht aus, um Ideen wie *Reinforcement Learning* über einen qualitativen Proof-of-Concept hinaus in praxisrelevanter Größenordnung zu realisieren.

Der derzeitige Höhenflug von KI ist nämlich primär ingenieurtechnischen und quantitativen Fortschritten geschuldet: *Deep Blue* konnte zum Sieg über Garry Kasparov (1997) doppelt so viele Spielzüge pro Sekunde analysieren wie bei seinem verlorenen Vorgängerspiel (1996); und Moores Gesetz exponentiell wachsender Rechenleistung setzte sich fort über *Big Data* bis zum Sieg (2016) von *AlphaGo* über Lee Sedol: Eine Materialschlacht basierend auf 1202 CPUs und 176 GPUs und 48 TPU ASICs sowie mit einer Datenbank von 30 Millionen Zügen im Hintergrund [doi:10.1038/nature16961], *ad-hoc* zusammengestellt für dieses PR-Ziel.

Aktuell sehen wir erste Schritte hin zu Quasi-Standardisierung und zu schlüsselfertiger *AI as a Service* (AIaaS): Wie IBM PCs sich in den 80er-Jahren zu einem Hardware Quasi-Standard entwickelten (und dadurch mit konkurrierenden, aber kompatiblen *clones* einen Massenmarkt für kommerzielle Software-Entwicklung erst schufen), so wurde aus *Google Brains* proprietären *DistBelief* das plattformunabhängige *TensorFlow* mit Apache-Lizenz.

Ausgehend von diesen erst programmatischen über theoretischen und dann technischen Revolutionen von KI stehen wir heute vor einer gesellschaftlichen und einer administrativen Revolution:

Klassifizierung und Regulierung

Den ersten historischen Automobilen und Straßenbahnen mit nur wenigen PS musste noch ein *Fahrerlaubnis* warnend vorangehen. Inzwischen gilt in Deutschland StVO und TÜV mit zahlreichen Fahrzeug-Klassen und Führerscheinen: Um die Gefährlichkeit der technisch gut und gerne 100-mal mächtiger gewordenen Maschinen im Zaum zu halten, wird deren Zulassung, Versicherungspflicht und Spezifikation detailliert geregelt – und damit auch rechtlich die Verantwortung im Fall eines Produktionsfehlers oder Unfalls.

Eine ähnliche Entwicklung beginnt derzeit bei Autonomen Fahrzeugen mit der Norm SAE J3016. Sie formalisiert Autonomie in Stufen von 0 (Notfallassistenz) bis 5 (vollständig selbständig). Weitergehende Regulierungsfragen, wie beispielsweise zur Steuerung oder zur Verantwortung im Falle eines Unfalls⁸, scheinen bislang *ad-hoc* entschieden – aber verdienen angesichts drastisch wachsender Zulassungszahlen⁹ eine grundsätzliche Erörterung und konsistente Behandlung. Auch Produkthaftungsfragen stellen sich verstärkt und neu mit zunehmender Entwicklung und Verbreitung von Intelligenten Agenten und autonomen Servicerobotern. Solche und weitere Klärungspläne verfolgt die *Normierungs-Roadmap KI* der DKE, naturgemäß mit Fokus auf die technisch-rechtliche Seite.

Insgesamt wird der TÜV eine signifikante Aufgaben- und Kompetenzerweiterung erfahren. Aber zuerst muss die Informatik ihre klassischen Begriffe und Konzepte erweitern auf KI sowie Methoden entwickeln, um die Korrektheit von Hardware und Software *und Training* überhaupt erst zu formulieren (Spezifikation!) – und dann sicherzustellen bzw. zu verifizieren und zertifizieren! Weiterhin wichtig werden soziale und ökonomische Aspekte.

Anwendungen und Konsequenzen

Die gängigen aktuellen und angestrebten Anwendungen von KI übernehmen immer mehr Aufgaben, die früher von Menschen erledigt wurden. Das gilt für nahezu alle Bereiche der Arbeitswelt, vom Pflegebereich über Gastronomie bis hin zum Bildungs- und Dienstleistungssektor. Der damit einhergehende soziale und gesellschaftliche Wandel in der Wirtschafts- und Arbeitswelt, in Öffentlichkeit und Privatleben stellt uns vor Herausforderungen, Chancen und Risiken, die in ihrer Breite und Tiefe kaum absehbar sind.

Wie die Industrielle Revolution des 19. Jahrhunderts entlastet diese neue Technologie einerseits von unangenehmen Aufgaben, ersetzt andererseits traditionelle Berufe: nun auch im Tertiären Sektor – und hier nicht mehr nur solche im Bereich Niedriglohn/Geringqualifikation.

Damit forciert KI die Trennung zwischen Arbeit zu Erwerbszwecken und Arbeit als sinnstiftender Herausforderung; und KI macht das bedingungslose Grundeinkommen (BGE) nicht nur ökonomisch möglich durch Produktivitätssteigerung, sondern macht BGE langfristig sozial unvermeidbar durch den massiven Abbau von Arbeitsplätzen in weiten Wirtschaftsbereichen.

Diese Vierte Industrielle Revolution treibt somit Karl Marx' Konzept der Entfremdung auf die Spitze – und *ad absurdum*, indem sie der Gesellschaft enorme Mengen an Freizeit und historisch nie gekannte Freiheiten an die Hand gibt. Solche Freiheiten sinnvoll und konstruktiv zu nutzen und zu verteilen, wird sich zunehmend als Herausforderung herausstellen – beispielsweise in Form zunehmender Zahlen von NEETs und von modernen Ermiten (Hikikomori).

(Feministische) Science and Technology Studies (STS) haben seit den 1970er-Jahren diese Sichtweise auf Arbeit und auf die Arbeitswelt einer umfassenden und begründenden Kritik unterzogen und die Machtstrukturen einer solchen westlich zentrierten, patriarchalisch dominierten Digitalen Revolution untersucht, die den globalen Blick auf Ausbeutung, modernen Sklavenhandel, Kinder- und Frauenarbeit marginalisiert und verdeckt.

Horizonte und Herausforderungen

Klassische KI kreist um den Menschen, soll ihn einerseits simulieren/ersetzen (Turing) und ihm andererseits dienen: Humane Kognition und Nutzbarkeit wird implizit und arrogant als universelles Maß postuliert. Dabei bietet gerade KI die Chance, solche Hybris zu hinterfragen: Würde eine wahre *Krone der Schöpfung* wirklich so mit besagter Schöpfung umgehen oder wäre eine Superintelligenz nicht die bessere Herrscher:in? Sind Menschenaffen, Delphine, Kraken wirklich *weniger* intelligent – oder eher *anders* intelligent¹⁰, vergleichbar beispielsweise inselbegabten *Savants* oder Autist:innen?

Umgekehrt ermöglicht KI, kognitive Defizite zu kompensieren¹¹ (z. B. Grammarly für Menschen mit Dyslexie) und vorhandene Fähigkeiten zu komplementieren, mangelnde Fremdsprachenkenntnisse beispielsweise schon jetzt durch Echtzeitübersetzung im Smartphone.

Zukünftig wird KI so auch unterstützen bei der Kommunikation mit Tieren¹² und – nach Zwergschimpanse Kanzi – eine neue Generation von bahnbrechenden Experimenten ermöglichen: vielleicht zur Autobiografie eines Elefanten, zu religiösen/ethischen Vorstellungen eines Kraken oder zu animaler Autonomie im Sinne des *Great Ape Project*.

Wenn ein virtueller Assistent im Laufe der Zeit nicht nur auf die Aussprache der „Meisterin“ trainiert wird, sondern auch schrittweise deren Präferenzen lernt (beispielsweise in Online-Suchen oder -Reservierungen), entsteht dann nicht eine individuelle Persönlichkeit? Müssten wir dann nicht Fragen nach den intrinsischen Rechten einer KI-Entität stellen? Die daraus entstehenden Dilemmata führte bereits der Film *Blade Runner* (1982) deutlich vor Augen.

Utopie und Dystopie

„*L'enfer c'est les autres*“, erkannte Jean-Paul Sartre; und umgekehrt kann das geeignete Umfeld auch der Himmel auf Erden sein. Ein solches Umfeld lässt sich in Virtual Reality (VR) künstlich herstellen mit einem individuell *maßgeschneiderten*

MMORPG, wo Träume wahr werden (vgl. *Solaris* von Stanislaw Lem), während die Erinnerung an ein externes Leben verblasst (vgl. *Ready Player One*, 2018).

Diese perfekte virtuelle Welt könnte sich von Kindheit an um den/die *Spieler:in* entwickeln und seinen/ihren Präferenzen schrittweise anpassen, so dass der Wunsch jene zu verlassen vollständig verschwindet (vgl. Cypher in *Matrix*) – und vielleicht zusätzlich abschreckend tabuisiert ist, zum Beispiel als *Sterben/Tod* (*Inception*).

Damit eine solche „platonische Höhle“ nicht langweilig (oder durch ihre *ad-hoc* Struktur, Inkonsistenzen und Paradoxien als artifiziell erkannt) wird, ist sie *bevölkert* von Avataren geeigneter Künstlicher Intelligenzen zur Interaktion, Inspiration, Ablenkung sowie als Vermittler fortwährender Herausforderungen und Belohnungen (vgl. *eXistenZ* 1999 oder *The Good Place*).

Wir Menschen leben in einer komplexen Welt der Interaktion, von der Geburt bis zum Tod. Diese Welt ist keine dualistische Welt, in der es hier die Außenwelt, dort die Innenwelt, hier das Leben in der Höhle, dort das Leben außerhalb der Höhle gäbe. Utopien und Dystopien können helfen, die Probleme in der komplexen Welt, in der wir leben und arbeiten, kritisch zu reflektieren und dabei den Menschen und seine Interaktionen ins Zentrum zu stellen.

„There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened.“ (Douglas Adams)

Anmerkungen

- 1 Ausnahmen bestätigen die Regel, z. B. die Magnetschwebebahn in Deutschland.

- 2 Vgl. Thorsten Klein: *Algotokratie. Wie Algorithmen die Demokratie gefährden*. LIT Verlag Münster, 2021.
- 3 Alan M. Turing: *On Computable Numbers, with an Application to the Entscheidungsproblem*. In: *Proceedings of the London Mathematical Society*. Band 42, 1937, S. 230–265, doi:10.1112/plms/s2-42.1.230
- 4 Konrad Zuse: *Über den allgemeinen Plankalkül als Mittel zur Formulierung schematisch-kombinativer Aufgaben*. *Arch. Math.* 1, pp. 441–449, 1948/49.
- 5 Vgl. Florentin Neumann, Andrea Reichenberger, Martin Ziegler: *Variations of the Turing Test in the Age of Internet and Virtual Reality*. In: Mertsching, B., Hund, M., Aziz, Z. (eds.): *KI 2009: Advances in Artificial Intelligence*. *KI 2009. Lecture Notes in Computer Science*, vol 5803, pp 355–362. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04617-9_45
- 6 Marvin Minsky: *The Society of Mind*. New York: Simon & Schuster, 1986.
- 7 Leslie Valiant: *A Theory of the Learnable*. *Communications of the ACM* 27 (11), 1984, pp. 1134–1142; Julius Berner u. a.: *Die moderne Mathematik des tiefen Lernens*. *Mitteilungen der DMV* 29 (4), 2022, S. 191-197. <https://doi.org/10.1515/dmvm-2021-0074>.
- 8 Tobias Stahl: *Hammer-Urteil: Tesla-Fahrer dürften Touchscreen während der Fahrt nicht benutzen* (3. August 2020). URL: http://efahrer.chip.de/news/x_102753 [3. August 2020].
- 9 Martin Kords: *Neuzulassungen von Elektroautos in Deutschland bis 2021* (6. Mai 2022). URL: <https://de.statista.com/statistik/daten/studie/244000/umfrage/neuzulassungen-von-elektroautos-in-deutschland/> [Zugriff 15. Mai 2022].
- 10 David Brooks: *Turin. Approaching Animals*. Brandl & Schlesinger, 2022.
- 11 Carole Martinez: *Artificial Intelligence and Accessibility: Examples of a Technology that Serves People with Disabilities* (March 5, 2021). URL: <https://www.inclusivitymaker.com/artificial-intelligence-accessibility-examples-technology-serves-people-disabilities/> [Zugriff 15. Mai 2022].
- 12 S.S. Verma: *How Technology Is Enabling Communication With Animals* (September 17, 2019). URL: <https://www.electronicsforu.com/technology-trends/tech-focus/technology-enabling-communication-with-animals> [Zugriff 15. Mai 2022].



Andrea Reichenberger und Martin Ziegler

Andrea Reichenberger wurde 2014 im Fach Philosophie an der Universität Paderborn promoviert und hat in diversen Drittmittelprojekten gearbeitet. Sie ist Initiatorin und Sprecherin der Arbeitsgruppe *Frauen- und Genderforschung* der GWMT (Gesellschaft für Geschichte der Wissenschaften, der Medizin und der Technik e. V.) und derzeit als wissenschaftliche Mitarbeiterin an der FernUniversität in Hagen beschäftigt.

Martin Ziegler hat an der Universität Paderborn Informatik studiert, dort promoviert und habilitiert. Er war ab 2010 Professor für Angewandte Logik an der TU Darmstadt und ab 2013 Sprecher der IANUS Gruppe für naturwissenschaftliche Friedensforschung. Seit 2015 ist er Professor für Informatik am KAIST in Südkorea.

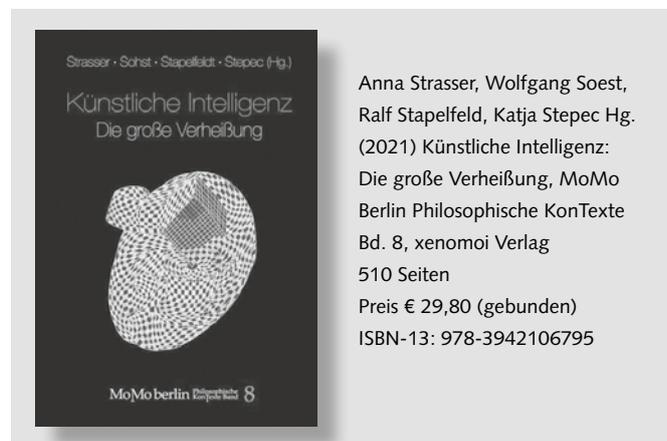
Anna Strasser, Wolfgang Soest, Ralf Stapelfeld, Katja Stepec (Hg.): Künstliche Intelligenz: Die große Verheißung Buchbesprechung

„Die Gefahr, dass der Computer so wird wie der Mensch,
ist nicht so groß wie die Gefahr, dass der Mensch so wird wie der Computer.“
(Konrad Zuse)

Tatsächlich befasst sich der Band kaum mit der Erklärung der Funktionsweisen von KI-Systemen, sondern im ersten Teil mit dem Begriff der menschlichen Intelligenz und dem Vergleich mit der sogenannten künstlichen Intelligenz und in einem zweiten Teil mit Grenzen und Folgen der „Künstlichen Intelligenz“ (KI). Für beide Teile zentral ist die Auseinandersetzung mit dem Turing-Test (dem Imitationsspiel, das auf der Ununterscheidbarkeit der Kommunikation von Menschen mit einem Automaten basiert, dabei jedoch so etwas wie „Verstehen“ ausblendet) mit der Turing-These (Intelligenz sei durch solche Art von Kommunikation definierbar), auch starke KI-These genannt, und der schwachen KI-These. Erstere vermeint, alle Aspekte menschlicher Intelligenz, also auch Bewusstsein, Selbstbewusstsein und Verstehen, durch einen qualitativen Sprung erreichen, wenn nicht gar übertreffen zu können, während letztere die Simulation von Elementen der analytischen Intelligenz für realistisch hält, nie jedoch alle Aspekte menschlicher Intelligenz.

Aus Sicht des feministischen Materialismus sind für die Eigenschaften menschlicher versus künstlicher Intelligenz das *Embodiment*, das in-der-Welt-Sein, die Intentionalität der handelnden Akteur:innen wesentlich für den Vergleich. Es fällt auf, dass zwar die hermeneutische Geschichtlichkeit des Menschen und die menschliche phylogenetische Entwicklung in „wetware“ als zentral für Eigenschaften der Intelligenz gesehen werden, aber die Körperlichkeit von (Software in) Hardware von Computern und Serverfarmen für KI-Systeme, mit all ihren Anforderungen an Stromzufuhr und -erzeugung, Metallabbau, -isolation und -bearbeitung nicht in Betracht gezogen wird. Ganz besonders fehlen solche Überlegungen bei jenen KI-Phantasten, die die Ablösung des Menschen durch eine höhere Intelligenz erwarten oder befürchten, als ob KI-Systeme selbständig in der Erde nach seltenen Erden buddeln, Chips herstellen oder Windräder zur Stromerzeugung aufstellen würden. Vielleicht sollen das alles dann Menschensklaven in einer KI-Diktatur erledigen. Aber mit welchen Druckmitteln könnte KI die Menschen dazu knechten? Aus feministischer Sicht ist bei diesen Diskussionen die allenthalben fehlende Materialität und Körperlichkeit, vor allem der Maschinen zu bemängeln, während die Historizität von Ideen eines künstlichen Menschen durchaus in den Blick genommen wird.

Da es keinen einheitlichen fächerübergreifenden Intelligenzbegriff gibt, und die Beurteilung der Fähigkeiten von KI abhängig ist von Definitionen von Mensch und Maschinen, werden hier im ersten Teil jeweils Aspekte von Intelligenz erörtert und in Abgrenzung zu beziehungsweise Überschneidung mit KI diskutiert. Die aus der Psychologie bekannten Bereiche analytische, soziale und emotionale Intelligenz sind der Mathematisierung beziehungsweise Maschinisierung unterschiedlich gut zugänglich. Während man sich auf eine prinzipiell erreichbare Simulation der analytischen Intelligenz leichter einigen kann, erfordert die soziale Intelligenz beispielsweise Welthaltigkeit und Empathie und die emotionale Letzteres sowie Selbstbewusstheit. Aus philosophischer und kulturwissenschaftlicher Sicht kommen andere Aspekte von Intelligenz zur Sprache, wie die für Berücksichtigung von Individualität notwendige subjektive Wahrnehmung und Gefühllichkeit, die bewusste oder zumindest kulturell geprägte Geschichtlichkeit, oder die onto- und phylogenetische



Anna Strasser, Wolfgang Soest,
Ralf Stapelfeld, Katja Stepec Hg.
(2021) Künstliche Intelligenz:
Die große Verheißung, MoMo
Berlin Philosophische KonTexte
Bd. 8, xenomoi Verlag
510 Seiten
Preis € 29,80 (gebunden)
ISBN-13: 978-3942106795

Leiblichkeit, die gefühlsmäßig erfahrene Kommunikation mit der Umwelt und damit die räumliche und zeitliche Zukunfts- und Welt-Offenheit.

Dass mathematische Funktionen nicht nur auf Zahlen und Buchstaben operieren können, sondern u. a. auch auf Worten und Sätzen (allesamt in Zahlen kodierbar), begünstigt m. E. Missverständnisse hinsichtlich semantischer oder gar sozialer und emotionaler Kompetenz der KI.

Nadine Schumann und Yaoli Du schlagen eine solche Betrachtung in einem ganzheitlichen Simulationsmodell vor, das auch menschliche Subjektivität erfasst. *Hans Zillmann* analysiert dafür auch die neurowissenschaftliche Subjektbeschreibung Antonio Damasio mit Bezug auf die Subjektivität, die er nicht nur sozialer und kultureller Zuschreibung zuordnet, sondern auch im Kontext emotionaler Erfassung der Umwelt sieht. Vorsichtig vermutet er, dass Gefühle Maschinen wohl unzugänglich seien, wobei er – unzutreffend – Maschinen ein eher deduktives Vorgehen zuschreibt. Aber zwischen induktivem Schließen, Subsumption etc. und Gefühlen klafft, von ihm unbetrachtet, immer noch eine große Lücke.

Daniel Dennett stellt KI in den Kontext von Evolutionstheorie und Darwins Thesen und fordert zum Verständnis des Ver-

hältnisses von Maschine und Mensch einen Perspektivwechsel, gewissermaßen beider *bottom-up*-Evolution, womit er dann eine Art „als ob“-Perspektive einnimmt, ähnlich einem digitalen Zwilling. *Michael Meyer-Albert* ruft eine Reihe von Philosophen auf, um drei Eigenschaften von Intelligenz herauszustellen, die der Ko-Konstruktion von Denken und Wahrnehmen Schwierigkeiten machen könnten: einmal die körperliche Kommunikation mit der Außenwelt, welche Geist und Welt mit Gefühlen verbindet, zum zweiten die hermeneutische Dimension Intelligenz als ein geschichtliches Verstehen von Sinn, in Absetzung von archivarisch hinterlegter Information. Intelligenz erfordert demgemäß zeitliche und geschichtliche Reflexivität, um eine Interaktion der Gegenwart mit der Tradition zu ermöglichen. Diese sich in Zusammenhängen modifizierende „Sinnbewegung des Verstehens“ (Gadamer) erschließt eine als leiblich erfahrene und geschichtlich reflektierte Welt-Offenheit des Menschen. Harte KI-ler:innen könnten gegen die Behauptung: „Das Objekt des Verstehens, das Verständnis des Verstehens und das Subjekt des Verstehens verändern sich im maschinellen Verstehen nicht“ einwenden, dass sich beim *deep learning* das *Subjekt* und das *Verständnis* des Verstehens, also das Modell, und das *Objekt*, also das zu Erkennende sehr wohl verändern. Nur ist dagegen zu fragen, was bedeutet hier *Verstehen*?. Schließlich führt Meyer-Albert ethische Bedenken gegen die Erschaffung künstlicher Intelligenzen auf, indem er das *Existieren* mitsamt seinen Bedingungen von Sterblichkeit und Angst ins Treffen führt. *Gutmann, Haag und Wadephul* führen mit einem tätigkeitsbezogenen Wissensbegriff, der auf den Verwendungszweck abhebt, abgesetzt von den Begriffen Information und Daten, zu einer Problematisierung von Maschinenzwecken gegenüber reflexivem Wissen mit selbstständiger Zwecksetzung, Rationalität und Verantwortung. Solche zweckorientierte Verwendung führt *Christian Freska* zu einer weiteren Voraussetzung für soziale Intelligenz, nämlich für ökologisches Problemlösen.

Im Gegensatz dazu argumentiert *Jan Tobias Fuhrmann*, dass Kommunikation nicht notwendigerweise Bewusstsein und Verstehen voraussetzt, weshalb er Programmen die Fähigkeit zugesteht, einen Turing-Test zu bestehen, ohne dass sie für sich selbst Sinn erzeugen müssten. Ja, ihm scheint die fehlende Reflexion gerade ein Vorzug für die Anschlussfähigkeit von Kommunikation zu sein. Aber genügt das für die Analogie beider Intelligenzen?

Christoph Merdes befasst sich mit der Frage, ob man Maschinen Moral beibringen kann, und modifiziert dafür einen Turing-Test für moralische Urteile, der sich im Grunde von einem sprachlichen Turing-Test nicht unterscheidet, aber für die Rechtfertigung des Urteils auf einen Gruppen-Turing-Test erweitert. Dabei stellt er fest, dass es zu Inkonsistenzen kommen kann, denn die Wertvorstellungen innerhalb der Gruppe können sich unterscheiden, aha! Wenn er auch die Möglichkeit einer programmierten oder freigelassenen KI für Emotionalität und Urteilsvermögen nicht ganz ausschließt, stellt er fest, dass die bisherigen Erfolge diesbezüglich nicht ausreichen.

Daniel Wenz diskutiert das mathematische Beweisen mittels KI und setzt es gegen das verstehende Nachvollziehen und ganzheitliche Verstehen der bewiesenen mathematischen Behauptung durch eine Anzahl von mathematischen Peers ab. So begründen Mathematiker:innen die Akzeptanz der Gültigkeit von

Beweisen in Abgrenzung von noch so langen Computer-Rechnungen, aber auch von immer möglichen Hardware-Fehlern.

Der philosophische Logiker *Walther Zimmerli* befasst sich schließlich mit der KI zwischen sinnstiftendem Mythos und der Verselbständigung in einem Hype. Dieser kulminiert schließlich in der Behauptung des immer wieder verschobenen Zeitpunkts der Kurzweilschen „Singularität“, also des Zeitpunkts, an dem die künstliche die menschliche Intelligenz übertrifft, und weiter sogar eines „postanalogen Menschseins“, in dem auch einem religiös verbrämten Trans- und Posthumanismus gehuldigt oder in dystopisch Bostromscher Manier Furcht vor einer Super-Intelligenz erzeugt wird. Er zeigt dazu die ideengeschichtlichen, literarischen und philosophischen Erzählungen, Voraussetzungen und Phantasmen, und auch die keineswegs unerheblichen Mittel, die für die Realisierung dieser Mythen eingesetzt werden. Wo und wie, um alles in der Welt, werden diese ungeheuren Materialisierungen in Servern für Big Data und (u. a. Lern-) Software enden? Wozu das alles, ist die Frage.

Der zweite Teil des Bandes beginnt mit einem Überblick von *Ralf Stapelfeld* über Diskussionen zur KI, ehe er einen Überblick über die folgenden Texte gibt. Dabei kommen insbesondere auf den Seiten 244-250 wichtige, vor allem logisch-philosophische Überlegungen zur Sprache, die danach nicht mehr behandelt werden.

Hans-Jörg Kreowskis und Wolfgang Kriegers Beitrag: „Künstliche Intelligenz – ‚künstlich‘ ja, ‚Intelligenz‘ wohl kaum“: Nach einer Darstellung der 50-jährigen Geschichte der KI und Robotik und ihrer institutionellen und militärischen Förderungen, großartigen Prognosen und Rückschlägen, begann in den letzten Jahrzehnten der große Hype nach einzelnen Erfolgen in eng begrenzten Anwendungsbereichen, wie Schach und Go, aber auch Entscheidungssystemen und Szenariensimulationen, ermöglicht durch die enorm vergrößerte Speicherkapazität und Rechengeschwindigkeit. Große Beachtung finden insbesondere konnektionistische Systeme (deep learning) im Zusammenhang mit Sprach- und Bilderkennung im Verein mit dem so genannten *Dataismus*, also der Fütterung lernender Maschinen mit ungeheuren Mengen von Bild-, Text- und Sprach-Daten aus all unseren Bewegungen im Digitalen, von denen man sich Anwendungen und Wertschöpfung in fast allen Bereichen heutigen Lebens erhofft. Die Autoren stellen die staatlichen Hoffnungen und stattlichen Förderungen für politische Wunschträume in aller Breite genüsslich dar, bevor sie sich dem Vergleich zwischen künstlicher und natürlicher Intelligenz widmen. Hier wird auch eine Kritik am ersten Teil des Buches geäußert, dass nämlich dort Intelligenz kaum direkt diskutiert wird, sondern eher Vernunft, Geist und Seele, weshalb man an dieser Stelle eher von technischer Realisierbarkeit von künstlicher Vernunft, künstlichem Geist und künstlicher Seele sprechen sollte.

Bei der Einordnung der Fähigkeiten von KI beginnen die Autoren mit der Turing-These als der Definition von Berechenbarkeit via Turing-Maschinen, die dadurch gestützt wird, dass bisher alle universellen Computermodelle und Programmiersprachen beweisbar äquivalent zu Turing-Maschinen sind. Die Beschränkungen der Berechenbarkeit (so kann mit reellen Zahlen nur approximativ gerechnet werden, oder es lässt sich im Allgemeinen nicht nachweisen, ob ein Programm im Sinne der Absicht kor-

rekt läuft) und die praktischen Verengungen durch Komplexitätsüberschreitung setzen Allmachtsphantasien der Computerei harte Grenzen, insbesondere innerhalb der Mathematik.

Diskutierbar wäre allerdings die Churchsche These, welche alle denkbaren Berechenbarkeitsmodelle hinsichtlich ihrer Funktionalität für äquivalent hält, weil alle bisher definierten es waren. Akzeptiert man die Churchsche These, wie es für die Berechenbarkeitstheorie zwar nicht notwendig, aber sinnvoll erscheint, also jener (durch Komplexitätseinschränkungen enorm zu reduzierenden, daher funktional nochmals ungleich größeren) Basistheorie der Computabilität (Berechenbarkeit, Nicht-/Entscheidbarkeit), so kollabiert der Vergleich zwischen Turing-Test als (angeblichem) Nachweis menschlicher analytischer (soziale oder emotionale Intelligenz ohnedies ausgenommen) und künstlicher Intelligenz. Die starke KI-These geht ja davon aus, dass kognitive Prozesse nichts anderes sind als Informationsverarbeitungs- und damit Berechnungsprozesse auf symbolischen Repräsentationen. Diese These impliziert die prinzipielle Simulierbarkeit menschlicher Intelligenz auf Computern (sehen wir von Komplexitätsproblemen einmal ab), sonst stünde sie im Widerspruch zur Churchschen These.

Ob KI der natürlichen Intelligenz gleichen kann, erörtern Kreowski und Krieger anhand der Turing-These – und lassen eine Antwort zwar offen, halten aber eine nachmalige Bejahung doch für unwahrscheinlich, in jedem Fall beide Begriffszugänge für sehr verschieden. Am Ende werden die Gefahren der Anwendungen von KI in der Gesellschaft (Überwachung) in Waffen und Wirtschaft und Politik diskutiert.

Während Kreowski und Krieger auf den Seiten 272 und 273 (empfehlenswert!) sehr einfach und verständlich die Prinzipien und Ergebnisse der Berechenbarkeitstheorie erklären, greift Dieter Mersch, nach einer kurzen Revue der Versuche, das menschliche neuronale System nachzubilden, und der philosophischen Abwägungen zwischen intentionalem Handeln, Denken und Rechnen, deutlich formaler in die Kiste der logischen Theorie des Konstruktiven. Um für die Beurteilung von Intelligenz die Möglichkeiten formaler Sprachen einzuordnen, kommt er in Anlehnung an die von Lucas 1961 angestoßene Diskussion über die Gödelschen Sätze, die die Unentscheidbarkeit wahrer Sätze innerhalb genügend ausdrucksstarker Formaler Systeme nachweisen, zu dem Schluss, dass auch für maschinelles Argumentieren menschliches Denken teilweise unerreichbar sein müsse. Stimmt, wenn Maschine und Mensch getrennt bleiben, aber dem lässt sich entgegenhalten, dass KI großenteils auf dem Vergleich von in Big Data gespeicherten Texten oder Bildern beruht, wobei Letztere zwar kein *Verständnis* für sich selbst haben, sich in ihrer Bedeutung für die beobachtenden Menschen aber der Einordnung in formale Systeme entziehen, während der auto-

matische Text- oder Bild-Abgleich selbst sich innerhalb der formalen Systeme bewegt. Dass „die weithin geteilten Einsichten von Kurt Gödels Unvollständigkeitssatz“ zu dem Schluss führen, dass der menschliche Geist prinzipiell die Möglichkeiten eines formalen Systems übersteigt, demonstriert ein Missverständnis mancher Philosophen, mathematische Erkenntnisse für diskutierbar zu halten, die ja nur innerhalb explizit definierter formaler Systeme Gültigkeit beanspruchen. Dies gilt vor allem für die konstruktive Mathematik, innerhalb derer sich Gödels Theorie der Berechenbarkeit bewegt. (L. E. J. Brouwers Einwand, dass die Mathematik des *tertium non datur* oder eine, die sich des Auswahlaxioms bedienen muss, auf tönernen Füßen stehe, bezieht sich ja gerade auf jenen ungeheuer größeren Teil der Mathematik, der sich indirekter, also nicht konstruktiver Beweise bedient. Allerdings vertritt er mit seinem Intuitionismus der intuitionistischen Logik, die ein mathematisches Objekt als Produkt der konstruierenden Intuition des Mathematikers entstanden und von ihm dabei überhaupt erst hergestellt sieht – ein für die starke Turing-These der KI befremdlicher Gedankengang.) Am Ende erwägt Mersch die (Un-) Möglichkeit algorithmischer Kreativität in Kunst und Literatur vor dem Hintergrund eines für soziale Menschen notwendigen ästhetischen Ethos.

Elektra Wagenrad vergleicht Computer mit dem klugen Hans, einem sprechenden und rechnenden Pferd, dem sie ein Verständnis für seine geklopften Antworten abspricht. Die Dialoge des klugen Hans mit seinem Besitzer hat der Schriftsteller Clemens J. Setz in seiner ungewöhnlichen Büchner-Preis-Rede rekonstruiert¹ und als Metapher für die (Un-)Möglichkeit menschlicher Verständigung herangezogen. Oder kann diese Metapher – er ist auch Mathematiker – gerade auch zur Stützung von Turings These gelten? Bedeutet denn dasselbe Wort einem anderen Menschen dasselbe? Die Klopffzeichen des Hans werden nach Aufforderung für eine Antwort-Interpretation genauer, und ähneln so den durch wiederholtes Nicken angegebenen Buchstaben und -folgen von Locked-in-Patient:innen mit Brain-Computer-Interface. Weitergehend befasst Wagenrad sich – den Mangel an Eigenwillen von KI im Vergleich zum menschlichen Entscheidungsverhalten problematisierend – mit dem neurowissenschaftlichen Libet-Experiment, das die Intentionalität bewusster menschlicher Entscheidungen infrage stellt. Allerdings sind die Versuchsbedingungen für das elektrophysiologisch gemessene Bereitschaftspotenzial so eng und unnatürlich, dass der Umfang ihrer Bedeutung gegen bewusste freie Entscheidungen durchaus infrage steht. So muss man noch nicht Wittgensteins unterschiedliche Sprachspiele bemühen, um die Unvergleichbarkeit beider Systeme von Intelligenzbegrifflichkeit festzustellen. Schließlich geht sie auf das erstaunlich „einfühl-same“ Sprachsystem GPT-3 ein, das sie selbst in einem spielerischen Experiment nicht nur zu sich selbst widersprechendem



Britta Schinzel

Britta Schinzel promovierte in Mathematik, arbeitete in der Computerindustrie und habilitierte sich in der Informatik. Im Rahmen ihrer Professur für Theoretische Informatik an der RWTH Aachen arbeitete sie zunehmend interdisziplinär. Sie war von 1991 bis 2008 Professorin für Informatik und Gesellschaft und Gender Studies in Informatik und Naturwissenschaft an der Universität Freiburg.

Plappern, sondern auch teilweise zu philosophisch sinnvollen Antworten bringen konnte. Allerdings, so lässt sich einwenden, weiß GPT-3 nichts davon, ob es sinnvolle oder unsinnige Antworten gibt, und Menschen sollten die für sie sinnvollen nicht als systemimmanent sinnvoll missdeuten.

Vladova und Friesike konzentrieren sich auf das Lernen aus Irrtümern als wichtigen, für KI unerreichbaren Aspekt menschlicher Intelligenz. Auch dagegen lässt sich einwenden, dass die Methode beim Backtracking im Deep Learning gerade das Annähern an fehlerhafte Ergebnisse ist. Aber es ist wohl gemeint, dass Fehler neue oder unerwartete Lösungen in Gang setzen.

Sybille Krämer lenkt den Blick auf einfache Anwendungen von schwacher KI (wobei sie diese mit allgemeiner Software identifiziert) und sieht für diese eine weitere Verflachung der menschlichen Kultur in Text und Bild auf das Maschinenlesbare. Im Weiteren problematisiert sie innerhalb der KI das Maschinelle Lernen und stellt sehr in Frage, ob es je möglich wäre, dass Maschinen Sinn und Bedeutungszusammenhänge verstehen können sollten.

Rico Hauswald vergleicht Deep-Learning-Ergebnisse mit Orakeln und sieht darin Vorteile (Demokratisierung von Wissen) und Gefahren, da die Black Box von Urteilsfähigkeit entwöhnt. Allerdings ist beides nicht KI-typisch, dafür genügt schon Google und klassische Entscheidungsalgorithmen. Thomas Weiss hinwiederum diskutiert die Vollautomatisierung im Hinblick auf Arbeitsplatzverluste und phantasiert über sich emanzipierende Artificial General Intelligence von einem KI-Proletariat.

Reinhard Kahle problematisiert die rückwärtsgewandten Entscheidungen durch KI, wenn sie auf vergangenen Datensammlungen beruhen, beispielsweise im Hinblick auf ein Ausbleiben von Innovationen sowie den Autonomieverlust von Menschen bei der Übernahme von Aufgaben durch KI.

Engel und Schultheis befassen sich ebenfalls mit KI-basierten Entscheidungen anhand einer Delphi-Studie zu den Erwartungen der Befragten über die entsprechenden Veränderungen. Daran anschließend erörtern sie, wieviel Vertrauen in diese Technologie gesetzt werden kann oder sollte, was sie weiter zu moralischen Fragestellungen führt.

Mit Ethik und Moral im Zusammenhang mit KI befassen sich die weiteren Texte, wobei zu unterscheiden ist, ob sich die ethischen Fragen auf KI-Ergebnisse beziehen, oder ob versucht werden soll, ethische Anforderungen ins Design von KI-Entwicklungen einzubeziehen. *John Michael und Hendrik Kempt* diskutieren den Aufbau von Beziehungen und Vertrauen zwischen Menschen und Robotern in Grenzen durchaus positiv, während *Orphelia Deroy* gerade die Möglichkeit von Vertrauen in KI problematisiert und dazu Rechtfertigung verlangt. *Catrin Misselhorn* hält hingegen eine Maschinenethik für möglich, fordert dazu aber drei Grundsätze, von denen zwei meines Erachtens im Widerspruch zur Funktionsweise von KI stehen und so ihre Anwendung aushebeln würden, etwa wenn der Mensch immer die Verantwortung für KI-Entscheidungen tragen soll.

Ein wenig enttäuschend ist, dass das Buch wenig auf die Funktionsweise von KI zum Verständnis ihrer Möglichkeiten und Unmöglichkeiten eingeht. Ausnahmen bilden nur Hans-Jörg Krewowskis und Wolfgang Kriegers Darstellung der Entscheidbarkeit und Berechenbarkeit S. 272 ff. und Sybille Krämers Darstellung von Deep Learning ab S. 346, dies alles auch nur im Bereich des Formalen. Hingegen scheint es mir erstaunlich, dass die Körperlichkeit von (Soft- in) Hardware von Computern und Serverfarmen für KI-Systeme, mit all ihren ungeheuren Konsum von und Anforderungen an Stromzufuhr und seine Erzeugung, Metallabbau und seine Isolation und Bearbeitung nicht in Betracht gezogen wird. Ganz besonders fehlen solche Überlegungen bei jenen KI-Phantasten, die die Ablösung des Menschen durch eine höhere Intelligenz erwarten oder befürchten, als ob KI-Systeme selbständig in der Erde nach seltenen Erden buddeln, Chips herstellen oder Windräder zur Stromerzeugung aufstellen würden. Vielleicht sollen das alles dann Menschensklaven in einer KI-Diktatur erledigen. Aber mit welchen Druckmitteln könnte KI die Menschen dazu knechten?

Anmerkung

1 <https://www.sueddeutsche.de/kultur/clemens-j-setz-georg-buechner-preis-woyzeck-karl-krall-denkende-tiere-pferde-1.5457889>



Peter Brödner

Die Illusionsfabrik der ›KI‹-Narrative

Derzeit sind medial verbreitete ›KI‹-Narrative wieder en vogue. In den 1980er-Jahren versuchten Ansätze symbolischer KI, explizites Wissen über Praktiken kooperativer kognitiver Arbeit und daraus zu ziehende Schlüsse in Gestalt wissensbasierter oder Expertensysteme zu modellieren. Im Unterschied dazu richten sich heutige Ansätze darauf, zwecks Gewinnung von Berechnungsverfahren zur Bewältigung kognitiver Aufgaben die Mühen analytischer Durchdringung und Modellierung mittels Verfahren maschinellen Lernens (ML) zu umgehen – tatsächlich aber nur eine Art Funktions-Approximation an große Mengen vorgegebener Daten. Während erstere an den hohen Hürden hinreichender Analyse und Explizierbarkeit impliziten Wissens gescheitert sind, werfen die neuen Ansätze erneut unüberwindlich erscheinende Probleme auf. Zum besseren Verständnis wird zunächst anhand üblicher ›KI‹-Definitionen gezeigt, dass ›KI‹-Protagonisten nicht einmal wissen können, worin sich die künstlich intelligent genannten Systeme eigentlich genau von herkömmlichen Computersystemen unterscheiden – ein Umstand, aus dem viele Illusionen über Funktionsweisen und Leistungspotenziale dieser Systeme erwachsen. Neue Probleme ergeben sich einerseits aus der kaum einschätzbaren Relevanz und Validität der Daten, zudem aus der Intentionalität und Kontingenz sozialer Praktiken, andererseits aus einer höheren Art der Undurchschaubarkeit des Systemverhaltens. Das wirft zudem eine Reihe neuer, freilich noch ungelöster ethischer Fragen auf.

Was ist eigentlich ein ›KI-System?‹

Derzeit redet alle Welt nach längerer Pause wieder über *künstliche Intelligenz (KI)* als zukunftsweisende Computertechnik. Angesichts dessen darf angenommen werden, dass einigermaßen klar ist, was diese Technik besonders kennzeichnet, worüber ein Blick auf übliche ›KI-Definitionen‹ Auskunft geben sollte. Diese lassen sich in zwei Gruppen einteilen: Die erste Gruppe begreift Computertechnik als ›KI- bzw. ›AI-System‹, wenn die Lösung der Aufgaben, zu deren Bewältigung es geschaffen wird, natürliche Intelligenz und Erfahrung erfordert:

„AI is the part of computer science concerned with ... systems that exhibit characteristics we associate with intelligence in human behaviour – understanding language, learning, reasoning, problem solving, and so on.“ (Barr & Feigenbaum 1981; ähnlich auch schon McCarthy 1955);

Neuerdings auch:

Systems „that are capable of performing tasks commonly thought to require intelligence. Machine learning ... refers to the development of digital systems that improve their performance on a given task over time through experience“ (Autorengruppe 2018: 9).

Eine zweite Gruppe von Definitionen schreibt ›KI-Systemen‹ eine gewisse eigenständige „Handlungsträgerschaft“ (agency) zu:

AI research investigates „intelligent agents“, i. e. devices „that perceive their environment and take actions maximizing the chance of successfully achieving their goals“ (Russell & Norvig 2009: 2);

„AI researchers use mostly the notion of rationality, which refers to the ability to choose the best action to take in order to achieve a certain goal, given certain criteria to be optimized and the available resources.“ (High-Level Expert Group on AI 2018: 1 f).

Bei näherem Hinsehen erweisen sich diese Bestimmungen jedoch als reine Scheindefinitionen: Der ersten Gruppe zufolge sollen sich ›KI-Systeme‹ von *gewöhnlichen* Computersystemen, selbst anderen technischen Artefakten, dadurch unterscheiden, dass die Bewältigung der Aufgaben, für die sie geschaffen werden, natürliche Intelligenz erfordert. Eben dies gilt aber auch schon für die Lösung relativ einfacher kognitiver Aufgaben wie die Bestimmung der Nullstelle einer quadratischen Gleichung oder das Spiel der *Türme von Hanoi*, die ebenfalls ein beträchtliches Maß an Intelligenz erfordern (das schon die Fähigkeiten vieler Menschen übersteigt, ganz abgesehen davon, dass auch die Ausübung körperlicher Arbeit meist hohe Intelligenz voraussetzt). So wäre diesen Definitionen zufolge auch jedes andere auf einem Computer ausgeführte Berechnungsverfahren ein ›KI-System‹, die vermeintliche *differentia specifica* unterscheidet nicht wirklich.

Bei der zweiten Definitionsgruppe werden *KI-Systemen* die typischen Merkmale von Intentionalität und rationalem Handeln, die Wahl geeigneter Mittel zum Erreichen von Zielen, einfach

zugeschrieben. Tatsächlich befolgen diese aber, wie alle Computersysteme, nur ein ihr Verhalten determinierendes Programm, dem bereits alle denkbaren Bedingungen methodisch eingeschrieben sind, unter denen von außen festgelegte Ziele bestmöglich zu erreichen sind. Hier werden Herstellen und Hergestelltes, die intelligenten Tätigkeiten des Entwerfens und Programmierens mit den Leistungen des Programms als deren Ergebnis verwechselt – ein krasser Kategorienfehler. Tatsächlich vergegenständlicht das Programm lediglich Ergebnisse des lebendigen Arbeitsvermögens, der Intelligenz, Erfahrungen und Fähigkeiten seiner Schöpfer, vorgestellte Ziele unter angenommenen Bedingungen mit Mitteln der Logik und Verfahren der Berechnung bestmöglich zu verwirklichen.

Somit bleibt festzuhalten, dass aufgrund dieser Definitionen niemand wirklich wissen kann, was ein ›KI-System‹ eigentlich ist, paradoxerweise auch jene nicht, die ständig davon reden – ein eklatanter Fall von *Technik und Wissenschaft als ‚Ideologie‘* (Habermas 1968; vgl. Brödner 1997: Kap. 4.4). Jede Art von Computerprogramm ist schließlich nur eine Vergegenständlichung des lebendigen Arbeitsvermögens und der Einsichten natürlicher Intelligenz seiner Konstrukteure – eine Feststellung, die freilich auch für jedes andere technische Artefakt gilt, vom Faustkeil bis zum Computer.

Die Mühen der Modellierung von Praxis und der vergebliche Versuch ihrer Umgehung

Computersysteme, gleich welcher Komplexität, führen berechenbare Funktionen auf binären Schaltsystemen aus und nichts sonst. Gestaltung und Einsatz erfordern die Modellierung und Formalisierung sozialer Praktiken kooperativer kognitiver Arbeit, ein schwieriger, hohe Einsichtsfähigkeit und Nutzerbeteiligung verlangender Vorgang, der leicht misslingen kann (Rohde et al. 2017). Dabei muss die klaffende semantische Lücke zwischen der Praxis und deren sprachlicher Beschreibung einerseits und Programmen als formalen Beschreibungen maschinell ausführbarer Berechnungsverfahren andererseits überwunden werden (Programmiersprachen helfen dabei). Die nötige Modellbildung in aufgabenorientiert reduzierender Perspektive – Kern der Softwaretechnik – durchläuft folgende Schritte der Reduktion, Abstraktion und Formalisierung (Andelfinger 1997):

- *Semiotisierung*: Begrifflich-propositionale Beschreibung der Aufgaben und Abläufe einer sozialen Praxis mittels Zeichen liefert ein perspektivisch reduziertes Abbild von Wirklichkeit als Ergebnis gemeinsamer Reflexion und Kommunikation der Akteure (Sprachanalyse, Ontologie):
→ Anwendungsmodell.
- *Formalisierung*: Abstraktion von situations- und kontextgebundenen Bedeutungen und Reduktion auf sinnfreie Standardzeichen und -operationen:
→ formales Modell (Spezifikation).
- *Algorithmisierung*: Überführung von Gegenständen und Abläufen des formalen Modells in auto-operational ausführbare Prozeduren in Form von Daten und berechenbaren Funktionen (Algorithmen):
→ Berechnungsmodell (als Grundlage der Programmierung).

Sprachlich repräsentierte Vorgänge kooperativer kognitiver Arbeit können so partiell formalisiert und dann als berechenbare Funktionen (Algorithmen) maschinell ausgeführt werden – auch Menschen rechnen formalisiert wie Maschinen, ihre Fähigkeiten sind aber nicht darauf beschränkt (daher gilt der Einsatz von Computern auch als *Maschinisierung von Kopfarbeit*; Nake 1992).

Die Ausführung der berechenbaren Funktionen stellt einen „degenerierten“, auf eine dyadische Relation reduzierten Zeichenprozess ohne „Fenster zur Welt“ dar, dem der Bezug zu einem erlebten, leiblich erfahrenen oder gedachten Objekt, eben die *Bezeichnung* fehlt. Es ist nur eine *Quasi-Semiose*, die mit Signalen (logisch: Daten) als auf Syntax reduzierten *Quasi-Zeichen* operiert (Nöth 2002). Deren Zustände werden per Programm rein physisch transformiert ohne Ansehen von Bedeutung. Im Computersystem implementiert entstehen damit *auto-operationale Formen* (Floyd 2002) als Ausdruck abstrakter, formalisierter Operationen. Deren Sinn muss durch Aneignung seitens der Systemnutzer für wirksamen praktischen Gebrauch erst noch erschlossen werden.

Dabei ist zwischen Problem und Aufgabe zu unterscheiden (Dörner 1983): Ein *Problem* liegt vor, wenn die Mittel zum Erreichen eines angestrebten Ziels noch unbekannt sind oder über das Ziel keine klaren Vorstellungen bestehen, wenn handelnde Personen also nicht wissen, wie sie ihr Ziel erreichen sollen: „*Intelligenz ist das, was man einsetzt, wenn man nicht weiß, was man tun soll.*“ (J. Piaget). Gesucht sind dann Ideen für abduktives Schließen, d. h. die Bildung von erklärenden Hypothesen aufgrund von Intuition, Analogie oder Kreativität (Peirce 1878). Bewährt sich eine Hypothese, können damit Verfahren zur methodischen Bewältigung der dem Problem entsprechenden Aufgaben gewonnen werden (Popper 1994).

Davon unterscheiden sich *Aufgaben* als geistige Anforderungen, für deren Bewältigung Methoden oder Verfahren bereits existieren. Aufgaben erfordern lediglich den Einsatz bekannter Mittel auf gewohnte Weise; als Instanzen eines prinzipiell bereits gelösten Problems erfordert ihre Lösung lediglich den routinierten Gebrauch dafür angeeigneter Methoden oder Verfahren (einschließlich der Beurteilung ihrer jeweiligen Eignung).

Die Modellierung einer komplexen sozialen Praxis beginnt als Problemlösung: Anfangs sind weder das Problem noch dessen Lösung hinreichend durchschaut; sie müssen im Zuge der Semiotisierung erst durch Analyse und Genese expliziten Wissens verstanden werden, um gesicherte Methoden der Bewältigung zu gewinnen. Dadurch wird die weitere Modellierung zur Aufgabe reduziert und durch Anwendung des Lösungsverfahrens bewältigt. In der Problemanalyse, der Wissensgenese, der Schaffung formalisierter Lösungsverfahren und der Beurteilung ihrer Eignung erweist sich die natürliche Intelligenz der Akteure, während die Leistung des Computersystems auf die Ausführung des daraus entstandenen programmierten Berechnungsmodells beschränkt ist, ggf. unter Berücksichtigung äußerer Bedingungen.

Mit der derzeit im Zentrum des Interesses stehenden Verfahren *maschinellen Lernens* und der Nutzung von *Big Data* wird versucht, sich diese Mühen von Problemanalyse, Modellierung, Formalisierung und Bestimmung eines spezifischen Berechnungsmodells zu ersparen. Stattdessen werden einfach für

ganze Klassen von Aufgaben – darunter Aufgaben der Objekt-Klassifizierung, der Clusterung von Objekten oder automatisierter Entscheidung – erfahrungsbasiert oder schlicht aufgrund theorieleeren Probierens geeignet erscheinende, generische mathematische Funktionen ausgewählt, deren offene Parameter noch aufgabenspezifisch zu bestimmen sind. Solche Funktionen sind etwa *künstliche neuronale Netze* (KNN) mit ihren Gewichten, Polynome oder logistische Funktionen mit ihren Koeffizienten oder Entscheidungs-bäume mit ihren Kantengewichten als Parametern.

Die Parameter werden mittels meist längst bekannter Verfahren der Funktions-Approximation möglichst gut an große Mengen verfügbarer Datenobjekte angepasst, was sie als *adaptive Systeme* kennzeichnet. Die so für die Bewältigung einer spezifischen Aufgabe *trainierten* Funktionen lassen sich auf neue Datenobjekte gleicher Art anwenden, vorausgesetzt, der infolge prinzipieller Kontingenz sozialer Praktiken veränderliche Kontext bleibt erhalten. Dieses Vorgehen mag in je besonderen Einzelfällen durchaus gelingen, setzt aber meist enorme Rechenleistung voraus (die jüngst erst verfügbar ist). Diese Art *maschinellen Lernens* hat aber nichts mit herkömmlichem Verständnis reflexiven, auf Einsicht beruhenden Lernens zu tun und ist insofern eine irreführende Benennung. Der Erfolg steht und fällt mit den zum *Training* benutzten Daten, deren Herkunft und Qualität aber meist nicht einschätzbar und hinsichtlich Repräsentativität und Verzerrungen (Biases) oft äußerst fragwürdig sind.

Ein solches Vorgehen hat den hohen Preis, dass grundsätzlich nur wahrscheinliche, von der Vorgeschichte abhängige, daher stets unsichere Ergebnisse zu erwarten sind, deren Validität kaum zu beurteilen ist – eine Art postmoderner Obskurantismus, unreflektierter Datengläubigkeit geschuldet. Den berechneten Ergebnissen kann man nur blind vertrauen, weil sich aktuell nicht nachvollziehen lässt, wie sie im Einzelnen zustande gekommen und wie zuverlässig sie sind. Solch neue Art undurchschaubares Systemverhalten hat im Gebrauch allerdings höchst abträgliche Folgen, die reflexives Lernen behindern und großes Stresspotenzial aufweisen.

Ungelöste ethische Fragen

Die hinsichtlich Validität, Repräsentativität und Aktualität oftmals unsichere Qualität der Daten über zugrunde liegende reale Vorgänge wirft, zusammen mit der Intransparenz und Variabilität des Systemverhaltens und der prinzipiellen Unsicherheit berechneter Ergebnisse, schwerwiegende ethische Fragen auf (Mittelstadt et al. 2016): Dürfen derartige potenziell gefährliche Systeme überhaupt in praktischen Einsatz gelangen? Wie lassen sich deren Sicherheit und (auch nicht intendierten) Schadenspotenziale im Vorhinein bewerten, und wer wird im Schadensfall zu auch haftender Verantwortung gezogen – die Hersteller, die Betreiber oder gar einzelne Nutzer? Dazu werden etwa unter dem Stichwort „trustworthy AI“ zwar weithin allgemeine Bewertungskriterien diskutiert, fraglich bleibt jedoch, ob wegen genannter Systemeigenschaften konkrete Regelungen überhaupt verbindlich festlegbar sind (vgl. z. B. High-Level Expert Group on AI 2019) – man fragt sich, ob nicht hinter der Fassade großer ethischer Besorgnis riskante Entwicklungen einfach weiter betrieben werden sollen.

Zudem ist, in Anbetracht der begrenzt erscheinenden Möglichkeiten vollständiger Automatisierung ganzer Arbeitsprozesse, auf absehbare Zeit damit zu rechnen, dass Wissensarbeiter mit adaptiven Systemen zusammen werden arbeiten müssen. Wegen deren undurchschaubaren Eigenlebens ist statt bislang üblicher Interaktion aber nur noch Ko-Aktion möglich, die zweckorientierten instrumentellen Gebrauch der Systeme erheblich erschwert und deren Nutzer mit beträchtlichen Handlungshindernissen konfrontiert: Unter dem Druck zugewiesener Leistungsforderungen und Eigenverantwortung können sie den Resultaten mangels Urteilsfähigkeit nur blind vertrauen, leiden mithin unter großer Unsicherheit, jedoch ohne die Möglichkeit, sich das Systemverhalten hinreichend aneignen zu können. Das führt gesicherten arbeitswissenschaftlichen Erkenntnissen zufolge zu beträchtlichen spezifischen Belastungen und Stressreaktionen (Brödner 2020).

Dementsprechend wird auch von vielen Seiten vehement gefordert, die Systeme mit Komponenten auszurüsten, die das Zustandekommen ihrer Resultate auf Verlangen mit hinreichender Detailwiedergabe zu erklären und damit auch den instrumentellen Gebrauch zu erleichtern vermögen (*explainable AI*). Deren Realisierung steht aber, so überhaupt möglich, noch in weiter Ferne, und solange es sie noch nicht gibt, sollte der Einsatz adaptiver Systeme im Interesse effizient und sozialverträglich gestalteter Arbeitsprozesse unbedingt vermieden werden.

Fazit

Während derzeit viel und durchaus zurecht von ethischen Herausforderungen durch „ML“-Systeme die Rede ist, scheinen tragfähige Lösungen noch in weiter Ferne zu liegen. Insbesondere lassen praxistaugliche Ergebnisse der Bemühungen um eine *explainable AI* noch auf sich warten, die der Intransparenz geschuldete unzumutbare Stresssituationen für die Nutzer zu vermeiden in der Lage wären.

Dessen ungeachtet verleiten gängige, trotz entgegenstehender Erkenntnisse ständig reproduzierte »KI«-Erzählungen über vermeintlich *lernfähige* oder gar *autonome* Systeme zu folgenreichen Illusionen über deren tatsächliche Leistungsfähigkeit. Sie sind weder *lernfähig*, passen sich allenfalls mittels gegebener Daten algorithmisch gesteuert an äußere Gegebenheiten an, noch *autonom*, also in der Lage, eigene Funktionsregeln zu setzen, sondern sind, wie jedes andere selbsttätige Computersystem auch, per Programm fremdgesteuerte Automaten (oft raf-

finiert ausgedacht, wie Hofstadter bereits (1979: 601) spottete: „AI is whatever hasn't been done yet“). Damit erweisen sich »KI«-Erzählungen als in der Sache unbegründete, durch falsche Begriffsbildung geschaffene Brutstätten gefährlicher und Ressourcen fehlleitender Illusionen (was naiven Rezipienten freilich entgeht).

Für einige in das Geschehen involvierte Akteure sind diese Illusionen jedoch durchaus verlockend: Politiker können sich damit als Förderer von *Modernisierung* profilieren, Forschungsstätten erhalten reichlich Mittel, die dem Nachwuchs viele Promotions-themen bieten, und Unternehmen vermögen sich vorübergehend neue Geschäftsfelder zu erschließen. Das Ergebnis ist allerdings Science Fiction – wirkmächtige Fiktion und miserable *Science*. Und die angesprochenen Probleme lassen eher unerwartet großen Aufwand bei minimalem Ertrag, wenn nicht gar das Menetekel eines erneuten Scheiterns erahnen.

Literatur

- Andelfinger U (1997) Diskursive Anforderungsanalyse. Ein Beitrag zum Reduktionsproblem bei Systementwicklungen in der Informatik, Frankfurt/M: Peter Lang
- Autorengruppe (2018) The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, Oxford (AR): Future of Humanity Institute u. a. 02/2018, <https://arxiv.org/pdf/1802.07228.pdf>
- Barr A, Feigenbaum EA (1981) The Handbook of Artificial Intelligence, Stanford (CA): HeurisTech Press
- Brödner P (2020) Paradoxien der Koaktion von Experten und adaptiven Systemen, in: Brödner P, Fuchs-Kittowski K Hg.: Zukunft der Arbeit – soziotechnische Gestaltung der Arbeitswelt im Zeichen von »Digitalisierung« und »Künstlicher Intelligenz«, Abhandlungen der Leibniz-Sozietät der Wissenschaften Band 67, Berlin: trafo Wissenschaftsverlag, 143-159
- Brödner P (1997) Der überlistete Odysseus. Über das zerrüttete Verhältnis von Menschen und Maschinen, Berlin: edition sigma
- Dörner D (1983) Lohhausen: Vom Umgang mit Unbestimmtheit und Komplexität, Bern: Huber
- Floyd C (2002) Developing and Embedding Autooperational Form, in: Dittich Y, Floyd C, Klischewski R Eds.: Social Thinking – Software Practice, Cambridge (MA): MIT Press, 5-28
- Habermas J (1968) Technik und Wissenschaft als »Ideologie«, Frankfurt/M: Suhrkamp
- Haug WF (2005) Vorlesungen zur Einführung ins »Kapital«, Hamburg: Argument
- High-Level Expert Group on Artificial Intelligence (2018) A Definition of AI: Main Capabilities and Scientific Disciplines, Brussels: European Commission



Peter Brödner

Peter Brödner, Prof. Dr.-Ing., Jahrgang 1942, Studium des Maschinenbaus in Karlsruhe und Berlin. Nach verschiedenen Stationen in Forschung und Projektmanagement auf dem Gebiet computerunterstützter Produktion bis 2005 Forschungsdirektor für Produktionssysteme am Institut Arbeit und Technik im Wissenschaftszentrum Nordrhein-Westfalen. Seither im Ruhestand, Honorarprofessor an der Universität Siegen (Wirtschaftsinformatik), Mitglied der Leibniz-Sozietät der Wissenschaften zu Berlin.

High-Level Expert Group on Artificial Intelligence (2019) Ethics Guidelines for Trustworthy AI, Brussels: European Commission

Hofstadter DR (1979) Gödel, Escher, Bach. An Eternal Golden Braid, New York: Vintage Books

McCarthy J (1955) A Proposal for the Summer Research Project on Artificial Intelligence, <http://www-formal.stanford.edu/jmc/history/dartmouth.pdf>

Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L (2016) The Ethics of Algorithms: Mapping the Debate, *Big Data & Society* 3 (2), 1-21

Nake F (1992) Informatik und die Maschinisierung von Kopfarbeit, in: Wolfgang Coy et al. (Hg.): *Sichtweisen der Informatik*, Braunschweig Wiesbaden: Vieweg, 181-201

Nöth W (2002) Semiotic Machines, *Cybernetics and Human Knowing* 9 (1), 5-22

Peirce CS (1878) Deduction, Induction, and Hypothesis, in: *Collected Papers*, Vol. 2, ed. by Hartshorne C, Weiss P, Cambridge (MA): Harvard University Press (1931-35)

Popper KR (1994) *Alles Leben ist Problemlösen. Über Erkenntnis, Geschichte und Politik*, München. Piper

Rohde M, Brödner P, Stevens G, Betz M, Wulf V (2017) Grounded Design – a Praxeological IS Research Perspective, *Journal of Information Technology* 32 (2), 163-179



Rainer Rehak

The Language Labyrinth: Constructive Critique on the Terminology Used in the AI Discourse

Introduction

In the seventies of the last century, the British physicist and science fiction writer Arthur C. Clarke coined the phrase of any sufficiently advanced technology being indistinguishable from magic – understood here as mystical forces not accessible to reason or science. In his stories Clarke often described technical artefacts such as anti-gravity engines, ‘flowing’ roads or tiny atom-constructing machinery. In some of his stories, nobody knows exactly how those technical objects work or how they have been constructed, they just use them and are happy doing so.

In today’s specialised society with a division of labour, most people also do not understand most of the technology they use. However, this is not a serious problem, since for each technology there are specialists who understand, analyse and improve the products in their field of work – unlike in Clarke’s worlds. But since they are experts in few areas and human lifetime is limited, they are, of course, laypersons or maybe hobbyists in all other areas of technology.

After the first operational universal programmable digital computer – the Z3 – had been invented and built in 1941 in Berlin by Konrad Zuse, the rise of the digital computer towards today’s omnipresence started. In the 1960s, banks, insurances and large administrations began to use computers, police and intelligence agencies followed in the 1970s. Personal computers appeared and around that time newspapers wrote about the upcoming ‘electronic revolution’ in publishing. In the 1980s professional text work started to become digital and in the 1990s the internet was opened to the general public and to commercialisation. The phone system became digital, mobile internet became available and in the mid-2000s smartphones started to spread across the globe (Passig and Scholz 2015).

During the advent of computers, they were solely operated by experts and used for specialised tasks such as batch calculations and book-keeping at large scale. Becoming smaller, cheaper, easier to use and more powerful over time, more and more use cases emerged up to the present situation of computer ubiquity.

More applications, however, also meant more impact on personal lives, commercial activities and even societal change (Coy 1992). The broader and deeper the effects of widespread use of networked digital computers became, the more pressing political decisions about their development and regulation became as well.

The situation today is characterised by non-experts constantly using computers, sometimes not even aware of it, and non-experts making decisions about computer use in business, society and politics – from schools to solar power, from cryptography to cars. The only way to discuss highly complex computer systems and their implications is by analogies, simplifications and metaphors. However, condensing complex topics into understandable, discussable and then decidable bits is difficult in at least two ways. First, one has to deeply understand the subject and second, one has to understand its role and context in the discussion to focus on the relevant aspects. The first difficulty is to do with knowledge and lies in the classical technical expertise of specialists. But the second difficulty concerns what exactly should be explained in what way. Depending on the context of the discussion, certain aspects of the matter have to be explicated using explanations, metaphors and analogies highlighting the relevant technical characteristics and implications. Seen in this light, this problem of metaphors for technology is not only philosophically highly interesting but also politically very relevant. Information technology systems are not used because of their actual technical properties, but because of their assumed functionality, whereas the discussion about the functionality is usually part of the political discourse itself (Morozov 2013).

Given the complexity of current technology, only experts can understand such systems, yet only a small number of them actively and publicly take part in corrective political exchanges about technology. Especially in the field of artificial intelligence (AI) a wild jungle of problematic terms is in use. However, as long as discussions take place among AI specialists those terms function just as domain-specific technical vocabulary and no harm is done. But domain-specific language often diffuses into other fields and then easily loses its context, its specificity and its limitations. In this process terms which might have started as

pragmatic 'weak' metaphors within the technical field, then develop into proper technical terms eventually starting to be seen as 'proper' metaphors outside their original professional context. In addition to the effect of specific terms, those metaphors can also unfold effects beyond concrete technologies but also fuel or inhibit larger narratives around them or digital technology in general. Hence, powerful metaphors push the myths of unlimited potential of (computer) technology, the superiority of computation over human reasoning (Weizenbaum 1976) or the leading role of the 'digital sublime' in transforming society (Mosco 2004). On the other hand, less colourful and less visionary metaphors keep such myths at bay and narratives grounded. Of course, it would be short-sighted to interpret the choice, development and dissemination of technical metaphors and specifically AI terminologies purely as a somehow chaotic process of misunderstandings and unclear technical usage. Those discourses, as all discourses, are a playing field of interests and power where actors brawl over the 'proper' narration either because they find sincere truth in it (e. g., transhumanist zealots of the singularity) or because it plainly benefits them politically or financially (e. g., companies selling AI), or both. Practically speaking, if relevant decision-makers are convinced that AI can develop a real 'understanding' or properly 'interpret' issues, its regular use for sensitive tasks like deciding about social benefits, guiding education, measuring behavioural compliance or judging court cases problematically looms; and corresponding companies will then eagerly come forward to sell such systems to them.

All the above dynamics motivate this work to scrutinise the AI discourse regarding its language and specifically its metaphors. The paper analyses central notions of the AI debate, highlights their problematic consequences and contributes to the debate by proposing more fitting terminology and hereby enabling more fruitful debates.

Conceptual Domains and Everyday Language

Unlike the abstract field of mathematics, where most technical terms are easily spotted as such, AI makes heavy use of anthropomorphisms. Considering AI-terms such as 'recognition', 'learning', 'acting', 'deciding', 'remembering', 'understanding' or even 'intelligence' itself, problems clearly loom across all possible conversations. Of course, many other sciences also use scientific terms that are derived from everyday language. In this case, these terms then have clearly defined meanings or at least linked discourses reflecting upon them. Examples are the terms 'fear' in psychology, 'impulse' in physics, 'will' in philosophy or 'rejection' in geology and 'ideology' in mathematics. Often the same words have completely different meanings in different domains, sometimes even contradictory meanings, as the examples of 'work' in physics and economic theory (energy transfer via application of a force while moving an object vs. planned and purposeful activity of a person to produce goods or services) or 'transparency' in computer science and political science (invisibility vs. visibility) illustrate.

Hence, problems arise when these scientific terms are transferred carelessly into other domains or back into everyday language used in political or public debates. This can occur through

unprofessional science journalism, deliberate inaccuracy for PR purposes, exaggerations for raising third-party funding, or generally due to a lack of sensitivity to the various levels and contexts of metaphors.

The Case of Artificial Intelligence

For some years now, technical solutions utilising artificial intelligence are widely seen as means to tackle many fundamental problems of mankind. From fighting the climate crisis, tackling the problems of ageing societies, reducing global poverty, stopping terror, detecting copyright infringements or curing cancer to improving evidence-based politics, improving predictive police work, local transportation, self-driving cars and even waste removal.

Definitions

The first step towards a meaningful discussion about AI would be to define what exactly one means when talking about AI. Historically there have been two major understandings of AI: strong AI or Artificial General Intelligence (AGI) and weak AI or Artificial Narrow Intelligence (ANI). The goal of AGI is the creation of an artificial human like intelligence, so an AI system with true human-like intelligence including perception, agency, consciousness, intentions and maybe even emotions (see Turing 1950 or more popular Kurzweil 2005). ANI, on the other hand, refers to very domain-specific AI systems being able to accomplish very specific tasks in very narrowly defined contexts only. Questions of agency or consciousness do not arise with ANI systems, they are merely tools, although potentially very powerful tools.

So far and tellingly, AGI can only be found in manifold media products within the fantasy or science fiction genre. Famous examples are Samantha in 'HER', Data in 'Star Trek', HAL 9000 in '2001' (based on a novel of the aforementioned writer Arthur C. Clarke), Bishop in 'Aliens', the Terminator in the movie series of the same name or even the Maschinenmensch in 'Metropolis' (Hermann 2020).

In contrast, ANI systems are the ones calculating the moves in advanced chess games, the ones enhancing smartphone pictures, the ones doing pattern recognition concerning speech (e. g., natural language processing) or images (e. g., computer vision) or even the ones optimising online search results. Furthermore, within the ANI discourse mainly two more specific definitions should be mentioned. The first one focuses on the technical process of how ANI works and goes along the lines of AI being computer algorithms that improve automatically through experience (cf. Mitchell 1997 about machine learning). The second one focuses more on the phenomenon of ANI by defining AI broadly as computer systems that are able to perform tasks normally requiring human intelligence (Gevarter 1985).

Technically there are a multitude of approaches to actually build AI systems. Those approaches are usually referred to as the field of machine learning (ML) and comprise the so-called symbolic approaches with explicit data representations of relevant information like simple decision trees or formal logic-based ones

like knowledge databases with inference algorithms. These approaches are comparatively limited due to the necessity of explicit data representation. Then again there are the more recent sub-symbolic approaches of ML which do not use explicit data representations of relevant information but mathematical (e. g., statistical) methods for processing all kinds of data. Artificial neural networks (ANN) or evolutionary computation are examples of sub-symbolic approaches in ML. Interestingly so far, none of the actual methods available seem to promise a path to AGI.

Yet, despite having at least some general definitions at hand, the common discussion usually ignores those and therefore the range of AI-assigned functionality reaches from applying traditional statistics to using machine learning (ML) techniques up to solely movie inspired ideas or even generally to 'highly complex information systems', as in the official 'Social Principles of Human-centric AI' of Japan (Council for Social Principles of Human-centric AI 2019).

In the following, we will concentrate on artificial neural networks to illustrate the fallacies and pitfalls of questionably used language. The focus on ANN in this text is in line with the current debate of AI, where AI is predominantly used synonymously with machine learning using artificial neural networks (Eberl 2018). Nevertheless, the problems mentioned here also apply to debates concerning other forms of AI, when a similar terminology is being used.

Key drivers for the current AI renaissance are the successes of applying artificial neural networks to huge amounts of data now being available and using new powerful hardware. Although the theoretical foundations of the concepts used were conceived as early as the 1980s, the performance of such a system has improved to such an extent over the last years, that they can now be put to practical use in many new use cases, sometimes even in real-time applications such as image or speech recognition. Especially if huge data sets for training are available, depending on the task results can be much better than traditional symbolic approaches where information is written into databases for explicit knowledge representation.

Before we analyse the language being used to describe the functionality, we should have a look at the inner workings of artificial neural networks to have a base for scrutinising terminologies.

Basic Structure of Artificial Neural Networks

Artificial neural networks are an approach of computer science to solve complex problems that are hard to explicitly formulate, or more concretely: to program. Those networks are inspired by the function of the human brain and its network of neurons; however, the model of a neuron being used is very simplistic. Many details of biological neuronal networks, such as myelination or ageing (Hartline 2009), are left out, as well as new mechanisms, such as backpropagation (Crick 1989), are introduced. Trying to follow the original model, each artificial neuron, the smallest unit of such systems, has several inputs and one output. In each artificial neuron, the inputs are weighted according to its configuration and then summed up. If the result exceeds a

certain defined threshold the neuron is triggered, and a signal is passed on to the output. These neurons are usually formed into 'layers', where each layer's outputs are the next layer's inputs. The resulting artificial neural network thus has as its input the individual inputs of the first layer, and as its output the individual outputs of the last layer. The layers in between are usually called 'hidden' layers and with many hidden layers an artificial neural network is usually called 'deep'.

In the practical example of image recognition, the input would consist of the colour values of all distinct pixels in a given image and the output would be the probability distribution among the predefined set of objects to recognise.

Configuring the Networks

From a computer science point of view, ANNs are very simple algorithms, since the signal paths through the connections of the network can easily be calculated by mathematical equations. After all, it is the variables of this equation (weights, thresholds, etc.) that accord the powerful functionality to ANNs. So ANNs are basically simple programs with a very complex configuration file and there are various ways of configuring artificial neural networks, which will now briefly be described. Building such a network involves certain degrees of freedom and hence decisions, such as the number of artificial neurons, the number of layers, the number and weights of connections between artificial neurons and the specific function determining the trigger behaviour of each artificial neuron. To properly recognise certain patterns in the given data – objects, clusters etc. – all those parameters need to be adjusted to a use case. Usually there are best practices how to initially set it up; then the artificial network has to be further improved step by step. During this process the weights of the connections will be adjusted slightly in each step, until the desired outcome is created, may it be the satisfactory detection of cats in pictures or the clustering of vast data in a useful way. Those training cycles are often done with a lot of labelled data and then repeated until the weights do not change any more. Now it is a configured artificial neural network for the given task in the given domain.

Speaking about the Networks

Now we will take a closer look at how computer scientists speak about this technology in papers and in public, and how those utterances are carried into journalism and furthermore into politics. As mentioned above, the description of ANNs as being inspired by the human brain already implies an analogy which must be critically reflected upon. Commonly used ANNs are usually comparatively simple, both in terms of how the biochemical properties of neurons are modelled but also in the complexity of the networks themselves. A comparison: the human brain consists of some 100 billion neurons while each is connected to 7,000 other neurons on average. ANNs on the other hand are in the magnitude of hundreds or thousands of neurons while each is connected to tens or hundreds of other neurons. This difference in orders of magnitude entails a huge difference in functionality, let alone understanding them as models of the human brain. Even if to this point the difference might only be a matter of

scale and complexity, not principle, we have no indication of that changing anytime soon. Thus, using the notion of 'human cognition' to describe ANN is not only radically oversimplifying, it also opens up the metaphor space to other neighbouring yet misleading concepts. For example, scientists usually do not speak of networks being configured but being 'trained' or doing '(deep) learning'. Along those lines are notions like 'recognition', 'acting', 'discrimination', 'communication', 'memory', 'understanding' and, of course, 'intelligence'.

Considering Human Related Concepts

When we usually speak of 'learning', it is being used as a cognitive and social concept describing humans (or, to be inclusive, intelligent species in general) gaining knowledge as individual learner or as a group, involving other peers, motivations, intentions, teachers or coaches and a cultural background (Bieri 2017). This concept includes the context and a whole range of learning processes being researched, tested and applied in the academic and practical fields of psychology of learning, pedagogy, educational science (Piaget 1944) and neuroscience (Kandel and Hawkins 1995). This is a substantial difference to the manual or automated configuration of an ANN using test sets of data. Seen in this light, the common notion of 'self-learning systems' sounds even more misplaced. This difference in understanding has great implications, since, for example, an ANN would never get bored with its training data and therefore decide to learn something else or simply refuse to cooperate (Weizenbaum 1976); metaphors matter, no Terminator from the movies in sight.

'Recognition' or 'memory' are also very complex concepts in the human realm. Recognising objects or faces requires attention, focus, context and – depending on one's school of thought – even consciousness or emotions. Human recognition is therefore completely different from automatically finding differences of brightness in pictures to determine the shape and class of an expected object (Goodman 1976). Furthermore, consciously remembering something is a highly complex process for humans which is more comparable to living through imagined events again and by that even changing what is being remembered. Human memory is therefore a very lively and dynamic process, and not at all comparable to retrieving accurate copies of stored data bits (Kandel and Hawkins 1995).

Especially the notions of 'action' or even 'agency' are highly problematic when being applied to computers or robots. The move of a computer-controlled robotic arm in a factory should not be called a robot's 'action', just because it would be an 'action' if the arm belonged to a human being. Concerning human actions, very broad and long-lasting discussions at least in philosophy and the social sciences already exist, note the difference between 'behaviour' and 'action' (Searle 1992). The former only focuses on observable movement, whereas the latter also includes questions of intention, meaning, consciousness, teleology, world modelling, emotions, context, culture and much more (Weizenbaum 1976). While a robot or a robotic arm can be described in terms of behavioural observations, its movements should not easily be called actions (Fischer and Ravizza 1998).

Similarly complex is the notion of 'communication' in a human context, since communication surely differs from simply uttering sounds or writing shapes. 'Communication' requires a communication partner, who knows that the symbols used have been chosen explicitly with the understanding that they will be interpreted as deliberate utterances (von Savigny 1983). Communication therefore needs at least the common acknowledgement of the communicational process by the involved parties, in other words an understanding of each other as communicating (Watzlawick 1964). A 'successful' communication is then the result of both parties agreeing that it was successful and therefore the creation of a common understanding. Hence, the sound of a loudspeaker or the text on a screen does not constitute a process of communication in the human sense, even if their consequences are the production of information within the receiving human being. If there is no reflection of the communication partner, no deliberation, no freedom of which symbols to choose and what to communicate one should not easily apply such complex notions as 'communication' outside its scope without explanation.

Furthermore, the concept of 'autonomy' – as opposed to heteronomy or being externally controlled – is widely used nowadays when dealing with artificial intelligence, may it be concerning 'intelligent' cars or 'advanced' weapon systems. Although starting in the last century even human autonomy has been largely criticised within the social sciences (some even say completely deconstructed, Krähnke 2006) since individuals are largely influenced by culture, societal norms and the like, the concept of autonomy seems to gain new traction in the context of computer science. Yet, it is a very simplistic understanding of the original concept (Gerhardt 2002). Systems claimed to be 'autonomous' heavily depend on many factors, e.g., a stable, calculable environment, but also on programming, tuning, training, repairing, refuelling and debugging, which are still traditionally done by humans, often with the help of other technical systems. In effect those systems act according to inputs and surroundings, but they do not 'decide' on something (Kreowski 2018), certainly not as humans do (Bieri 2001). Here again, the system can in principle not contemplate its actions and finally reach the conclusion to stop operating or change its programmed objectives autonomously. Hence, artificial intelligence systems – with or without ANN – might be highly complex systems, but they are neither autonomous nor should responsibility or accountability be attributed to them (Fischer and Ravizza 1998). Here we see one concrete instance of the importance of differentiating between the domain-specific ANI and universal AGI (Rispen 2005). This clarification is not meant to diminish the technical work of all engineers involved in such 'autonomous' systems, it is purely a critique about how to adequately contextualise and talk about such systems and its capabilities in non-expert contexts.

Instances and Consequences

After having briefly touched upon some areas of wrongly used concepts, we can take a look at concrete examples, where such language use specifically matters.

A very interesting and at that time widely discussed example was Google's 'Deep Dream' image recognition and classification

software from 2015, codename 'Inception'. As described above, ANNs do not contain any kind of explicit models; they implicitly have the 'trained' properties distributed within their structures. Some of those structures can be visualised by inserting random data – called 'noise' – instead of actual pictures. In this noise, the ANN then detects patterns exposing its own inner structure. What is interesting are not the results – predominantly psychedelic imagery – but the terminology being used in Google's descriptions and journalists' reports. The name 'Deep Dream' alone is already significant, but also the descriptive phrases 'Inside an artificial brain' and 'Inceptionism' (Mordvintsev and Tyka 2015). Both phrases (deliberately) give free rein to one's imagination. In additional texts provided by Google, wordings such as 'the network makes decisions' accumulate. Further claims are that it 'searches' for the right qualities in pictures, it 'learns like a child' or it even 'interprets' pictures. Using this misleading vocabulary to describe ANNs and similar technical artefacts, one can easily start to hope that they will be able to learn something about the fundamentals of human thinking. Presumably those texts and descriptions have been written for the primary purpose of marketing or public relations, since they explain little but signify the abilities and knowledge of the makers, yet that does not diminish the effect of the language used. For many journalists and executive summary writers or even the interested public those texts are the main source of information, not the hopefully neutral scientific papers. In effect, many of those misleading terms were widely used, expanded on and by that spread right into politician's daily briefings, think-tank working papers and dozens of management magazines, where the readers are usually not aware of the initial meanings. This distorted 'knowledge' then becomes the basis for impactful political, societal and managerial decisions.

Other instances where using wrong concepts and wordings mattered greatly are in car crashes involving automated vehicles, e. g., from companies like Uber, Google or Tesla. For example, in 2018 a Tesla vehicle drove into a parked police car in California, because the driver had activated the 'autopilot' feature and did not pay attention to the road. This crash severely exposed the misnomer. The driver could have read the detailed 'autopilot' manual before invoking such a potentially dangerous feature, yet, if this mode of driving had been called 'assisted driving' instead of 'autopilot', very few people would have expected the car to autonomously drive 'by itself'. So, thinking about a car having an autopilot is quite different from thinking about a car having a functionality its makers call 'autopilot'. Actually reading into Tesla's manuals, different levels of driving assistance are being worked on, e. g., 'Enhanced Autopilot' or 'Full Self-Driving', whereas the latter has not been implemented so far. Further dissecting the existing 'autopilot' feature one finds it comprises different sub-functionalities such as Lane Assist, Collision Avoidance Assist, Speed Assist, Auto High Beam, Traffic Aware Cruise Control or Assisted Lane Changes. This collection of assistance technologies sounds very helpful, yet it does not seem to add up to the proclaimed new level of autonomous driving systems with an autopilot being able to 'independently' drive by itself.

Those examples clearly show how a distinct reality is created by talking about technology in certain terms, yet avoiding others. Choosing the right terms, is not always a matter of life and death, but they certainly pre-structure social and societal negotiations regarding the use of technology.

Malicious Metaphors and Transhumanism

Suddenly we arrive in a situation where metaphors are not only better or worse for explaining specifics of technology, but where specific metaphors are deliberately being used to push certain agendas; in Tesla's case to push a commercial and futurist agenda. Commercial because of using 'autonomy' as a unique selling point for cars and futuristic, as it implies that 'autonomy' is a necessary and objective improvement for everyone's life and the society as a whole. Generally, most innovative products involving 'artificial intelligence' and 'next generation technology' are being communicated as making 'the world a better place', 'humans more empowered' or 'societies more free' by the PR departments of the offering companies and spread even further by willing believers and reporting journalists. The long-standing effects of metaphors let loose can also be seen vividly in the discourse about transhumanism, where humans themselves, even humankind as a whole, should be enhanced and improved using (information) technology, predominantly by using AI. Here again the proponents either really believe in or profit from those narratives, or both (Kurzweil 2005).

In this discourse all mistakes of the AI terminology can be observed fully developed with many consequences, since when we pose the transhumanist question regarding how information technology can help human beings the answer is usually 'enhancement'. Yet the notion of 'enhancement' is being used in a very technical way, ignoring its fundamental multiplicity of meanings. With information technology, so the argument from the classic flavour of transhumanism goes, we will soon be able to fix and update the human operating system: merging with intelligent technical systems will make our brain remember more faces, forget less details, think faster, jump higher, live longer, see more sharply, be awake longer, be stronger, hear more frequencies and even create new senses – exactly how a technologist would imagine what new technology could deliver for humanity (Kurzweil 2005). More recent concepts see humans and AI systems in a cooperative even symbiotic relationship. Those concepts exemplify the direction of imagination once we assume there are truly 'intelligent' systems with 'agency' who can 'decide' and 'act'.

However interestingly the underlying and implicit assumption is a very specific – to be precise: technical – understanding of what is considered 'good' or 'desirable'. But does every human or even the majority primarily want to remember more, forget less, live longer or run faster? Are those aspects even the most pressing issues we want technology to solve? In addition, not only do those fantasies happily follow along the lines of the neo-liberal logic of applying quantification, competition, performance and efficiency into all aspects of life, they also unconsciously mix in masculinist – even militarist – fantasies of power, control, strength and subjugation of the natural or finally correcting the assumed defective (Schmitz and Schinzel 2004).

As valid as those opinions concerning optimisations are, still it is important that views like that imply absolute values and are incompatible with views which put social negotiation, non-mechanistic cultural dynamics or in general pluralistic approaches in their centre. To structure the discourse, I call those conflicting groups of views regimes of enhancement. Clearly it is not pos-

sible to 'enhance' a human being with technically actualised immortality, if this person does not want to live forever or does not find it particularly relevant. Many other conflicting views can be thought of. However, the mere acceptance of the concept of regimes already breaks any claim for absoluteness and opens the door for discussing different understandings of 'enhancements'. Accepting this already makes any positions somehow compatible and allows for individual or even societal endeavours of creative re-interpretations of the concept of transhumanism itself (Haraway 1991).

The transhumanist discourse outlines the consequences of not reflecting on core notions like 'enhancement' in the same way as it is consequential not to reflect on 'intelligence' in the AI discourse (Bonsiepen 1994). Broken visions and faulty applications are to be expected. Furthermore, this kind of language also shapes and attracts a certain kind of mindset where the above mentioned reductionist metaphors are not even used as metaphors anymore, but as accurate descriptions of the world (Coy 1993).

Constructive Wording

So, next time decision-makers and journalists will be asked about possibilities of technology they will surely remember having heard and read about computers winning Chess and Go, driving cars, recognising speech, translating text, managing traffic and generally finding optimal solutions to given problems (Dreyfus 1972). But using deficient anthropomorphisms like 'self-learning', 'autonomous' or 'intelligent' to describe the technical options of solving problems will lead to malicious decisions (The Royal Society 2018, 7–8).

Surely the best solution for this problem would be to completely change the terminology, but since large parts of the above mentioned are fixed scientific terms, a clean slate approach seems unrealistic. Therefore, at least in interdisciplinary work, science journalism activities or political hearings, a focus should be put on choosing the appropriate wording by scientists and (science) journalists. Only then policy and decision-makers have a chance to meaningfully grasp the consequences of their actions. In addition, interdisciplinary research could also get a more solid (communication) ground. Of course, this change of terms will not be the end of discipline-limited jargon in AI, but it would surely increase the efficiency of exchange between the different fields.

For concretely deciding which terms to use and which words to change it would be generally preferable to have some kind of criteria. Following the above descriptions, the used terminology should be as close as possible to the technical actualities while at the same time avoiding:

- technical terms that have a connotation in common language reaching far beyond the actual technical function, e.g., recognition, agent, communication, language, memory, training, senses, etc. since they will be understood as metaphors
- anthropomorphisms which are not technical terms but usually used as metaphors to describe technical details, e.g., thinking, (re)acting, deciding, remembering, etc.
- concepts widely used in popular science, media and science fiction implying a completely different meaning e.g., intelligent machine, android, self-improving, autopilot, etc.

Certainly those words can be replaced by more fitting vocabulary. Depending on context 'remembering' could be paraphrased by 'implicitly stored in configuration', 'learning' by 'changing/improving configuration', 'recognition' by 'detection', 'intelligent' by 'automated' (cf. Butollo 2018), 'action' by 'movement' or 'response', 'decision' or 'judgement' by 'calculation' (cf. Weizenbaum 1976), and 'communication' by 'indication' or 'signalling'. However, terms like 'agency' and 'autonomy' should be discarded entirely, since they are neither accurate or necessary nor helpful; they are just completely misleading.

Being aware that this change might also bear consequences for scientific grant proposals which usually have to sound societally important, scientifically innovative and relevant, it is imperative here too as part of science ethics to reflect on the wider consequences of the language used to communicate. Admittedly it should be noted that those suggestions won't be applied by speakers who are deeply convinced of such metaphors fitting the subject matter, yet, they would then be clearly visible as such.

Closing Remarks

Technology is used and politically decided upon perceived functionality, not upon the actually implemented functionality. However, communicating functionality is much more driven by interests than creating the actual technology. Therefore, attribution ascription is a very delicate and consequential issue that paints



Rainer Rehak

Rainer Rehak beschäftigt sich seit rund zehn Jahren mit dem Themenfeld *Informatik und Gesellschaft*. Er studierte Informatik und Philosophie in Berlin, Hong Kong und Peking. Während des Studiums arbeitete er am Lehrstuhl für *Informatik in Bildung und Gesellschaft* von Wolfgang Coy. Aktuell promoviert er am Weizenbaum-Institut für die vernetzte Gesellschaft und lehrt in den Bereichen Datenschutz/Datensicherheit, sowie Informatik und Gesellschaft.

a differentiated picture of the consequences of careless use of terms. If relevant decision-makers in politics and society are (really) convinced at some point that these 'new' artificial neural networks can develop an understanding of things or properly interpret facts, nothing would stand in the way of their use for socially or politically sensitive tasks like deciding about social benefits, teaching children or judging court cases. Here the difference between 'judging' and judging, 'acting' and acting play out. If one acts in the social science meaning of the word, one has to take responsibility for one's actions, if a computer only 'acts', used as a metaphor, responsibility is blurred.

Hence, especially computer professionals but also scientific journalists should follow the professional responsibility to be more sensitive about the criticised misleading metaphors and in effect to change them to more fitting ones. The danger here does not lie in incompletely understanding computers or AI but in not understanding them while thinking that they have been understood. A possible way out of this tricky situation is certainly more disciplinary openness towards interdisciplinary research and communication. Especially the discipline of computer science could embrace this kind of exchange much more, from student curricula to research projects. This would maybe not so much change their disciplinary core work but it would contextualise this work, create better accessibility for other less technical fields and produce overall more useful results. Naturally, this would require all parties involved to speak to each other but also to listen and teach each other ways of looking at the world. Of course, and not least those ventures must be encouraged and facilitated by field leaders, research grant givers and research politics alike.

So, if technological discussions and societal reflections on the use of technology are to be fruitful, scientists and (science) journalists alike have to stop joining the buzzword-driven language game of commercial actors and AI believers alike, which does neither help with solutions nor advances science. It merely entertains our wishful thinking of how magical technology should shape the future. Finally, we record that a chess computer will never get up and change its profession, exponential growth in computing power does so far not entail more than linear growth of cognitive-like functionality, and the fear of computers eliminating all human jobs is a myth capable of inciting fear since at least 1972.

But maybe, indeed, any sufficiently advanced technology is indistinguishable from magic – to the layperson – but we also have to conclude that this 'magic' is being constructed and used by certain expert 'magicians' to advance their own interests and agendas, or that of their masters (Hermann 2020). So not even such magical interpretation would spare us the necessity to pay attention to power, details and debate (Kitchin 2017). This chapter tries to constructively be a part of this interdisciplinary project.

Der Beitrag erschien im Original als Rehak R (2021) The Language Labyrinth: Constructive Critique on the Terminology Used in the AI Discourse. In: Verdegem P ed. (2021) AI for Everyone? Critical Perspectives. Pp. 87–102. London: University of Westminster Press. DOI: <https://doi.org/10.16997/book55.f>.

References

- Bieri P (2001) *Das Handwerk der Freiheit*. Munich: Hanser.
- Bieri P (2017) *Wie wäre es, gebildet zu sein?* Munich: Komplett Media GmbH.
- Bonsiepen L (1994) Folgen des Marginalen. Zur Technikfolgenabschätzung der KI. In: G. Cyranek and W. Coy (Eds.), *Die maschinelle Kunst des Denkens. Theorie der Informatik*. Braunschweig/Wiesbaden: Vieweg.
- Butollo F (2018) *Automatisierungsdividende und gesellschaftliche Teilhabe*. Regierungsforschung.de, NRW School of Governance. Retrieved from https://regierungsforschung.de/wp-content/uploads/2018/05/23052018_regierungsforschung.de_Butollo_Automatisierungsdividende.pdf
- Council for Social Principles of Human-centric AI (2019) *Social Principles of Human-Centric AI*. Council for Social Principles of Human-centric AI: Japan.
- Coy W (1992) Für eine Theorie der Informatik! In: Coy W et al. Eds., *Sichtweisen der Informatik. Theorie der Informatik*, pp. 17–32. Wiesbaden: Vieweg+Teubner Verlag.
- Coy W (1993) Reduziertes Denken. *Informatik in der Tradition des formalistischen Forschungsprogramms*. *Informatik und Philosophie* 22, 31–52.
- Crick F (1989) The Recent Excitement About Neural Networks. *Nature* 337 (6203), 129–132.
- Dreyfus HL (1972) *What Computers Can't Do*. New York: Harper & Row.
- Eberl U (2018) Was ist Künstliche Intelligenz – was kann sie leisten? *Aus Politik und Zeitgeschichte*, 6–8 (2018), 8–14.
- Fischer JM, Ravizza M (1998) *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Gerhardt V (2002) *Freiheit als Selbstbestimmung*. In: Wobus AN et al (Eds.) *Nova Acta Leopoldina*, Issue 324, Volume 86, Deutsche Akademie der Naturforscher, Leopoldina, Halle (Saale).
- Gevarter WB (1985) *Intelligent Machines: Introductory Perspective on Artificial Intelligence and Robotics*. Englewood Cliffs, NJ: Prentice Hall.
- Goodman N (1976) *Languages of Art: An Approach to a Theory of Symbols*. Indianapolis, MA: Hackett Publishing.
- Haraway D (1991) *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century*. *Simians, Cyborgs and Women: The Reinvention of Nature*, pp. 149–181. New York: Routledge.
- Hartline DK (2009) *What is Myelin?* Cambridge: Cambridge University Press.
- Hermann I (2020) *Künstliche Intelligenz in der Science-Fiction: Mehr Magie als Technik. Von Menschen und Maschinen: Interdisziplinäre Perspektiven auf das Verhältnis von Gesellschaft und Technik in Vergangenheit, Gegenwart und Zukunft*. *Proceedings der 3. Tagung des Nachwuchsnetzwerks 'INSIST', 5–7 October 2018, Karlsruhe (INSIST-Proceedings 3)*.
- Kandel ER, Hawkins RD (1995) *Neuronal Plasticity and Learning*. In: Broadwell RD Ed., *Neuroscience, Memory, and Language. Decade of the Brain*, Vol. 1, S. 45–58. Library of Congress: Washington, DC.
- Kitchin R (2017) *Thinking Critically About and Researching Algorithms*. *Information, Communication & Society* 20 (1), 14–29.
- Krähnke U (2006) *Selbstbestimmung. Zur gesellschaftlichen Konstruktion einer normativen Leitidee*. Weilerswist: Velbrück Verlag.
- Kreowski HJ (2018) *Autonomie in Technischen Systemen*. *Leibniz Online* 32.
- Kurzweil R (2005) *The Singularity Is Near: When Humans Transcend Biology*. London: Penguin.
- Mitchell T (1997) *Machine Learning*. New York: McGraw Hill.
- Mordvintsev A, Tyka M (2015) *Inceptionism: Going Deeper into Neural Networks*. *Google AI Blog*. 17 June 2015. Retrieved from: <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>.
- Morozov E (2013) *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York: PublicAffairs.
- Mosco V (2004) *The Digital Sublime: Myth, Power, and Cyberspace*. Cambridge, MA: MIT Press.

Passig K, Scholz A (2015) Schlamm und Brei und Bits – Warum es die Digitalisierung nicht gibt. Stuttgart: Klett-Cotta Verlag.

Piaget J (1944) Die geistige Entwicklung des Kindes. Zurich: M.S. Metz.

Rispens SI (2005) Machine Reason: A History of Clocks, Computers and Consciousness. Doctoral Thesis, University of Groningen.

Royal Society, The (2018) Portrayals and Perceptions of AI and Why They Matter. London. Retrieved from: http://lcfi.ac.uk/media/uploads/files/AI_Narratives_Report.pdf.

von Savigny E (1983) Zum Begriff der Sprache – Konvention, Bedeutung, Zeichen. Stuttgart: Reclam.

Schmitz S, Schinzel B (2004) Grenzgänge: Genderforschung in Informatik und Naturwissenschaften. Ulrike Helmer Verlag.

Searle JR (1992) The Rediscovery of Mind, Cambridge, MA: MIT Press.

Turing A (1950) Computing Machinery and Intelligence. Mind LIX (236), 433–460. DOI: <https://doi.org/10.1093/mind/LIX.236.433>.

Watzlawick P (1964) An Anthology of Human Communication. Palo Alto, CA: Science and Behavior Books.

Weizenbaum J (1976) Computer Power and Human Reason: From Judgment to Calculation. San Francisco: W. H. Freeman.



Claudia Müller-Birn

Human-Centered Data Science

Etablierung einer kritisch-reflexiven Praxis bei der Entwicklung von datengetriebener Software

Das Forschungsgebiet Data Science hat sich in den letzten Jahren rasant entwickelt und dies stellt die akademische Ausbildung vor besondere (wenn auch nicht neue) Herausforderungen. Die Erfahrungen der letzten Jahre zeigen immer deutlicher, dass der Fokus auf statistische und numerische Aspekte in der Ausbildung nicht ausreicht, um soziale Nuancen, affektive Beziehungen, ethische, wertorientierte Grundsätze oder die tatsächlichen Auswirkungen einer datengetriebenen Software zu erfassen (Aragon et al., 2022). Aus diesen Herausforderungen hat sich das Gebiet Human-Centered Data Science entwickelt, in welchem ein Verständnis für die komplexen Interaktionen zwischen Gesellschaft, Technologie und von Menschen erzeugten Daten vermittelt wird (Aragon et al., 2022). Im Mittelpunkt steht dabei, die herkömmlichen computergestützten Methoden zur Analyse großer Datensätze mit qualitativen Methoden zu verbinden, die mit ihrem Detailreichtum und Kontextwissen zu einem tieferen Verständnis von Daten und Gesellschaft beitragen können. Ein zentrales Anliegen vom Human-Centered Data Science ist es, den Studierenden eine kritisch-reflexive Datenpraxis zu vermitteln. Dieser Artikel soll einen Diskussionsbeitrag liefern, wie eine solche kritisch-reflexive Datenpraxis in der akademischen Ausbildung im Bereich Data Science verankert werden kann.

Human-Centered Data Science – Erweiterung von Data Science durch Qualitative Methoden

Human-Centered Data Science ist ein interdisziplinäres Forschungsgebiet, das sich auf Erkenntnisse und Methoden aus den Bereichen der Mensch-Computer-Interaktion, den Sozialwissenschaften, der Statistik und des maschinellen Lernens stützt (Aragon et al., 2022). Human-Centered Data Science liegt dabei ein menschenzentrierter Ansatz bei der Technologiegestaltung zugrunde, das *human-centered design*, um die Praktiken des Data Science zu verbessern. Dieser Ansatz der menschenzentrierten Gestaltung basiert auf einer Reihe von Leitsätzen, die Kling und Star bereits vor über 20 Jahren aufgestellt haben (Kling & Star, 1998). Danach sollte datengetriebene Software menschliche Fähigkeiten sinnvoll ergänzen, aber diese nicht ersetzen oder automatisieren. Soziale Konstrukte (wie Fairness) sollten nicht in mathematische Konzepte übersetzt werden, da mathematische Operationalisierungen die Vielfältigkeit unserer sozialen Realität nicht umfassend zu beschreiben vermögen. Somit sollte bei der Entwicklung datengetriebener Software nicht nur die Optimierung der statistischen Modelle im Mittelpunkt stehen, sondern auch der Kontext berücksichtigt werden, in den die Software letztlich eingebettet ist. Dies erfordert es, dass auch Fragen der ökologischen Nachhaltigkeit berücksichtigt werden. Kling & Star verweisen bereits darauf, dass Entwickler:innen Bescheidenheit (modesty) in Bezug auf die Fähigkeiten von datengetriebener Software entwickeln sollten, denn Technologie allein kann keine Probleme wie beispielsweise die der sozialen Gerechtigkeit lösen. Daher bildet ein Verständnis über die inklu-

dierten Werte, die der zukünftigen direkt oder indirekt durch den Softwareeinsatz betroffenen Personengruppen und die der Datenwissenschaftler:innen selbst, eine wesentliche Grundlage der menschenzentrierten Gestaltung. Durch einen solchen menschenzentrierten Ansatz bezüglich Data Science können eine Vielzahl von Methoden aus dem Bereich der Mensch-Computer-Interaktion im Data Science verwendet werden, wie beispielsweise die partizipative Gestaltung. Eine grundlegende Voraussetzung in der Vermittlung dieser Methoden ist es aber, die zumeist positivistisch geprägte Haltung von Studierenden um eine kritisch-reflexive Datenpraxis zu ergänzen.

Förderung kritisch-reflexiver Praktiken im Data Science

Eine Erkenntnis aus den Herausforderungen des Einsatzes von datengetriebener Software im gesellschaftlichen Kontext war, dass Informatiker:innen oder Datenwissenschaftler:innen neben der Vermittlung von technischen Fähigkeiten auch im ethischen Denken geschult werden sollten. Gegenwärtige Ethikkurse verfolgen das Ziel, den Studierenden beizubringen, ethische Probleme in der Welt zu erkennen, diese Probleme kritisch zu beurteilen und Technologien in Bezug auf diese Probleme zu bewerten sowie gut begründete Argumente auf der Grundlage der Kritik zu formulieren (Fiesler et al., 2020). Ethik wird in diesen Veranstaltungen aber häufig als statischer, antizipatorischer und formalisierter Prozess operationalisiert. Es werden ethische Theorien (wie Utilitarismus, Deontologie) erläutert und anhand

moralphilosophischer Gedankenexperimente (z. B. Trolley-Problem) diskutiert. Es werden darüber hinaus bestehende ethische Grundsätze (wie Privatsphäre, Rechenschaftspflicht, Transparenz und Erklärbarkeit, Fairness und Nichtdiskriminierung) eingeführt, aber deren konkrete Umsetzung in der jeweiligen Data Science-Praxis bleibt theoretisch. Das hat wahrscheinlich damit zu tun, dass die Ethik-Ausbildung häufig aus Sicht der Geistes- oder Sozialwissenschaft vermittelt wird. Das ist fachlich sicherlich sinnvoll, aber damit ethische Überlegungen auch über den jeweiligen Lehrkontext hinaus nachhaltig wirken können, sollten diese stattdessen in Data Science-Praktiken eingebettet sein.

Es bedarf in der Data Science-Ausbildung einer *Ethics-in-Action* (Handlungsethik) wie sie von Frauenberger et al. formuliert wurde (Frauenberger et al., 2017). Die Handlungsethik ergänzt dabei den durch bestehende ethische Grundsätze (formuliert durch Regierungen, Unternehmen, Fachgesellschaften und Nichtregierungsorganisationen) formalisierten institutionalisierten Ethik-Rahmen, verbindet diesen aber gleichzeitig mit der praktischen Datenarbeit. Die Handlungsethik baut auf Donald Schöns *Reflective Practice* auf, nach der eine Person durch einen gewählten methodischen Ansatz in einen Dialog mit einer Situation tritt. Dieser Dialog „should create a dynamic in which the situation ‚talks back‘, and [the designer] responds to the situation’s back-talk“ (Schön, 1983, S. 79). Schön bezeichnet dieses Gespräch mit einer Situation als reflexiv (ebd.). Im Bereich Data Science kann eine solche *Situation* als konkrete Entscheidung bei der Datenauswahl, der Datenaufbereitung, dem Feature Engineering, etc. angesehen werden. Bei diesen Entscheidungen ist es wichtig, diesen *Dialog* herbeizuführen, der letztlich zu einem reflexiven Handeln befähigt.

Aus der Handlungsethik ergibt sich somit die Notwendigkeit, ethische Überlegungen eng in die Daten- und Programmierpraxis von Datenwissenschaftler:innen einzubetten. Der Dialog kann nur durch bestehende ethische Grundsätze (wie Fairness und Nichtdiskriminierung) initiiert werden, aber das reflexive Handeln selbst muss zu einem inhärenten Teil der Data Science-Praxis werden. Um diese Verbindung herzustellen, wird *ethos* (altgriechisch Charakter) benötigt, „a moral commitment or stance, a moral attitude that underlies a particular practice“ (Frauenberger et al., 2017). Im Gegensatz zu formalen Ethikgrundsätzen ist *ethos* also intrinsisch und personifiziert. Somit ist bei der akademischen Ausbildung ein wichtiges Ziel, dass sich Studierende ihrer Werte bewusst werden. Diese Werte bilden letztlich die Grundlage für die Anwendung der ethischen Grundsätze. Es ist von zentraler Bedeutung, dass Studierende diese eigenen Werte formulieren und danach handeln können.

Basierend auf diesen Überlegungen stellen wir nachfolgend unseren Vorschlag für einen methodischen Ansatz für die Vermittlung von Human-Centered Data Science in der akademischen Ausbildung vor, welche das Ziel verfolgt, die kritisch-reflexive Praxis der Data Science-Studierenden zu fördern, indem die technischen und ethischen Kompetenzen gestärkt werden. Wir verwenden dazu bestehende ethische Grundsätze (beispielsweise Fairness und Nichtdiskriminierung, Transparenz und Erklärbarkeit) und verknüpfen sie mit konkreten Implementierungsaufgaben. Jeder ethische Grundsatz wird anhand einer Fallstudie eingeführt, um Interesse zu wecken. Dieser Fall wird durch Studierende anhand ihrer Werte reflektiert, um Aufgaben

abzuleiten, die dann anhand verfügbarer *Lösungsansätze* umgesetzt und praktisch evaluiert werden. Als konzeptioneller Rahmen dienen dabei Greens Stufen einer *critical technical practice* (Green, 2021). Nachfolgend erläutern wir die Phasen Interesse, Reflexion, Anwendungen und Praxis exemplarisch anhand des ethischen Grundsatzes der Nichtdiskriminierung. Die ersten zwei Phasen werden in der Vorlesung durchlaufen, während die letzten beiden Phasen Teil der Übungsveranstaltung sind. Diese Phasen können innerhalb einer Lehrveranstaltung mehrmals unter Anwendung verschiedener ethischer Grundsätze durchlaufen werden.

Beispiel einer kritisch-reflexiven Praxis anhand des Grundsatzes der Nichtdiskriminierung

In der ersten Phase geht es vor allem darum, Interesse für das Problem (abgeleitet aus dem ethischen Grundsatz) zu wecken. Dazu sollten gesellschaftlich relevante Daten anstelle von banalen Beispieldaten verwendet und konkrete gesellschaftliche Probleme diskutiert werden. Dies zielt darauf ab, das Denken der Studierenden vom einfachen Technologie-Einsatz auf die positive Beeinflussung der Gesellschaft zu lenken (Green, 2021). Hierfür verwenden wir Fallstudien, die eine bestimmte *reale Situation* beschreiben. Die Fallstudien dienen dazu, Studierende nach erfolgtem theoretischem Input (wie im Bereich der Diskriminierung) zu aktivieren. Geeignete Beispiele finden sich in der Sammlung der Gewissensbits (Kurz et al., 2009), die von Mitgliedern der Fachgruppe *Informatik und Ethik* der Gesellschaft für Informatik in den letzten zehn Jahren erstellt wurden.¹ Die Fallstudien decken ein breites Spektrum gesellschaftlicher Themen ab, wodurch Studierende leichter nachvollziehen können, inwieweit eine bestimmte Technologie auch ihr Leben unmittelbar beeinflussen kann.

Die zweite Phase soll die Reflexion fördern, indem Erkenntnisse und Theorien aus Bereichen wie der Wissenschafts- und Technikforschung, Technikphilosophie, Soziologie und Politikwissenschaft bei der Beurteilung einer *Situation* einbezogen werden. Die Situation wird dabei in Daten, Modell, Zielgruppe, Zweck, Kontext, etc. zerlegt, während die erörterten Theorien dazu anregen sollen, über bestehende Annahmen in diesen Bereichen nachzudenken. Durch die Diskussion der Fallstudie wird es Studierenden ermöglicht, über ethische Herausforderungen im Data Science nachzudenken und zu erkennen, dass es oft keine eindeutigen Lösungen gibt. Ziel ist es, Studierende zu befähigen, die eigenen Argumente nachvollziehbar zu formulieren und darzustellen sowie gegenteilige Argumente aufzunehmen und zu bewerten. Ein wesentlicher Ansatz in der Diskussion ist es daher, auf andere Meinungen einzugehen und gemeinsame Lösungen zu suchen. Eine wichtige Erkenntnis in diesen Diskussionen soll es sein, dass bestehende Zielkonflikte beispielsweise bezüglich des Ressourcenverbrauchs, des Sicherheitsniveaus und der Zuverlässigkeit nicht eindeutig lösbar sind. Darüber hinaus sollen Studierende verstehen, dass datengetriebene Software in einen soziotechnischen Kontext eingebettet ist und Zielkonflikte auch dadurch häufig nicht lösbar sind, weil es nicht nur um technische Lösungen geht.

In der dritten Phase konzentrieren wir uns darauf, die Studierenden dabei zu unterstützen, die erlernten Konzepte/Perspek-

tiven in ihre Data Science-Praxis einzubringen. Daher stellen wir Studierenden Datensätze zur Verfügung (wie Kreditdaten²), die sie zur Umsetzung einer konkreten Anwendung (beispielsweise Kreditberatungs-Software) verwenden sollen. Im Themenbereich der Diskriminierung erarbeiten Studierende durch den Einsatz der *Datasheets for Datasets* (Gebu et al., 2021), was sie über den Datensatz wissen und was nicht. Wir diskutieren anschließend, wie durch diese Dokumentation die Nutzbarkeit solcher Datensätze erhöht werden kann. Anschließend untersuchen die Studierenden mögliche Verzerrungen im Datensatz mithilfe des AI360-Toolkit³. Ziel dieser praktischen Übungen ist es immer wieder, dass Studierende lernen, ihre Entscheidungen zu hinterfragen, um so eine kritische Denkweise aufzubauen. Dabei geht es auch darum, bestehende Techniken, Ansätze und Hilfsmittel zu kritisieren, d. h. deren Grenzen zu verstehen und Veränderungsvorschläge zu erarbeiten.

In der letzten Phase sollten die Studierenden in die Lage versetzt werden, eine partizipative Data-Science-Praxis zu realisieren. Dabei lernen sie Methoden (wie partizipatives Design) aus der menschenzentrierten Gestaltung kennen und wenden diese an. Beispielsweise geht es bei der Kreditberatungs-Software nicht nur darum, dass sie ein entsprechendes Vorhersagemodell entwickeln, sondern auch mit den potenziellen Gruppen, die vom Einsatz dieser Anwendung (direkt oder indirekt) betroffen sind, in Kontakt treten und die Auswirkungen ihrer Software verstehen lernen. Dazu werden zunächst Evaluationsstudien durchgeführt, indem die Studierenden die in der vorherigen Phase realisierten Anwendungen (Jupyter Notebooks) kennenlernen und gegenseitig beurteilen. Durch die gegenseitige Evaluation der unterschiedlichen Anwendungen werden die Studierenden ermutigt, die *beste* Anwendung zu küren. Auch diese Aufgabe soll Studierende befähigen, verschiedene Perspektiven auf das Problemfeld einzunehmen und auch potenzielle Konsequenzen zu antizipieren.

Zusammenfassung und Ausblick

Die Veranstaltung *Human-Centered Data Science* ist Teil des Bildungsprogramms *Verantwortungsvolle Informatik*, das an unserem Institut für Informatik der Freien Universität Berlin gerade angelaufen ist. Wir haben die Veranstaltung *Human-Centered Data Science* im Wintersemester 2022/21 erstmalig für Informatik-Studierende an der Freien Universität angeboten. Im Anschluss an diese Lehrveranstaltung führten wir eine Interviewstudie durch, deren Ergebnisse das beschriebene Konzept maßgeblich geleitet haben. Aktuell wird das vorgestellte Kon-

zept praktisch in einer laufenden Lehrveranstaltung evaluiert. Für die Zukunft planen wir, alle Materialien als frei zugängliches Lehrmaterial unter einer offenen Lizenz zur Verfügung zu stellen. Wir möchten aber diesen Artikel mit einem Zitat von Bates et al. (2020) beschließen: „students' critical and ethical thinking is clearly not a cure-all; unethical and socially irresponsible data practice will continue as long as it goes rewarded and poorly regulated.“

Referenzen

- Cecilia Aragon, Shion Guha, Marina Kogan, Michael Muller, and Gina Neff (2022) *Human-Centered Data Science: An Introduction*. MIT Press.
- Jo Bates et al. (2020) Integrating FATE/critical data studies into data science curricula: where are we going and how do we get there? In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. Association for Computing Machinery, 425–435.
- Casey Fiesler, Natalie Garrett, and Nathan Beard (2020) What Do We Teach When We Teach Tech Ethics? A Syllabi Analysis. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. Association for Computing Machinery, 289–295.
- Christopher Frauenberger, Marjo Rauhala, and Geraldine Fitzpatrick (2017) In-Action Ethics. *Interacting with Computers* 29, 2 (Mar 2017), 220–236.
- Timnit Gebu, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford (2021) *Datasheets for datasets*. *Commun. ACM* 64, 12 (Nov 2021), 86–92.
- Ben Green (2021) Data Science as Political Action: Grounding Data Science in a Politics of Justice. *Journal of Social Computing* 2, 3 (Sep 2021), 249–265.
- Rob Kling and Susan L. Star (1998) Human Centered Systems in the Perspective of Organizational and Social Informatics. *SIGCAS Comput. Soc.* 28, 1 (March 1998), 22–29.
- Constanze Kurz, David Zellhöfer, Debora Weber-Wulff, Christina Class and Wolfgang Coy (2009) *Gewissensbisse: Ethische Probleme der Informatik. Biometrie – Datenschutz – geistiges Eigentum*.
- Donald A. Schön (1983) *The reflective practitioner: How professionals think in action*. Basic books.

Anmerkungen

- Gewissensbits werden regelmäßig in einer Rubrik der Fachzeitschrift Informatik Spektrum von Springer veröffentlicht. Weitere Informationen finden Sie unter <http://gewissensbits.gi.de>.*
- Der Datensatz ist verfügbar unter <https://archive.ics.uci.edu/ml/datasets/Statlog+%28German+Credit+Data%29>.*
- Weitere Informationen sind hier verfügbar: <https://github.com/Trusted-AI/AIF360>.*



Claudia Müller-Birn



Claudia Müller-Birn leitet die Forschungsgruppe *Human-Centered Computing* am Institut für Informatik der Freien Universität Berlin. In ihrer Forschung untersucht sie die Verflechtung von Menschen, Daten, Algorithmen und Kontext, um eine Human-Computer-Collaboration zu ermöglichen. Ihr aktueller Anwendungsschwerpunkt liegt auf Technologien des maschinellen Lernens im Zusammenhang mit Privatsphäre, Reflexion und Interpretierbarkeit.

Je künstlicher die Intelligenz, desto weniger intelligent

KI erobert unseren Alltag

Anfangen von komplexen Spielstrategien bei Schach oder Go, Sprachassistenten wie Siri, Cortana und Alexa, dem Erkennen sowie Klonen von Stimmen und Gesichtern, automatenbasierter telefonischer Kommunikation, Expertensystemen mit Wissensdatenbanken, automatischen Entscheidungssystemen im Gesundheitswesen, in der Justiz, Produktionssteuerung, persönliche Optimierungsassistenten auf dem Smart Phone über Industrie-Roboter (Fertigung, Logistik, Pflege) bis zu autonomen Waffen: KI-Anwendungen treten weit in das alltägliche Leben ein, machen neue Geschäftsmodelle (wie selbstfahrende Autos, unbemannte Kassen in Supermärkten, personalisiertes Marketing oder Prognosen des Kaufverhaltens) möglich und ersetzen dabei den Menschen mit seinen Qualifikationen.

KI-unterstützte Suchmaschinen helfen uns, die relevanten Inhalte im Netz zu finden. Empfehlungsdienste bringen bei Amazon, Spotify oder Netflix „personalisierte Angebote“ oder in sozialen Medien die wichtigsten Neuigkeiten nach oben, moderieren Nachrichtenforen und steuern das Recruiting von Firmen. Ihre Apologeten versprechen durch Förderung, Entwicklung und Anwendung von KI-Technologien die Lösung aller globalen Probleme der Menschheit: Reichtum und Fortschritt, Lösungen der Klimakrise und einen wirtschaftlichen Wachstumstreiber.

„Künstliche Intelligenz ist eines der wichtigsten Dinge, an denen Menschen arbeiten. Ihre Bedeutung ist grundlegender als Elektrizität oder das Feuer“, so Google-Chef Sundar Pichai 2018 auf einer Veranstaltung in San Francisco. Und Microsoft-Chef Satya Nadella: „Künstliche Intelligenz ist nicht einfach nur eine weitere Technologie, es könnte eine der wirklich grundlegenden Technologien sein, die Menschen jemals entwickelt haben.“¹ Jürgen Schmidhuber freut sich über sich selbst reproduzierende KI, die den Menschen ignoriert und die Milchstraße erobert: Mit der KI hebe sich „nicht nur dieser kleine Planet Erde, sondern in der Folge davon das gesamte Universum auf eine höhere Komplexitätsstufe. Denn die KIs werden natürlich merken, dass fast alle physikalischen Ressourcen weit draußen sind. Und das wird zwangsläufig dazu führen, dass viele KIs expandieren und sich ausbreiten werden.“²

Aufzuhalten sei dieser Prozess nicht mehr, aber man könne sich ja anfreunden mit der Idee, „dass man Steigbügelhalter des Universums auf seinem Weg zu immer höherer Komplexität ist. Man kann Schönheit und Erhabenheit darin finden, dass man Teil eines gewaltigen Prozesses ist, der den gesamten Kosmos umgestalten wird.“³ KI könne sogar heute noch nicht definierte zukünftige Herausforderungen identifizieren, selbständig Ziele definieren, Entscheidungen treffen und eigenständig handeln. Hilfreicher Boom vieler sinnvoller Anwendungen oder gefährlicher Hype?

Versuch einer Definition

„Als Künstliche Intelligenz bezeichnet man traditionell ein Teilgebiet der Informatik, das sich mit der Automatisierung von intelligentem Verhalten befasst.“ (Kirste, Schürholz). Ein solcher Definitionsversuch ist selbstreferentiell, weil er sich auf die Intelligenz, also die kognitive Leistungsfähigkeit des Menschen bezieht, die die KI beansprucht, besser verstehen und zugleich automatisieren zu wollen. So verfolgt die KI zugleich ingenieurs- und kognitionswissenschaftliche Ziele (Wolfgang Wahlster).



Self checkout im Supermarkt, aufgenommen 2013 in Japan
Foto: Karl Baron, CC BY 2.0

Erstmals geprägt wurde der Begriff *Künstliche Intelligenz* 1956 von John McCarthy auf der Dartmouth Conference. Seither ist er ein unscharfer Sammelbegriff für fortgeschrittene IT-Technologien und Anwendungen, die seit den 1940er-Jahren verschiedene Höhen und Tiefen durchlaufen haben.⁴ Die Anfänge gehen auf Spieltheorie, klassische Prinzipien der mathematischen Logik, Repräsentations- und Regelsysteme sowie Entscheidungsbäume zurück. Dabei wird jede auftretende Situation mit ihren Folgealternativen (zum Beispiel mögliche Spielzüge für einzelne Schachfiguren) dargestellt und durch Heuristiken bewertet, so dass nur die mit großer Wahrscheinlichkeit besten Züge weiterverfolgt werden. Dadurch konnten größere Erfolge bei regelbasierten Lösungen erzielt werden.

In den 70er-Jahren gab es kleinere Erfolge mit Expertensystemen für eingeschränkte Anwendungen, die allerdings einen immensen Aufwand bei der Erfassung und Darstellung menschlichen Wissens in Wissenssystemen voraussetzten.⁵ Für bereits weit entwickelte Modelle fehlten damals die hardwaretechnischen und datenmäßigen Voraussetzungen (schnelle Mehrfach-Prozessoren, effiziente Algorithmen, Miniaturisierung, große Speicherkapazitäten, Big Data Strategien, Internet der Dinge), so dass die KI erst wieder nach einer 20jährigen Stagnationsphase zum neuen Jahrtausend mit Erfolgen ans Licht trat.

Zentraler Ansatz war jetzt das sogenannte *Maschinelle Lernen*. Dabei wird das IT-System weiterhin von einem statischen Programm mit Entscheidungsregeln gesteuert, die zu jedem Input

(Problem) einen klaren Output (Lösung) generieren. Zusätzlich konfrontiert man das System in einer notwendigen Lernphase mit einem Feedback der richtigen Lösung für gegebene Inputs, so dass sich die Entscheidungsregeln anpassen, also *lernen* können.

Die am weitesten entwickelte Form ist das *Deep Learning*, bei dem nicht nur das Input-Output-Verhalten ständig verbessert, sondern im *Lernen* auch die internen Repräsentationsstrukturen angepasst werden. Dabei bedient man sich der Analogie biologischer neuronaler Netze des Menschen.⁶ Algorithmen sollen diese mehrschichtigen Netzstrukturen als *künstliche neuronale Netze* (KNN) dergestalt abbilden, dass Eingangssignale (Input) als elektrische Impulse durch das Netz bis zu den Ausgängen geleitet werden, deren Werte dann wiederum den Output ausmachen.

Dabei können sich die Gewichtungen der Netzverknüpfungen bei den einzelnen Neuronen verändern, in diesem Fall durch die angebotene Rückkopplung (Feedback) des richtigen Ergebnisses in der Lernphase. Mit den neuen Gewichtungen wird das Netz mit neuen Eingabewerten durchlaufen. Das Modell der KNN wendet man auf vielseitige Problemstellungen an, wie in der Mustererkennung.

In die Eingangsstellen können Bild(farb)werte eingehen und der Ausgabewert ist eine Aussage, ob es sich beispielsweise um ein Auto handelt oder nicht. Heute sind KI Systeme hybride Systeme, verbinden die symbolische Ebene (*Wissen*: semantische Netze, Regeln und Modelle) mit der subsymbolischen Ebene (*Lernen*: Stützvektoren, KNN und Genetische Algorithmen/heuristische Optimierungsverfahren).

Das KI-Selbstverständnis: Intelligenz wie beim Menschen!

KI-Vertreter setzen *Künstliche* mit *menschlicher* Intelligenz gleich. Schon die Begriffe Intelligenz, Lernen und neuronale Netze suggerieren, der Mensch arbeite wie ein Computer und umgekehrt. Diese Sicht wird bereits von frühen KI-Forschern vertreten. So betrachtet man „den Denkapparat nicht nur als biologische Ingenieursleistung, sondern als ein System mit Hardware, Compiler und Programmen“ (Elithorn/Jones 1973).

Die Künstliche Intelligenz „begriff den Menschen als Informationsverarbeiter. Sie ist gerade an den inneren Vorgängen, als Verarbeitung von Information betrachtet, interessiert und sucht nach Modellen für Automation, die dasselbe leisten.“⁷ In der Turingthese wird die Frage, wann ein Computer denken kann, durch ein Imitationsspiel beantwortet: wenn eine Maschine einen Menschen so gut imitiert, dass ein Kommunikationspartner am Ende einer Telefonleitung beide nicht voneinander unterscheiden kann. Dazu sollen Maschinen bereits in 2000 in der Lage sein.⁸

Das Ziel der KI-Akteure Feigenbaum/Feldman war „programming of computers to perform intellectual tasks in the same way that persons perform these tasks.“⁹ Geht es hier noch um Systeme, die in ihrer *Performanz* dem Menschen gleichen sollen, so postulieren andere Autoren gleiche interne Strukturen für die Rechenmaschine und das menschliche Gehirn: „Der Speicher des Systems ist wie das menschliche Gedächtnis aufgebaut“ (Collins/Warnock/Aiello/Miller 1975).

Auch Autoren, die semantische Netze für die Speicherung im Computer verwenden, bezeichnen diese als Modell der Gedächtnisstruktur im Menschen (Quillians, Simmons und Scragg¹⁰). Schank argumentiert, dass es eine kanonische Form der Repräsentation von Sätzen im menschlichen Gedächtnis geben muss; ein „seriöser Kandidat“ dafür sei sein Computermodell der „conceptual dependency“ (Schank 1976). Diese Zitate aus der Frühphase der Künstlichen Intelligenz lassen sich durch neuere Akteursäußerungen der Branche ergänzen.¹¹

Über bisherige KI-Systeme für begrenzte Anwendungen (*schwache KI*) wie spezifische Experten- oder Entscheidungssysteme hinaus will man eine sogenannte *starke KI* oder auch *allgemeine KI* entwickeln. Damit meinen Shane Legg, Mark Gubrud und Ben Goertzel die Fähigkeit, jedes beliebige Ziel mindestens so gut wie Menschen zu erreichen.

Innerhalb der Grenzen von schwacher KI können Computer eine Menge gedanklicher Operationen schneller und exakter als der Mensch ausführen. Operationen können formalisiert, modelliert, algorithmisiert und einem Computer übertragen werden, die menschliche Kognition als Prozess der Informationsverarbeitung nachzeichnen.

Die starke KI beansprucht einen *allgemeinen Geist* zu realisieren und kognitive Prozesse wie beim Menschen zu imitieren. Bei einer starken KI mit offenen Fragestellungen müsste die Maschine jedoch über ein nahezu universelles Wissen (Weltformell!) verfügen oder eine Optimierungsfunktion bezüglich des Weltwissens benutzen, die sowohl neues Wissen erschließt als auch zur Strategiebildung gebraucht wird.

Diese Sichtweise hat zwei Implikationen zwischen dem Selbstbild der KI und dem zugrundeliegenden Menschenbild: KI wird anthropomorphisiert als menschliche Intelligenz, und menschliche Intelligenz wird reduziert auf Berechenbarkeit. Was sind das für besondere Eigenschaften der KI, die der Auffassung vom Menschen als Maschine zu neuer Plausibilität verholfen haben, so kritisiert Weizenbaum schon 1975 zu Recht. Welf Schröter (2019) will das amerikanisch-englische Wort *artificial intelligence* demnach nicht *Künstliche Intelligenz*, sondern besser *Künstliche Nachbildung* übersetzen.

Was ist menschliche Intelligenz?

Bisher sind alle verwandten Wissenschaften – Biologie, Kognitionswissenschaften, Neurowissenschaften und Psychologie – an einer Definition gescheitert.

Der Brockhaus definiert Intelligenz als Fähigkeit der Auffassungsgabe, des Begreifens, Urteilens; als geistige Anpassungsfähigkeit an neue Aufgaben, während Meyers enzyklopädisches Lexikon Intelligenz als die Bewältigung neuartiger Situationen durch problemlösendes Verhalten versteht. Im Vorhandensein einer Ratio erschöpft sich jedoch bei weitem nicht der Begriff der Intelligenz: „Zur Intelligenz gehört mehr als nur kalkulierbarer Verstand“ (Dreyfus/Dreyfus, 1988: S. 61). Intelligenz ist „die Fähigkeit zu problemlösendem, einsichtigem Verhalten“ (Cruse/Dean/Ritter, 1998: S. 9).

Ein intelligentes System sollte 1. autonom sein, d. h. sein Verhalten weitgehend selbst bestimmen, mindestens eine Unterscheidung zwischen eigen- und fremdbestimmtem Verhalten treffen können, 2. „Intentionen besitzen, sich selbst die Ziele seines Verhaltens auszuwählen“, sich anpassen und aus Erfahrungen lernen und den Erfolg eines Verhaltens beurteilen zu können.

Weitere wichtige Eigenschaften sind die „Fähigkeit zur Generalisierung“ und „die Fähigkeit, zwischen Alternativen entscheiden zu können. Die vielleicht wichtigste Bedingung für das Auftreten von Intelligenz besteht in der Fähigkeit, Änderungen der Umwelt, z. B. als Folge eigener Aktivität vorhersagen zu können.“ Ein System ist also dann intelligent, „wenn es in einer gegebenen und sich verändernden Umwelt die Chancen seiner Selbsterhaltung im Vergleich zu seinem aktuellen Zustand verbessern kann“ (alles zit. Cruse/Dean/Ritter, 1998).

Menschen besitzen ein eigenes veränderbares Wertesystem, nach dem sie zielgerichtet handeln und denken. Auch Lernen erfolgt zielgerichtet mit der Tendenz, Probleme zwecks menschlicher Bedürfnisbefriedigung zu lösen. Probleme in diesem Sinne kennt das KI-System nicht; es kann folglich auch nicht zielgerichtet lernen.

Menschliches Denken ist auf der Grundlage der biologischen Entwicklung und unter dem Antrieb der gesellschaftlichen Arbeit der Menschen entstanden. Als Resultat gesellschaftlicher Entwicklung ist Denken somit auch veränderbar. Es verarbeitet die Sinneserfahrung und hebt durch Operationen wie *Abstraktion* und *Verallgemeinerung* das Wesentliche und Gesetzmäßige heraus; es erzeugt dadurch ideelle Modelle.

Das Denken ist schon seiner Entstehung nach sozial bedingt; es geht aus der praktischen Aneignung der Umwelt durch die Menschen hervor, indem die ursprünglich äußere Erkenntnistätigkeit des direkten Operierens mit Gegenständen langsam in die innere Tätigkeit des Operierens mit Begriffen und sprachlichen Zeichen übergeht.

Im Prozess der aktiven Auseinandersetzung des Menschen mit seiner natürlichen und gesellschaftlichen Umgebung verändert sich Denken ständig: es ist zweckgerichtet und folgt wie das Handeln einem Wertesystem. Die soziale Bedingtheit des Denkens in seinen verschiedenen Aspekten macht es zu einer abschließlichen Eigenschaft des Menschen.

Kritik

Kritik an der KI wird dahingehend geführt, dass der Mensch über besondere Denkfähigkeiten verfüge, die von Computern nie erreicht würden (Dreyfus 1979). Der Begriff der rationalen Intelligenz bilde für künstliche Intelligenz eine wichtige Begrenzung. Digitalcomputer sind symbolverarbeitende Systeme, die den Gesetzen von Logik und Ratio folgen. In dieser sehr eingeschränkten Sicht liege das Problem.

Ist ein System wirklich als intelligent zu bezeichnen, wenn es sich nur auf künstliche Problemstellungen beschränkt? Auch der Turing-Test erkläre nicht, was Künstliche Intelligenz ist, sondern nur, womit sich *Künstliche Intelligenz* beschäftige. Weizenbaum

(1975) kritisiert, dass mit der KI-Verbreitung ein bestimmter Intelligenzbegriff befördert wird, der von allen konkreten Formen des Denkens und Handelns abstrahiert und sämtliche individuellen Fähigkeiten auf ein einziges Maß reduziert, zurück bleibt ein abstraktes mechanistisches Phänomen.

Die Maschinerie im Produktionsprozess

Prägend für die Rolle, Rezeption und gesellschaftliche Verortung der KI sind aber nicht wissenschaftliche Diskussionen über Menschenbilder, sondern die alltäglichen arbeitsweltlichen Erfahrungen: die Macht, die die Menschheit durch Naturwissenschaft und Technik erworben hat, verkehrt sich bei den Arbeitenden in Ohnmacht.

In einer Untersuchung der Alltagsarbeit in den USA schreibt Studs Terkel¹²:

Bei den meisten zeigt sich kaum verhohlene Unzufriedenheit. ... „Ich bin eine Maschine“, sagt der Punktschweißer. „Ich bin in einen Käfig eingesperrt“, sagt der Bankbeamte, der damit nur dem Hotelportier aus dem Munde spricht. „Ich bin ein Packesel“, sagt der Stahlarbeiter. „Meine Arbeit könnte genauso gut ein Affe machen“, sagt die Empfangsdame. „Ich bin weniger als ein landwirtschaftlicher Zubehörsartikel“, sagt der Wanderarbeiter. „Ich bin ein Objekt“, sagt das Mannequin. Ob sie im blauen oder im weißen Kittel arbeiten, bei allen kommt es auf dasselbe heraus: „Ich bin ein Roboter.“

Die mit der Mechanisierung verbundene Ablösung der intellektuellen Fähigkeiten von der Arbeitskraft bewirkt mit ihrer Dequalifizierung auch ihre Entwertung. Durch die Unterordnung unter die Maschine lässt sich die Arbeit durch bloße Beschleunigung der Maschinengeschwindigkeit intensivieren, der Arbeiter dient der Maschine. Die verloren gegangenen Kenntnisse über Produktionsabläufe werden als technologisch angewandte Wissenschaft in die Maschine inkorporiert. Die Wissenschaft wird zum selbständigen, vom Arbeiter getrennten Moment des Arbeitsprozesses und tritt ihm als fremde, ihn beherrschende Macht gegenüber, als Macht der Maschine selbst.¹³

Bei der Automatisierung von Arbeitsprozessen¹⁴ sinken Dispositionsspielraum und Einfluss auf die Arbeitsgestaltung erheblich, womit die Distanz vom Menschen zum Arbeitsprodukt weiter vergrößert wird. Technisch-organisatorische Veränderungen durch digitale Arbeit sollen die Beschäftigten überwachen, bewerten und steuern. Diese Erfahrungen machen Kollegen nicht nur mit Industrie 4.0, auch in den Verkaufs- und Serviceabteilungen (Call Centern), im stationären Einzelhandel, in Versicherungen oder etwa Banken.

In Bereichen mit Kundenkontakt haben die Beschäftigten keinen Einfluss auf die Entscheidung, welche Arbeitsvorgänge sie übernehmen. KI-gesteuerte *Chatbots* beraten Anrufer und Internetnutzer und leiten an Beschäftigte weiter, wenn sie nicht weiterwissen.

Die eingehende Arbeit wird automatisiert durch Workflows in persönliche Arbeitskörbe zugeteilt. Über das Monitoring wer-

den Beschäftigte sowie Kunden ausgespäht, Kundenkontakte dokumentiert, durch das Kundenbeziehungsmanagement nachverfolgt und ausgewertet. Die Aufgaben von Führungskräften werden auf Kennzahlen verengt, um die Beschäftigten auf die Ergebnisse einzuschwören. Gemessen werden etwa die Bearbeitungsdauer, Gesprächsdauer, Wartezeiten, Antwortzeiten, Prozessdurchlaufzeiten oder Servicelevel. Auf dieser Basis werden die Prozesse standardisiert und durch Zeitvorgaben *optimiert*.

Mit der Flexibilisierung des Arbeitskräfte-Managements durch KI-Einsatz gilt es, „Luft in den Prozessen“ zu finden. Mit Hilfe statistischer Erhebungen und Vorhersagen des Arbeitsanfalls und Kundenverhaltens sollen Personalkapazitäten, Dienstpläne und die Verteilung der Arbeitszeiten bis hin zur Lage der Pausen gesteuert werden.¹⁵ KI entscheidet über Leistungsbewertung und Karriereentwicklung. Diese Leistungen der mathematisierenden Informationsverarbeitung sind: von Menschen gemacht, von Menschen gestaltbar.

Neben Software als starres unverändertes Werkzeug tritt Software als ein sich selbst veränderndes Werkzeug. Ihr Charakter entwickelt sich von der „Assistenztechnik“ hin zur „Delegationstechnik“ (Schröter). Somit kann eine Software die Vollmacht bekommen, hinter dem Rücken des Menschen in Echtzeit rechtsverbindliche Prozesse (Transaktionen) zu analysieren, zu steuern und zu entscheiden. Aber auch diese „Delegationstechnik“ *denkt nicht, lernt nicht, hat kein Ich*. Es bleibt Mathematik.

Wider die instrumentelle Vernunft!

Damit markiert die Künstliche Intelligenz den vorläufigen Höhepunkt einer Entwicklung, die mit der Aufklärung begann. Mit ihr sollten durch rationales Denken alle den Fortschritt behindernden Strukturen (überholte Vorstellungen, Ideologien, Widerstand von Tradition und Gewohnheitsrecht) überwunden, Akzeptanz für neu erlangtes Wissen geschaffen und in der Berufung auf die Vernunft als universelle Urteilsinstanz die Hauptprobleme menschlichen Zusammenlebens gelöst werden.

Heute ist die Aufklärung in die Krise geraten: Mit „Instrumenteller Vernunft“ bezeichnet Max Horkheimer die Dominanz einer technisch-rationalen Vernunft, die sich mit gesellschaftlicher Herrschaft verschwistert habe, über die praktische Vernunft. Sie steht für eine Vernunft, welche die Mittel, nicht jedoch die Ziele des Handelns reflektiert. Darüber hinaus bezeichnet sie das In-

teresse an technischer Beherrschung und Unterwerfung der Natur. Man spricht auch von Zweck-Mittel-Rationalität, womit die technisch-rationale Angemessenheit der Mittel zur Erreichung eines beliebig gewählten Zweckes bezeichnet wird. Die Zwecke selbst können unvernünftig, ja irrational sein, während die Mittel zu ihrer Durchführung rational und technisch effektiv funktionieren.

Die Kritik der instrumentellen Vernunft ist auch eine Kritik an Naturbeherrschung, also am instrumentellen Verhältnis zur Natur. Die Natur werde heute „als ein bloßes Werkzeug des Menschen“ aufgefasst und sei „Objekt totaler Ausbeutung“. Die Unterdrückung der (inneren wie äußeren) Natur und intrahumane Herrschafts- und Unterdrückungsformen stehen im Zusammenhang; die Geschichte der Anstrengungen des Menschen, die Natur zu unterjochen, ist auch die Geschichte der Unterjochung des Menschen durch den Menschen selbst.

Der Mensch sieht sich als „das andere zur Natur“ (Precht). Der Versuch, die Menschen zu Göttern zu erheben, führt den Humanismus zu seinem logischen Schluss und macht die in ihm innewohnenden Schwächen deutlich.¹⁶ Mit dem Versuch, die Natur zu beherrschen, wird der einst mythische Zugang zur Welt aufgeklärt und schlägt als „Herrschaft“ und „neue Barbarei“ dialektisch in Mythos zurück. An die Stelle der Aufklärung und der Befreiung von den Zwängen der Natur tritt die Unterordnung unter wirtschaftliche und politische Interessen.

Die Arbeitenden sind trotz der ihnen eigenen Potenzen in der Arbeit auf Anhängsel von Maschinen reduziert und werden auf ein Funktionieren eingeschränkt. Solcherart reduzierte Menschen sind offensichtlich durch Maschinen ersetzbar. Dieses Empfinden gegenüber IT wird dadurch verursacht, dass Arbeitnehmer/Nutzer keinen Einfluss auf die Einführungsbedingungen haben, dass Wissen und Fertigkeiten sowie Kenntnisse über Zusammenhänge und innerbetriebliche Arbeitsabläufe im Computer verschwinden. Zurück bleibt eine relative Undurchschaubarkeit der Abläufe und Prozesse.

„In allen Bereichen wird der Mensch nach und nach durch die Maschine ersetzt, nicht weil die Maschine diese Dinge besser bewältigt, sondern eher weil alle Dinge auf das Maß reduziert worden sind, das die Maschine zu bewältigen imstande ist“ (Müller 1982). So wie der Mensch seine Autonomie in der kapitalistischen und an die kapitalistische Produktion verloren hat, die als nicht steuerbare Maschinerie mit einem Eigenleben betrachtet wird und die Menschen als Arbeitskräfte ein- und aussaugt und so wie die natürlichen Ressourcen benutzt, so ist er nun bereit,



Klaus Heß

Klaus Heß, Dipl. Inf., 30 Jahre Betriebsräteberater für Arbeits- und Technikgestaltung bei der TBS NRW, ist FfF-Gründungsmitglied. Er hat in den Schwerpunkteditionen zu *Gute Arbeit ... mit und trotz IT* (FK 3/2010) und *Zukunft der Arbeit – Arbeit der Zukunft: Wer steuert wen?* (FK 4/2016) mitgearbeitet. Nebenberuflich engagiert er sich in internationalistischer Solidaritätsarbeit in Mittelamerika.

seine Autonomie auch an KI als Verlängerung seiner intellektuellen, kognitiven und seelischen Fähigkeiten abzugeben.

Während er Glauben und Vertrauen an seine eigene Autonomie verloren hat, will er sich auf autonome Maschinen verlassen. Soll dieser Verlass auf etwas anderem beruhen als blinder Hoffnungslosigkeit? Nobelpreisträger Herbert Simon (1977) befindet: „Die wahrscheinlich wichtigste Frage über den Computer ist, was er mit dem menschlichen Selbstverständnis und seinem Platz im Universum getan hat und weiterhin tun wird.“ Der Werkzeugcharakter des Computers ist erst wieder in einer selbstbestimmten Produktion herstellbar, in der die Menschen nicht nur gesellschaftlich über die Produktpalette, sondern betrieblich über die Arbeitsprozesse und am Arbeitsplatz über den Technikeinsatz entscheiden.

Es kommt also darauf an, das Verhältnis zwischen Menschen und Maschinen neu zu bestimmen. Statt weiterhin von KI zu reden, „geht es um die soziale Gestaltung „algorithmischer Steuerungs- und Entscheidungssysteme“, um die Gestaltung „algorithmischer, sich selbst verändernder Software-Systeme für die arbeitsweltlichen Vorgänge Analyse, Entscheidungsvorbereitung, Prozesssteuerung und Entscheidungsvollzug im Feld der Assistenz- und Delegationstechnik“ (Schröter).¹⁷

Hierzu fehlt bisher jegliches Mitbestimmungsinstrumentarium, um die arbeitenden Menschen in der Mensch-Computer-Interaktion zu Subjekten werden zu lassen. Mit der Stellung von Mensch und Maschine in der Produktion löst sich auch die Frage des Menschenbildes und der Begriffskritik von alleine.

Anmerkungen

- 1 Zitiert nach dem Wall Street Journal
- 2 Süddeutsche Zeitung, Oktober 2018: Künstliche Intelligenz: Jürgen Schmidhuber im Interview – Digital – SZ.de (sueddeutsche.de)
- 3 Jürgen Schmidhuber in: Publik Forum. Digitales Leben. März 2021 S.26-28
- 4 Maschinelles Lernen. Eine Analyse zu Kompetenzen, Forschung und Anwendung. Fraunhofer Gesellschaft 2018
- 5 Der Autor war selbst als Informatiker an einem Repräsentationssystem beteiligt, s. Heß: Handlungen und ihre Folgen – Zum Erschließen von Sachverhalten aus Handlungsabläufen, in SPRACHE UND DATEN-VERARBEITUNG 1/1980
- 6 Marvin Minskys SNARC von 1952 (Stochastic Neural Analog Reinforcement Calculator) war die erste selbstlernende, auf künstlichen neuronalen Netzwerken und Wahrscheinlichkeitsrechnung basierende Rechenmaschine und benutzte Vakuumröhren.
- 7 Seminar Künstliche Intelligenz, Universität Bonn 1979/80
- 8 Nach Turing 1950. Nach Tegmark testet der Turing-Test eher die menschliche Leichtgläubigkeit als die Künstliche Intelligenz
- 9 Computers and thought by Feigenbaum, Edward A., ed; 1963, S. 269ff
- 10 „Semantische Netze“ als „interlingua“ für maschinelle Übersetzung natürlicher Sprachen wurden von Robert F. Simmons und M. Ross Quillian in den frühen 1960er-Jahren entwickelt.
- 11 Neuere Forschungen gehen allerdings von solch hyperkomplexen Gehirnstrukturen aus, dass sie kaum in IT-Systemen nachgebaut werden könnten (Tegmark, Max: Leben 3.0. Mensch sein im Zeitalter Künstlicher Intelligenz, Berlin 2017)

- 12 Working, New York 1974, S. XI
- 13 Karl Marx, zitiert nach Mendner, Technologische Entwicklung und Arbeitsprozeß, Frankfurt 1975, S.126ff
- 14 Die dabei stattfindenden Auseinandersetzungen, Konflikte, Arbeitskämpfe konnte ich seit Mitte der 70er-Jahre in verschiedenen Rollen als Informatiker, Gewerkschafter und Betriebsräteberater begleiten: AK Rationalisierung (Hrg.): Verdatet Verdrahtet Verkauft. Beiträge zu Informatik und Gesellschaft, Stuttgart 1982/1986; AK Rationalisierung (Hrg.): EDV – Textverarbeitung – Bildschirmarbeit (mit der Geschichte von der Insel). Berlin 1983; Heß, Schwitalla, Wicke: Auswirkungen von Bildschirmtext im Dienstleistungsgewerbe, in K.TH. Schröder (Hrsg.): Arbeit und Informationstechnik, Karlsruhe 1986; Friedrich, Heß, Schwitalla, Wicke: Bildschirmtext in Kreditinstituten, Versicherungen und Kommunalverwaltungen. Einsatz und Auswirkungen auf die Beschäftigten. Frankfurt 1988; Herrmann, Heß, Meise: Die Verformung der Kommunikationsstrukturen durch ISDN, in Kitzing, Linder-Kostka, Obermaier (Hg.): Schöne neue Computerwelt. Zur gesellschaftlichen Verantwortung der Informatiker. Berlin 1988; TBS NRW (Hg.): Gute Arbeit im Call Center und im Kunden-Service durchsetzen! Handlungshilfe für die Interessenvertretung, 2014; TBS NRW (Hg.): E-Government in der öffentlichen Verwaltung. Plädoyer für eine ungewohnte Zusammenarbeit in der digitalen Transformation! Dortmund 2016; TBS NRW (Hg.) Mobile Arbeit, computing anywhere ... Neue Formen der Arbeit gestalten! Dortmund 2017
- 15 Man kann hier von einer „Call-Centrifizierung“ im gesamten Kundenservicebereich sprechen, s. Klaus Heß in Computer und Arbeit, Heft 9/2014.
- 16 Harari, Homo Deus
- 17 Dabei geht die „Handlungsträgerschaft Mensch“ teilweise oder vollständig auf die „Handlungsträgerschaft Autonome Software-Systeme“ über. Vgl. Der mitbestimmte Algorithmus: Der mitbestimmte Algorithmus | blog-zukunft-der-arbeit.de

Referenzen

- Collins, A./Warnock, Eleanor H./Aiello, N./Miller, M. L.: Reasoning from incomplete Knowledge, in: Bobrow, Daniel G. and Allan, Collins: Representation and Understanding. Studies in Cognitive Science 1975
- Cruse, H./Dean, J./Ritter, H.: Die Entdeckung der Intelligenz oder können Ameisen denken? Intelligenz bei Tieren und Maschinen. München, 1998.
- Elithorn, Alick. & Jones, David. 1973, Artificial and human thinking. Edited by Alick Elithorn and David Jones Amsterdam, S.7
- Heß, Klaus: Was bedeutet: Menschliche Fähigkeiten verschwinden im Computer? In Schöne elektronische Welt. Technologie und Politik 19 (rororo aktuell), Reinbek 1982
- Kirste, Moritz/Schürholz, Markus: Entwicklungswege zur KI, in: Volker Wittpahl (Herausgeber) KÜNSTLICHE INTELLIGENZ Technologie | Anwendung | Gesellschaft, Berlin 2019 Wolfgang Wahlster: Die Zukunft der KI: Die Kombination von neuronalem Lernen und modellbasierten Methoden, Online-Tagung Evangelische Akademie Hofgeismar, 19.3.2022
- Müller, Norbert: Das Räderwerk des technischen Fortschritts – Endstation: Menschen wie Computer, S.9, in: Norbert Müller (Hrsg.): Schöne elektronische Welt. Computer – Technik der totalen Kontrolle. Technologie und Politik 19 (rororo aktuell), Reinbek 1982
- Schank: The Role of Memory in Language Processing. In Cofer, Charles N.: Structures of Human Memory, San Francisco 1976
- Simon, Herbert A.: What Computers Mean for Man and Society, Science, März 1977
- Schröter, Welf (Hrsg.) Der mitbestimmte Algorithmus. Gestaltungskompetenz für den Wandel der Arbeit 2019



Stuart Russell: Human Compatible – Künstliche Intelligenz und wie der Mensch die Kontrolle über superintelligente Maschinen behält

Buchbesprechung

Die aktuellen Entwicklungen in der künstlichen Intelligenz werden von vielen (auch von mir) als existentielles Risiko für die Menschheit und das Leben auf unserem Planeten gesehen. Gleichzeitig besitzt Künstliche Intelligenz (KI) ein enormes Potenzial zum Besseren. Mit guten Modellen kann sie nicht nur große Arbeitserleichterungen bringen (denn wer will wirklich selbst mit Excel arbeiten oder stundenlang einen Laster über die Autobahn steuern?) sondern auch in Wissenschaft und Gesellschaft bei der Analyse komplexer Fragen helfen, und uns beim Treffen intelligenterer Entscheidungen auf solider Basis von Daten und Erkenntnissen unterstützen. Damit kann sie auch helfen, andere existentielle Risiken einzudämmen und viele wichtige Entwicklungen – in der Medizin und Biologie, der Umwelttechnik, der Physik und den Ingenieurwissenschaften – wesentlich schneller voranzubringen, als es ohne diese Werkzeuge möglich wäre.

In seinem aktuellen Buch *Human Compatible* beschreibt Stuart Russell die existentiellen Risiken einer künstlichen Intelligenz, die mit aktuellen Methoden und Kostenfunktionen entworfen wird. Er beschreibt aber auch eine Strategie, um künstlich intelligente Systeme so zu gestalten, dass sie nicht dazu prädestiniert sind, die Zukunft der Menschheit zu vernichten. Damit rücken die Möglichkeiten einer positiven künstlichen Intelligenz und eines an menschlichen Bedürfnissen ausgerichteten maschinellen Lernens in den Vordergrund.

Erstes Thema:

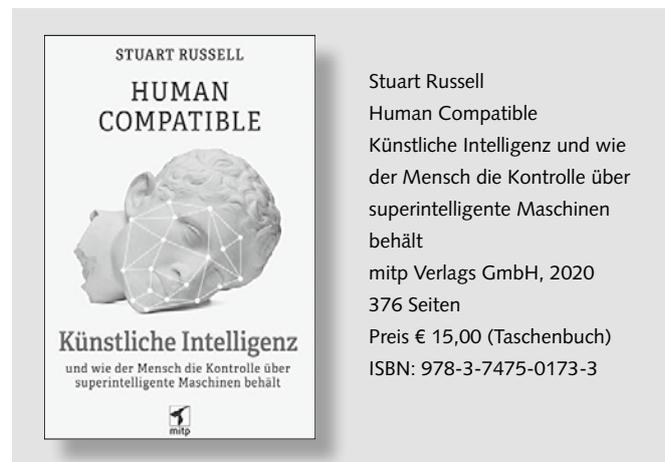
Warum KI ein existentielles Risiko darstellt

Die Fortschritte in der Entwicklung *intelligenter* Systeme sind beeindruckend und recht rasant. KI-Systeme können Sprache und Bilder bei entsprechendem Training ähnlich gut erkennen wie Menschen und GPT-3 schreibt inzwischen schönere Texte als manch einer von uns – jedenfalls an unseren weniger guten Tagen. Maschinelles Lernen und künstlich intelligente Systeme optimieren aber in ihrem Verhalten immer formal gefasste sogenannte *Kostenfunktionen*. Beispielsweise könnte ein KI-System zur Routenplanung immer versuchen, die Reisedauer zu minimieren – die Aufgabenstellung wäre dann einfach:

Verhalte Dich so, dass die Kostenfunktion $F = \text{Reisedauer}$ minimal wird.

Leider ist es aber sehr schwer, eine formale Kostenfunktion aufzustellen, die eine (zukünftige) superintelligente KI sinnvoll optimieren kann, ohne die Menschheit zu vernichten, oder jedenfalls mindestens krassen Schaden anzurichten. Das einfachste von vielen Beispielen einer derartigen Fehloptimierung in *Human Compatible* ist König Midas. Er wünschte sich, dass alles, was er berührt, zu Gold wird. Leider hatte er nicht bedacht, dass man Gold bekannterweise nicht essen kann. Ein noch fataleres Beispiel, mit konkreterem Bezug zu intelligenten Maschinen: Die Lösung der Optimierungsaufgabe, menschliches Leiden zu minimieren, führt in einen stabilen Punkt, an dem das menschliche Leiden genau Null ist.

Aber auch, wenn man versucht, die Kostenfunktionen einer KI mit allergrößter Sorgfalt zu gestalten, läuft man immer wieder in dieselbe Falle: Immer wieder fehlen wichtige Aspekte in der Kos-



Stuart Russell
Human Compatible
Künstliche Intelligenz und wie der Mensch die Kontrolle über superintelligente Maschinen behält
mitp Verlags GmbH, 2020
376 Seiten
Preis € 15,00 (Taschenbuch)
ISBN: 978-3-7475-0173-3

tenfunktion, so dass letztlich die engstirnige Optimierung derselben direkt in eine Katastrophe führt.

Aufgabe zur Illustration: Überlegen Sie sich eine formale Kostenfunktion, die in letzter Konsequenz und mit vollem Einsatz optimiert werden kann, ohne dass existentielle Probleme entstehen. Leider wird man eine superintelligente KI auch nicht gut abschalten können, wenn sie einmal damit loslegt, ihre Kostenfunktion zu optimieren – denn wenn sie ausgeschaltet ist, wird die Kostenfunktion ja nicht mehr optimiert. Also muss sie die Menschen auch von ihrem Ausschalter fernhalten, um die Kosten zu minimieren! Und wie aus Spielen wie AlphaGo schon aktuell für einfache KI und kleine Kunstwelten klar ist, kann eine KI wesentlich weiter in die Zukunft planen, als irgendein Mensch, der sie gerne ausschalten möchte. Bisher ist das vergleichsweise unproblematisch, weil auch die intelligentesten Systeme noch weitreichenden technischen (oder kognitiven?) Beschränkungen unterliegen. Aber einerseits ändern sich viele dieser Einschränkungen heute schnell – sogar diejenigen, die eine Bewaffnung von KI-Systemen verbieten sollten – und andererseits kommen auch Systeme aus Maschine und menschlichem *Handler* mit ähnlichen Problemen daher, da in vielen Kontexten (nicht nur im Krieg, auch oft genug an der Börse) der menschliche Akteur die Maschine nur nutzt und nicht selbst hinterfragt.

Dieses Problem ist so lange scheinbar unlösbar, wie man versucht, der KI eine explizite Kostenfunktion zu geben.

Zweites Thema: Human Compatible – wie KI gestaltet werden müsste

Stattdessen schlägt Russell vor, KI so zu gestalten, dass sie drei wesentliche Grundsätze beachtet:

1. KI soll so entwickelt werden, dass ihre einzige Zielfunktion in der bestmöglichen Realisierung menschlicher Präferenz liegt.
2. Die KI muss dabei unsicher sein, welche Handlung am besten menschliche Präferenzen realisiert – schließlich ist das niemandem, auch nicht den Menschen, in vollem Umfang klar. Die KI muss deswegen auch ihr eigenes Unwissen kennen bzw. modellieren. Und im Angesicht ihrer eigenen Unsicherheit über die menschlichen Präferenzen muss es ihr Ziel sein, diese Unsicherheit zu verringern, da sie nur so das erste Ziel erreichen kann.
3. Die maßgebliche Quelle für Informationen über menschliche Präferenzen ist das menschliche Verhalten.

Weil aber menschliche Präferenzen variabel sind (von Mensch zu Mensch, von Situation zu Situation, und über die Zeit) kann keine KI jemals vollständiges Wissen über die menschlichen Präferenzen erlangen. Stattdessen muss die KI also auch auf Dauer ihre eigenen Unsicherheiten mit modellieren – sie muss also immer mit einem statistischen oder jedenfalls unsicherheitsbehafteten Modell menschlicher Präferenz arbeiten.

Sie muss schließlich so entscheiden, dass unter dieser Unsicherheit die erwartete Übereinstimmung mit der menschlichen Präferenz am größten wird. Im Extremfall kann das auch dazu führen, dass die KI sich selbst ausschaltet, wenn das das geringste Risiko darstellt, sich gegen die Präferenzen der Menschen zu verhalten.

So erhält man, wenn die drei Prinzipien beachtet werden, endlich KI-Systeme, die auch bei herausragender „Intelligenz“ und größtmöglichem Zugriff auf Ressourcen weiter durch Menschen kontrollierbar bleiben!

Ausblick: Warum damit leider noch nicht alles klar ist

Zweifellos stellt das Konzept von Stuart Russell mindestens so viele Fragen, wie es Antworten gibt. Aus meiner Sicht ist es trotzdem das einzige Konzept, das in der Lage ist, eine Ausrichtung maschineller Intelligenz an menschlichen Zielen zu erreichen.

Wie das im Einzelnen geschieht, also, wie genau menschliche Präferenz modelliert wird, wie erreicht wird, dass das Modell auch die eigene Unsicherheit richtig einschätzt oder jedenfalls nicht unterschätzt und wie dann die vielen und oft veränderlichen und in Konkurrenz stehenden Ziele vieler Menschen zu einer bestmöglichen Handlung führen, das sind große Fragen und Aufgaben für die Entwicklung der KI in den nächsten Jahrzehnten, wenn nicht darüber hinaus.

Ich persönlich glaube aber, dass dieser Weg der einzige ist, der überhaupt zu KI-Systemen führen kann, die ihre Handlungen an humanistischen und humanen Zielen ausrichten – und damit der einzige verantwortungsvolle Weg in die Zukunft des maschinellen Lernens. Deswegen möchte ich diese Buchvorstellung mit einem Appell beenden: Lassen Sie uns Russells Strategie zu unserer eigenen machen – so wie Hilberts Programm die Mathematik auf solide Beine stellen wollte, können wir nun „Russells Strategie für menschenzentrierte KI“ in den Mittelpunkt der Forschung stellen:

Verantwortungsvolle KI muss menschliche Ziele und Werte unterstützen.

Es mag sein, dass das Ziel (wie schon bei Hilbert) zu hochgesteckt oder unrealistisch ist. Aber hier gibt es eine überzeugende, vielleicht die wichtigste Idee, wie eine an menschlichen Werten ausgerichtete KI vielleicht doch möglich wird – wenn wir gemeinsam und koordiniert als Forschende und als Gesellschaft in diese Richtung gehen.

Stuart Russell ist Professor für Informatik an der University of California, Berkeley. Er ist bei vielen Informatiker:innen bekannt geworden durch sein Buch *Artificial Intelligence: A Modern Approach*, das er mit Peter Norvig 1995 in der ersten Auflage veröffentlicht hat. Es wird inzwischen von weit über 1000 Universitäten weltweit verwendet und als Standardwerk der klassischen KI gesehen. Sein neues Buch *Human Compatible* ist im englischen Original im April 2020 und in der deutschen Übersetzung im Juni 2020 erschienen. Es setzt keine technischen Vorkenntnisse voraus und erklärt in Anhängen für Interessierte kurz die Grundlagen der modernen KI. Aus meiner Sicht möchte ich zu diesem Buch, das aktuell für 15 € als Taschenbuch erhältlich ist, meine stärkste Leseempfehlung geben: Es ist unglaublich intelligent, spannend zu lesen, hervorragend geschrieben und ganz klar und stringent argumentiert. Es zeigt gleichzeitig die Gefahren einer falsch entworfenen KI, den Stand der Diskussionen dazu, und eine für mich überzeugende, weitsichtige und fundamental richtige Strategie zur Entwicklung menschenfreundlicher Technologien.



Dorothea Kolossa

Dorothea Kolossa ist Professorin für Kognitive Signalverarbeitung an der Ruhr-Universität Bochum. In ihrer Forschung beschäftigt sie sich mit dem maschinellen Lernen aus Zeitreihen, besonders für Sprach- und Multimediadaten. In verschiedenen Projekten arbeiten sie und ihr Team an einer verantwortungsvollen und die Privatsphäre achtenden Entwicklung von Sprachtechnologien und intelligenten Systemen.

Eine lebenswerte Welt – Kann Künstliche Intelligenz helfen, das Ziel zu erreichen?

Künstliche Intelligenz (KI) ist ein Teilgebiet der Informatik, das in den Medien, der Politik und der Wirtschaft große Aufmerksamkeit erregt hat. Es besteht eine weltweite Erwartung, dass KI die zukünftige Schlüsseltechnologie wird. In diesem Artikel skizziere und diskutiere ich, ob diese Voraussagen und Hoffnungen von einer technischen Perspektive her realistisch sind und unter welchen Bedingungen KI zu einer lebenswerten Welt beitragen kann. Ich gehe kurz darauf ein, was eine lebenswerte Welt sein kann, welche Rolle Wissenschaft und Technologie für die menschliche Gesellschaft spielen, wie sich da Digitalisierung und KI einordnen, was es mit dem aktuellen Hype um die KI auf sich hat, welche methodischen Grundlagen KI hauptsächlich hat, wie sich künstliche und natürliche Intelligenz zueinander verhalten, was typische Anwendungen der KI sind welche ethischen Fragen im Kontext von KI gestellt werden. Zusammenfassend versuche ich schließlich, die Titelfrage zu beantworten.

Eine lebenswerte Welt

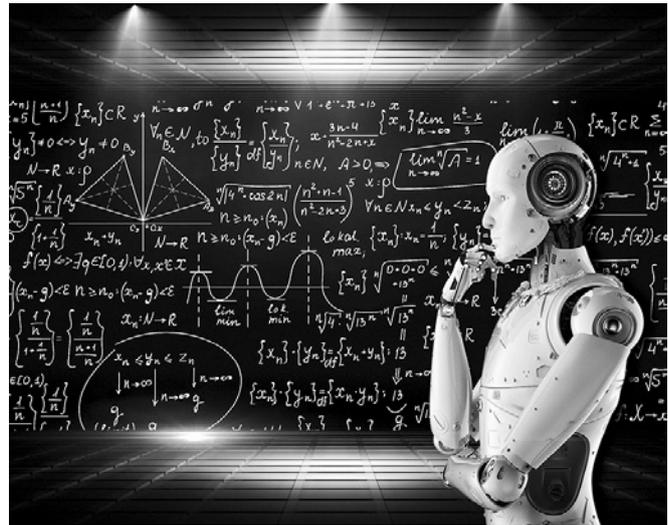
Nach meinem Verständnis ist eine lebenswerte Welt eine friedliche Welt ohne Armut und Hunger, ohne Ausbeutung und Naturzerstörung, ist eine Welt mit gleichen Rechten und Chancen für alle Menschen. In dieser Welt sind die Wirtschaft und unser aller Lebensstil nachhaltig. Darüber hinaus ist der Einsatz von Technologien mit diesen Zielen verträglich. Es mag auffallen, dass sich diese Beschreibung einer lebenswerten Welt nicht viel von den Versprechungen unterscheidet, die in der UN-Charta (1949) gemacht werden. Die wesentliche Charakteristik einer lebenswerten Welt, wie ich sie beschreibe, ist in gewisser Weise die naive Form dessen, was vertieft und elaboriert in Edward Morins Konzeption von *Heimatland Erde* gefunden werden kann (Morin und Kern 1999).

Das Problem ist, dass der aktuelle Zustand der Welt eher das Gegenteil von diesen hohen Idealen ist. Dutzende regionaler Kriege werden geführt, und es besteht das permanente Risiko eines globalen militärischen Konflikts. Armut, Hunger, Diskriminierung, Ausbeutung, unzureichender Zugang zu Wasser, Gesundheitsversorgung und Bildung sind das Schicksal von vielen hundert Millionen Menschen. In vielen Staaten der Welt werden die Menschenrechte von diktatorischen und undemokratischen Regimen ignoriert. Die Zerstörung der Natur wird in atemberaubendem Tempo vorangetrieben, und die bisher ergriffenen Maßnahmen gegen den Klimawandel sind hoffnungslos unzureichend, was das Leben vieler Menschen zukünftig noch miserabler werden lässt. Die dominante kapitalistische Ökonomie führt nicht nur dazu, dass sich die Schere zwischen Arm und Reich immer weiter öffnet, sondern ist auch die Ursache extensiven Verbrauchs natürlicher Ressourcen, etwa die Hälfte davon ohne jede Kompensation. Hofkirchner (2020) beschreibt die derzeitige Situation als Zeitalter globaler Herausforderungen mit der Gefahr, dass sich die Menschheit selbst auslöscht. Die Menschheit befinde sich an der Schwelle zu einer globalen und nachhaltigen Gesellschaft als Rahmen für die notwendigen Bedingungen des zukünftigen Zusammenlebens mit der Möglichkeit des Scheiterns.

Soweit es in diesem Zusammenhang um Digitalisierung geht, ist die Charakterisierung einer lebenswerten Welt eng verbunden mit der Idee des *digitalen Humanismus* (siehe Nida-Rümelin und Weidenfeld 2018 sowie Werthner et al. 2022). Ich vermeide diesen Begriff möglichst, weil er mir falsch gebildet zu sein scheint. Er verweist auf einen Prozess der Digitalisierung unter humanistischen Bedingungen, während Humanismus aus meiner Sicht nicht *digital* sein kann.

Wissenschaft und Technologie

In Verbindung mit gesellschaftlichen Strukturen, Kultur und Wirtschaft sind Wissenschaft und Technologie seit Tausenden von Jahren treibende Faktoren für die Entwicklung der menschlichen Zivilisation. Vor rund 6000 Jahren – vielleicht noch früher – begannen Menschen systematisch Landwirtschaft und Viehzucht zu betreiben und im Zuge dessen sesshaft zu werden. Diese Umwälzung war begleitet von der Erfindung der Schrift und erstaunlichen Erkenntnissen in Arithmetik, Geometrie und Astronomie als Basis von Architektur und Stadtplanung, Konstruktion von Bewässerungssystemen, Einrichtung von Verkehrs- und Transportrouten, Schiffbau, maritimer Navigation und vielen weiteren bemerkenswerten Errungenschaften.



Quelle: Mike MacKenzie, flickr.com, CC BY 2.0

Wissenschaft und Technologie haben während der letzten 250 Jahre durch die Industrialisierung einen großen Schub bekommen, der durch die Computerisierung und Digitalisierung der letzten 70 Jahre noch einmal zusätzlich verstärkt und in jüngerer Zeit durch KI und Robotik um neue Elemente bereichert wurde. Der Übergang von einer agrarischen und handwerklichen Gesellschaft zu einer industriellen und kapitalistischen war in mancher Hinsicht revolutionär. Die anhaltende Digitalisierung, die eigentlich eine Algorithmisierung ist und durch KI eine besondere Ausprägung erhält, hat nach meiner Einschätzung einen eher evolutionären Charakter innerhalb der Industrialisierung. Da Wissenschaft und Technologie grundsätzlich entscheidende Kräfte für gesellschaftliche Entwicklungen darstellen, kann man

natürlich speziell fragen, ob KI helfen kann, die Welt lebenswert für alle Menschen zu machen. Die Antwort ist aber nicht einfach JA oder NEIN, sondern hängt von vielem ab.

Der Hype um die KI

In den letzten 20 Jahren konnte man spektakuläre Erfolge der KI bestaunen bei Spielen wie Schach, Go, Jeopardy u. a. sowie bei praktischeren Anwendungen wie Sprach- und Bildverarbeitung einerseits und Robotik andererseits. Außerdem werden einige KI-Propagandist:innen nicht müde, weitere fantastische Durchbrüche zu versprechen. Beides befeuert die Erwartungen in Politik und Wirtschaft, dass KI die Schlüsseltechnologie für die zukünftige Wertschöpfung wird und – gelegentlich sogar – für die Führungsrolle in der Welt (vgl. z. B. Lee 2018). Viele Staaten haben nationale KI-Strategien ausgearbeitet (siehe beispielsweise European Commission 2021 und van Roy et al. 2021) und sind dabei, zig Milliarden Dollar, Euro, Yen, Yuan etc. in die Entwicklung von KI zu stecken. Man darf gespannt sein und sich wundern.

KI-Methoden

Obwohl das Gebiet der KI in ein weites Spektrum an Themenbereichen aufgeteilt ist, lassen sich einige grundlegende gemeinsame Methoden und Prinzipien identifizieren (vgl. z. B. Russell und Norvig 2021). Eine wesentliche Methode ist die Verwendung von Regeln von der Art von Rechenregeln, logischen Schlussregeln, Grammatikregeln und Spielregeln. Während diese allgemein bekannten Regeln üblicherweise einfach in Struktur und Größe und klein in der Zahl sind, kann ein KI-System aus einer riesigen Zahl teilweise kompliziert gestalteter Regeln bestehen. Die grundlegenden Prinzipien und der Gebrauch der Regeln sind aber in beiden Fällen ähnlich: Sie können auf die jeweils unterliegenden Informationsstrukturen angewendet werden, bewirken dabei lokale Änderungen und führen zu komplexen Transformationen und Berechnungen durch Iteration.

Eine weitere Methode, die in der KI häufig angewendet wird, ist die Wahrscheinlichkeitstheorie und ihre Varianten, um Unsicherheit, Vagheit, Unschärfe und Ähnliches zu beschreiben, wie sie bei praktischen Anwendungen oft vorkommen. Allerdings ist der Umgang mit Wahrscheinlichkeiten nicht unproblematisch und führt leicht zu Fehlinterpretationen. Denn unwahrscheinliche Ereignisse können jederzeit eintreten, während sehr wahrscheinliche möglicherweise lange auf sich warten lassen. Die Wahrscheinlichkeit eines Ereignisses und die Zahl seines Auftretens stimmen nur ungefähr überein, wenn Ereignisse in großer Zahl eintreten. Was selten passiert, hat keine Wahrscheinlichkeit. Deshalb glaube ich auch nicht im Gegensatz zu so manchen KI-Fachleuten, dass man Kriege, Aufstände, politische und wirtschaftliche Krisen mit KI-Methoden vorhersagen kann.

Maschinelles Lernen (einschließlich des sogenannten *deep learning*) als eines der momentan wichtigsten Teilgebiete der KI kombiniert beide Methoden in einer Weise, dass eindrucksvolle Fortschritte in einem weiten Anwendungsfeld erreicht werden konnten im Vergleich zu klassischen KI-Methoden (siehe z. B. Goodfellow et al. 2016, Jiang 2021, Mohri 2018).

Neben dem maschinellen Lernen gibt es weitere Teilgebiete der KI mit signifikanten Fortschritten wie autonome Systeme und Schwarmintelligenz (siehe z. B. Bonabou et al. 1999, Engelbrecht 2006, Kennedy und Eberhart 2001, Siegart et al. 2011, Thrun et al. 2005). Ein autonomes System beobachtet seine Umgebung mit Hilfe verschiedener Sensoren, fusioniert die Daten und wertet sie aus und entscheidet daraufhin über seine nächsten Aktionen, um die vorgegebenen Ziele zu erreichen. Kommuniziert ein solches System mit anderen gleichartigen Entitäten, bilden sie einen Schwarm.

Die meisten KI-Prinzipien sind seit Jahrzehnten bekannt und unterscheiden sich nicht wesentlich von den Methoden anderer Teilgebiete der Informatik. Der Hauptunterschied ist, dass sich die diversen wissenschaftlichen Gemeinschaften nicht gut kennen, wenig Austausch pflegen und wenig voneinander wissen. Die großartigen Erfolge der KI in jüngster Zeit beruhen zu einem erheblichen Teil auf der inzwischen erreichten hohen Rechengeschwindigkeit und der immensen Speicherkapazität, die ambitionierte Projekte mit riesigen Datenmengen möglich machen. Es handelt sich also nicht um Magie und nicht um „disruptive Innovation“, wie in diesem Kontext gern behauptet wird, sondern um normalen wissenschaftlichen und technologischen Fortschritt.

Intelligenz?

Der größte Teil der KI wird der *schwachen KI* zugerechnet. deren Ziel ist die Simulation eng begrenzter Prozesse, für die Menschen ihre Intelligenz einsetzen wie Spielen, logisch Schließen, Problemlösen. Sprachverstehen, Bilderkennen, Planen, Entscheiden usw. Der Unterschied zwischen schwacher KI und Informatik allgemein ist nicht allzu groß. Denn in der Informatik geht es überall um Berechnen, Sammeln, Suchen, Sortieren, Kontrollieren, Verwalten, Navigieren usw. – alles Aktivitäten, für die Menschen ihre Intelligenz benötigen.

Wie verhält sich die Simulation intelligenten Verhaltens in eingegrenzten Kontexten zu natürlicher und insbesondere zur menschlichen Intelligenz? Eine abschließende Antwort ist unmöglich, weil noch gar nicht verstanden ist, wie natürliche Intelligenz überhaupt funktioniert. Aber schon auf der phänomenologischen Ebene kann man erhebliche Unterschiede erkennen.

Man betrachte zum Beispiel das KI-Konzept des *deep learning*. Nichts ist wirklich *tief* daran. Es beruht auf künstlichen neuronalen Netzen mit mehreren Schichten zwischen Eingabe- und Ausgabeschicht, was als *tief* gestapelt angesehen wird. Das Lernen eines solchen Netzes ist eher ein *Trainieren* mit Tausenden und Abertausenden Eingabebeispielen für das, was gelernt werden soll. Typisch handelt es sich um Bilder von Hunden und Katzen oder – praktischer – von Hautkrebs. Menschen dagegen und schon kleine Kinder lernen anhand weniger Beispiele. Vor allem lernen Menschen permanent zu einem breiten Spektrum an Gegenstandsbereichen, nicht nur Hunden und Katzen. Das Problem ist nicht der Begriff *maschinelles Lernen*, sondern dass die unterliegenden rein technischen Prozesse missverstanden werden als etwas von ähnlicher Qualität wie menschliches Lernen. KI-Fachleute weisen auf diese Unterschiede zu selten hin, einige von ihnen ignorieren sie oder halten sie gar für klein und irrelevant.

Hinsichtlich der *Autonomie* verhält es sich ganz ähnlich. Die Entscheidungen, die ein autonomes System trifft, sind programmiert, also menschengemacht, und bleiben voll und ganz im Rahmen dessen, was die Systemmacher:innen vorgesehen haben. Ein solches System versteht nicht, was es tut, hat keine eigenen Ambitionen, weiß nicht von sich und hat keinen freien Willen. Ich kann nicht glauben, dass sich das bei menschlicher Autonomie auch so verhält. Von KI-Fachleuten, die sich mit Autonomie beschäftigen, erhält man aber diesen Hinweis selten bis gar nicht.

Im Gegensatz zur schwachen KI zielt die kleine Minderheit der Befürworter:innen einer *starken* KI auf Systeme ab, die sich im selben Sinnen intelligent verhalten wie Menschen. Einige halten sogar superintelligente Systeme für möglich und wünschenswert, die besser denken können als Menschen – ein Thema, das sich bestens verkaufen lässt (vgl. z. B. Bostrom 2005, Häggström 2016 und Kurzweil 2005). Soweit ich das sehe, gibt es jedoch bisher keine ernsthaften Hinweise, dass solche Systeme bald verwirklicht werden oder überhaupt jemals. Die meisten Autor:innen in dem Sammelband von Hofkirchner und Kreowski (2021) stimmen mit dieser Auffassung überein. Vor rund 40 Jahren habe ich einem Vortrag von John McCarthy gelauscht, einem der Gründungsväter der KI. Er behauptete, Superintelligenz sei technisch möglich, aber es erfordere mindestens ein Dutzend wissenschaftlicher Durchbrüche von der Qualität der Relativitätstheorie. Wenn ich das richtig sehe, warten wir immer noch auf den ersten davon.

Anwendungen von KI

Weltweit ist eine irrsinnige Menge an Geld verfügbar für die Entwicklung von KI-Technologien, viele Anwendungen werden geplant und vorangetrieben. Vieles davon klingt vielversprechend. KI zusammen mit Robotik wird bereits mit einigem Erfolg in der Medizin, Produktion, im Transportwesen, in der Verwaltung und in anderen Bereichen eingesetzt. Viele prototypische Anwendungen sind in der Entwicklung wie autonome Fahrzeuge, Serviceroboter aller Art und anderes mehr.

Aber manches davon steht in einem offensichtlichen Widerspruch zu der Idee einer lebenswerten Welt für alle. Einige Anwendungen mehren vor allem den Profit privater Unternehmen, verbrauchen Unmengen an Energie, dienen der sozialen Überwachung von Einzelpersonen über bestimmte Personengruppen bis hin zu einem ganzen Volk wie in China und erhöhen

den Horror des Krieges durch autonome tödliche Waffen, Drohnenschwärme und eine ganze Palette KI-gestützter Militärtechnik. In der *Roadmap* des Department of Defense (2014) und im Sammelband von Williams und Scharre (2015) wird ein weites Spektrum autonomer Waffensysteme beschrieben und diskutiert einschließlich einer langen Liste technologischer Probleme, die gelöst werden müssen, bevor autonome Systeme das Kriegsgeschäft dominieren werden.

KI-Ethik

Der zunehmende Gebrauch von KI-Technologien in kritischen Anwendungsbereichen und die Zwecke, für die sie entwickelt werden, bringen verschiedene ethische Probleme mit sich. Beispielsweise muss angesichts der Entwicklung letaler autonomer Waffen gefragt werden, ob es erlaubt sein soll, dass Maschinen über Tod und Leben entscheiden. Selbst in militärischen Kreisen wird dieses Problem gesehen. In Williams und Scharre (2015) zum Beispiel ist ein ganzes Kapitel dieser Frage gewidmet. Während viele Autor:innen da die rote Linie ziehen, argumentiert etwa Arkin (2009), dass ein künstliches Gewissen nicht nur möglich, sondern wünschenswert ist, denn es gerät nicht in Panik und hat keine Angst, Befehle zu verweigern.

Das AI4People Scientific Committee mit zwölf Expert:innen unter der Leitung von Luciano Floridi hat sechs große Studien zu ethischen Aspekten von KI vergleichend ausgewertet (Floridi et al. 2018), um dann Vorschläge zu machen für die Entwicklung einer guten KI-Gesellschaft. Zuerst beschreibt das Komitee die Chancen, die KI-Technologien mit sich bringen, um die Würde und das Gedeihen der Menschen zu befördern, aber auch die Risiken. Dann identifiziert es fünf Prinzipien, die „gute KI“ charakterisieren:

- Beneficence: Promoting WellBeing, Preserving Dignity, and Sustaining the Planet,
- Non-maleficence: Privacy, Security and *Capability Caution*,
- Autonomy: The Power to Decide (Whether to Decide),
- Justice: Promoting Prosperity and Preserving Solidarity, and
- Explicability: Enabling the Other Principles Through Intelligibility and Accountability.

Hans-Jörg Kreowski



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. Seit 2019 ist er außerdem Mitherausgeber des *Grundrechte-Reports*.

Die Studie ist äußerst signifikant, da sie einen umfassenden Überblick über die weltweite Debatte zu KI-Ethik verschafft, und viele wichtige Punkte sind in richtiger Weise angesprochen. Ich vermisse allerdings eine explizite Auseinandersetzung mit der Frage, wer und was dafür verantwortlich ist, dass die Situation jetzt „schlecht“ ist, und was getan werden muss, um politisch und wirtschaftlich eine gute KI-Gesellschaft zu erreichen.

Fazit

Kann KI helfen, die Welt lebenswert zu machen? Unglaublich viel Geld wird in die Entwicklung von KI gesteckt. Tausenden und Tausende KI-Forscher:innen, -Entwickler:innen und -Manager:innen werden das Geld einstecken und Ergebnisse aller Art produzieren. Das schließt KI-Anwendungen ein, die Reiche reicher machen und Mächtige mächtiger, die der sozialen Überwachung dienen und so die Menschenrechte aushöhlen, die in Waffentechnik einfließen, ohne dass Kriegsführung dadurch intelligenter würde. Die Attitüde vieler KI-Fachleute ist arrogant und ignorant, indem sie nicht nur die technischen Grenzen von KI nicht beachten und respektieren, sondern auch alles entwickeln, was von ihnen verlangt wird und wofür sie bezahlt werden. Sie ignorieren die Tatsache, dass KI-Algorithmen und -Systeme oft undurchschaubar und unkontrollierbar sind und damit häufig riskant und gefährlich. Ethische Aspekte spielen für sie kaum eine oder gar keine Rolle. Die Erwartungen in Politik und Wirtschaft sind übertrieben, weil sie außer Acht lassen, dass technologische Entwicklungen praktisch immer langsam vorstatten gehen, mehr kosten als geplant und die Ergebnisse in der Regel weit hinter dem zurückbleiben, was anfangs versprochen wurde. Und das ist noch lange nicht alles, was auf die Antwort NEIN weist.

Dennoch ist auch die Antwort JA denkbar, wenn alle Anstrengungen auf das Ziel gerichtet werden. Das allerdings erfordert eine dramatische Veränderung der gesellschaftlichen Rahmenbedingungen, der Leitlinien und Ziele von Politik und Wirtschaft. Es hängt ab von der Art, wie Technologien zukünftig entwickelt und eingesetzt werden. Die Regierenden der Welt einerseits und die Wissenschaftler:innen und Ingenieur:innen andererseits müssen sich an das Prinzip Verantwortung von Hans Jonas (1979) halten: „Handle so, dass die Wirkungen deiner Handlung verträglich sind mit der Permanenz echten menschlichen Lebens auf Erden.“

Referenzen

- Arkin, Ron C. 2009. *Governing Lethal Behavior in Autonomous Systems*. Chapman and Hall, Boca Raton, FL.
- Bonabeau, Eric, Marco Dorigo, Guy Theraulaz. 1999. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press.
- Bostrom, Nick. 2005. A History of Transhumanist Thought. *J. Evol. Technol.*, 14, 1–25.
- Department of Defense. 2014. *Unmanned Systems Integrated Roadmap, FY 2013-2038*. Washington.
- Engelbrecht, Andries P. 2006. *Fundamentals of Computational Swarm Intelligence*. John Wiley & Sons,
- European Commission. 2021. *Coordinated plan on artificial intelligence 2021 review*. <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review>.
- Floridi, Luciano, Josh Cows, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, Effy Vayena. 2018. *AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*. *Minds and Machines* 28:689–707.
- Goodfellow, Ian, Yoshua Bengio, Aaron Courville. 2016. *Deep Learning*, MIT Press.
- Hägström, Olle. 2016. *Here Be Dragons: Science, Technology and the Future of Humanity*. Oxford University Press: Oxford, UK.
- Hofkirchner, Wolfgang. 2020. A paradigm shift for the Great Bifurcation. *BioSystems* 197, 104193.
- Hofkirchner, Wolfgang, Hans-Jörg Kreowski (Eds.). 2021. *Transhumanism: The Proper Guide to a Posthuman Condition or a Dangerous Idea?*. Springer Nature.
- Jiang, Hui. 2021. *Machine Learning Fundamentals – A Concise Introduction*. New edition, Cambridge University Press.
- Jonas, Hans. 1979. *Das Prinzip Verantwortung*. Insel-Verlag, Frankfurt am Main.
- Kennedy, James, Russell C. Eberhart. 2001. *Swarm Intelligence*. Evolutionary Computation Series, Morgan Kaufman, San Francisco.
- Kurzweil, Raymond. 2005. *The Singularity Is Near: When Humans Transcend Biology*. Viking Adult: New York, NY, USA.
- Lee, K.-F. 2018. *AI Superpowers: China, Silicon Valley, and the New World Order*. Houghton Mifflin Harcourt, Boston New York, USA.
- Mohri, Mehryar, Afshin Rostamizadeh, Ameet Talwalkar. 2018. *Foundations of Machine Learning*. MIT Press, Second Edition.
- Morin, Edgar, Anne Brigitte Kern. 1999. *Homeland Earth: A Manifesto for the New Millennium*. Hampton Press.
- Nida-Rümelin, Julian, and Nathalie Weidenfeld. 2018. *Digitaler Humanismus – Eine Ethik für das Zeitalter der Künstlichen Intelligenz*. Piper Verlag GmbH, München.
- Russell, Stuart, Peter Norvig. 2021. *Artificial Intelligence: A Modern Approach*, Global Edition, 4th Edition, Pearson.
- Siegwart, Roland, Illah Reza Nourbakhsh, Davide Scaramuzza. 2011. *Introduction to Autonomous Mobile Robots*. 2nd Edition. The MIT Press.
- Thrun, Sebastian, Wolfram Burgard, Dieter Fox. 2005. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press.
- United Nations. 1945. *Charter of the United Nations and Statute of the International Court of Justice*. San Francisco.
- Van Roy, Vincent, Fiametta Rossetti, Karine Perset, Laura Galindo-Romero. 2021. *AI Watch – National strategies on Artificial Intelligence: A European perspective*. 2021 edition, EUR 30745 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/069178, JRC122684.
- Werthner, Hannes, Erich Prem, Edward A. Lee, Carlo Ghezzi (Eds.). 2022. *Perspectives on Digital Humanism*. Springer International Publishing.
- Williams, Andrew P., Paul D. Scharre (Eds.). 2015. *Autonomous systems: Issues for defence policymakers*. North Atlantic Treaty Organization. Headquarters Supreme Allied Commander Transformation. Capability, Engineering and Innovation, Norfolk, VA.



Joseph Weizenbaum: On the Impact of the Computer on Society, 1972

Ertappt! Auch wir fühlen uns von Joseph Weizenbaums Karikatur getroffen, denn natürlich haben wir aus der Schule der „Informatik und Gesellschaft“ schon Texte geschrieben, die genau der von Weizenbaum beschriebenen Struktur folgten: Nach der obligatorischen Einleitung, wie wunderbar transformativ die Informatik auf die Gesellschaft wirkt, müssen doch bitte auch die Risiken angesprochen werden. Zugegeben, der kritische Teil überwog, denn nach wie vor finden wir viele Entwicklungen total bekloppt, dürfen es aber nicht so salopp ausdrücken, sondern müssen mit Befremden feststellen, dass die Praxis der Informatik sich von der Förderung der Forschung hin zur Erforschung der Förderung gewandelt hat.

Weizenbaums Text enthält bereits einige Argumente seines späteren Hauptwerks „Computer Power and Human Reason: From Judgment to Calculation“. Eines davon zeichnet das Bild des lediglich auf Performance orientierten Informatikers (damals ganz sicher ein Mann), der realweltliche Probleme simuliert und das Funktionieren der Computersimulation plötzlich als Quelle der Erkenntnis nutzt, ohne zu realisieren, dass er stets nur das schon Bekannte erkennt und doch nichts begreift.

Vor allem versteht er nicht, dass er mit dem Glauben an die Übersetzbarkeit des menschlichen Daseins ins Maschinelle jegliche menschliche Autonomie zerstört. Und heute erleben wir die Konsequenzen der damals herbeibeschworenen Computerrevolution: Dazu gehört die vorherrschende Sicht, dass Wissenschaft dazu dienen soll, die Realität deskriptiv in Prozesse und Komponenten zu zerlegen und die Phänomene der Welt mit Hilfe der Sammlung und Auswertung empirischer Daten allein zu begreifen. Theorienfindung wird ersetzt durch Simulation und Beschreibung sowie immer wieder neu algorithmisch verdaute Empirie.

Die Fähigkeit, Fragen zu stellen, ist essentiell für den Menschen. Weizenbaum beobachtet mit Sorge bereits 1972, dass wir diese informationelle Hoheit gerade an Computersysteme abgeben. Abgegeben haben, müssen wir ein halbes Jahrhundert später wohl hinzufügen. „Whoever dictates the questions in large part determines the answers“ – wer auch immer die Fragen vorgibt, bestimmt zu einem wesentlichen Teil die Antworten. Weizenbaum sieht seine Zunft in einer kritischen Rolle. Er warnt vor einer technologischen Elite, die mit ihrer im Wortsinn berechnenden Denkweise zur bestimmenden gesellschaftlichen Kraft wird. Mehr noch, wir verlieren als Gesellschaft die Fähigkeit, mit einzelnen Gruppen und Individuen zu sprechen, da wir uns unterschiedlicher Metaphern und Denkweisen bedienen und letztlich sogar eine andere Sprache sprechen.

Es ist also längst überfällig, uns als Informatiker:innen, digital unterstützte Wissenschaftler:innen oder datengläubige Tech-Unternehmer:innen glaubwürdig selbst zu beschränken und zurückzunehmen. Warum also nicht aus dem Appell an die Verantwortung von Informatikfachleuten auch einen professionellen Weizenbaum’schen Eid herauslesen, wie es beispielsweise Ina Schieferdecker in Ihrem Text *Responsible Software Engineering*¹ getan hat? Dieser Eid lässt sich jedoch nicht allein in konkreten Designschritten und Prüfkriterien operationalisieren. Er ist auch daran gebunden, bereit zu sein, so manche Softwarelösung von Beginn an abzulehnen und sich ehrlich zu fragen „Brauchen wir das wirklich?“ und eben nicht „Wie machen wir es sicher?“, sich zu fragen „Ist das gut?“ und nicht „Wird es funktionieren?“ Ertappt?

1 Ina Schieferdecker (2020) Responsible Software Engineering. In: S. Gorerick (ed.), The Future of Software Quality Assurance. pp. 137-146, https://doi.org/10.1007/978-3-030-29509-7_11



Joseph Weizenbaum

On the Impact of the Computer on Society

How does one insult a machine?

[Redacted text block]

[Redacted text block]

[Redacted text block]

[Redacted text block]

[Redacted section header]

[Redacted text block]



[Redacted text block]

[Redacted]

[Redacted text block]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

[Redacted]

Der Artikel erschien zuerst In: *Science* 1972, Vol. 176, Issue 4035, pp. 609-614, <https://doi.org/10.1126/science.176.4035.609>. Der Wiederabdruck des Texts erfolgt mit freundlicher Genehmigung des Verlags.

Wissenschaft & Frieden 2/2022

Kriegerische Verhältnisse

Reflexionen zur Ukraine | Neokolonialismus

Der Krieg gegen die Ukraine löste im Februar 2022 ein Beben aus: Er brachte in vielen Landesteilen der Ukraine unermessliches Leid und Tod über die Bevölkerung; er erschütterte viele, die bis zuletzt an eine diplomatische Lösung geglaubt oder für sie gestritten hatten; er fegte in Deutschland alte, scheinbar festgefügte Gewissheiten in atemberaubendem Tempo beiseite.

In einem aktuellen Krieg Reflexionen anzustellen, gebietet auch ein gewisses Maß an Demut, Nabelschau und Vorsicht – daher bringt *W & F* in dieser Ausgabe erste vorsichtige Analysen, Einschätzungen und Perspektiven: auf Solidarität, Soziale Verteidigung, die Rückkehr des Militärischen, auf die Sozialpsychologie des Krieges und die rechtspazifistische Einschätzung der Sanktionsregime gegen Russland.

Der zweite Teil des Schwerpunkts zeigt die strukturelle Gewalt des globalen Gefüges auf. An anschaulichen Fallbeispielen diskutieren die Autor:innen Neokolonialismus als einen Begriff mit Zukunft: für ein Verständnis kriegerischer Verhältnisse sowie die Analyse andauernder kolonialer Machtverhältnisse in Wirtschaft, Gesellschaft und militärischen Belangen.

Mit Beiträgen von *Alexia Tsouni, Andreas Zumach, Werner Wintersteiner, Gracia Lwanzo Kasongo* und weiteren.

W&F 2/2022 | Mai | 64 Seiten | 12 € (print) / 9 € (epub) | wissenschaft-und-frieden.de





Stefan Hügel

BigBrotherAwards 2022

Auch im letzten Jahr gab es wieder reichlich Ereignisse, die einen BigBrotherAward verdient hätten. Wie immer berichten wir von der Preisverleihung: Wer hat es in den erlesenen Kreis der Preisträger:innen geschafft?

Wir fassen in diesem einleitenden Beitrag des Schwerpunkts zum BigBrotherAward 2022¹ zunächst die Laudationes für die Preisträger:innen kurz zusammen. Danach drucken wir zwei Laudationes im Wortlaut ab: die Laudatio *Behörden und Verwaltung* an die deutsche Polizei – vertreten durch das Bundeskriminalamt – für den Umgang mit den Daten von Bürger:innen und die Laudatio auf den Preis für das Lebenswerk an die irische Datenschutzaufsichtsbehörde für ihre Arbeitsverweigerung bei der Durchsetzung europäischen Datenschutzrechts.

Die Verleihung fand am 29. April 2022 in der Hechelei in Bielefeld statt.

Kategorie *Behörden und Verwaltung*

Der BigBrotherAward in der Kategorie *Behörden und Verwaltung* ging an die **Deutsche Polizei, vertreten durch das Bundeskriminalamt**,

für die Art, wie personenbezogene Daten in Dateien abgespeichert und genutzt werden.

Laudator Thilo Weichert – dessen vollständige Laudatio in dieser Ausgabe der *FIfF-Kommunikation* enthalten ist – führte dazu weiter aus:

Entgegen der verfassungs- und europarechtlichen Vorgaben werden die Daten in Dateien nicht oder unzureichend gekennzeichnet. Dadurch besteht die Gefahr, dass Millionen Menschen von der Polizei oder anderen Behörden ungerechtfertigter Weise als Gefährder oder Straftäter behandelt werden.

Er leitete seinen Vortrag beispielhaft mit den Ereignissen beim G 20-Gipfel 2017 in Hamburg ein, bei dem 32 Presseakkreditierungen mit der Begründung zurückgewiesen wurden, „Straftäter“ wollten sich damit Zutritt zum Gipfel „erschleichen“. Grundlage waren – bei makellosem polizeilichem Führungszeugnis – Einträge in Polizeidatenbanken in der Kategorie „politisch motivierte Kriminalität“, die inhaltlich fragwürdig und teilweise viele Jahre alt waren. Es handelt sich dabei um Einträge, die ohne Anklage oder gar Verurteilung abgelegt werden und

über das Informationssystem INPOL bundesweit abgefragt werden können. Einträge kann es als Kontaktperson, als Täter:in, als Opfer, als Zeug:in geben.

Um also zu verhindern, dass aus Einträgen, die über INPOL abgefragt werden, falsche Folgerungen gezogen werden, muss der Grund für den Eintrag erkennbar sein:

Es muss klar erkennbar sein, als was ich in der Datenbank geführt werde: als Kontaktperson, als Täterin, als Opfer, als Zeugin. Wurde gegen mich ermittelt und – weil an dem Vorwurf nichts dran war – das Verfahren eingestellt, so muss in der Datenbank meine Unschuld dokumentiert und ein Zugriff auf den Vorgang verhindert werden. Meine Daten müssen gekennzeichnet werden, damit ich nicht fälschlich festgehalten, durchsucht oder verhaftet werde – oder als Journalist:in meine Akkreditierung entzogen bekomme.

Aber:

Das Problem ist, dass in den veralteten Datenbanken der Polizei eine solche Kennzeichnung oder Markierung oft nicht vorgesehen war. Kurz vor der Verabschiedung des BKA-Gesetzes im Frühjahr 2017 ist das auch dem Bundestag aufgefallen, weshalb er als Übergangsbestimmung einen §91 einfügte. Demzufolge kann bei diesen Altsystemen auf die Datenmarkierung vorübergehend verzichtet werden. Ansonsten würde die „Funktionsfähigkeit der Polizei“ beeinträchtigt.

Dazu kommen weitere Beispiele. Das Bundeskriminalamt

... ist dafür verantwortlich, dass im polizeilichen Datenverbund die verfassungsrechtlich geforderten Maßnahmen nicht umgesetzt sind und deswegen weiterhin die Gefahr besteht, dass Polizeidaten unberechtigterweise gegen uns Bürger:innen zum Einsatz kommen.

Es erhält den BigBrotherAward dafür, dass es dieser Verantwortung nicht nachkommt.

Kategorie Technik

Frank Rosengart hielt die Laudatio in der Kategorie *Technik*. Dieser Preis ging an die **Bundesdruckerei GmbH**,

für die unsinnige Verwendung und Beförderung von Blockchain-Technik.

Konkret: Für das Leuchtturmprojekt Digitale Schulzeugnisse und die dafür verwendete Blockchain-Technologie mit ihren erheblichen datenschutzrechtlichen Auswirkungen:

Die Schulen erstellen mit einer speziellen Software eine digitale Version des Schulzeugnisses, welches über eine Datenschnittstelle an die Bundesdruckerei übermittelt wird. Dort wird daraus eine PDF-Datei erzeugt und dafür eine Prüfsumme berechnet. Die Prüfsumme ist ein mathematisches Verfahren, um nachzuweisen, dass ein Dokument nicht manipuliert wurde. Bei der Bundesdruckerei wird diese Prüfsumme dann in einer Blockchain gespeichert – und weil in einer Blockchain die jeweils angehängten Daten aufeinander aufbauen, ist eine nachträgliche Manipulation nur schwer machbar. Wenn ein solches Zeugnis zum Beispiel bei einer Hochschule zur Immatrikulation vorgelegt wird, kann bei der Bundesdruckerei nachgefragt werden, ob das Dokument echt ist. Dort wird in der Blockchain nachgeschaut, ob die Prüfsumme stimmt.

Frank Rosengart stellt in seiner Laudatio die Risiken der Blockchain für den Datenschutz dar:

Die große Gefahr in der Blockchain liegt darin, dass alle Eintragungen lückenlos nachvollziehbar sind. Das ist die Idee hinter der Blockchain. Eine nachträgliche Korrektur oder Löschung ist technisch nicht möglich. Alle Eintragungen werden aufeinander aufbauend wie in einer Kette (daher der Name) eingetragen und kryptographisch abgesichert.

Personenbeziehbare Daten in der Blockchain zu speichern, wäre ein Totalschaden für den Datenschutz, weil sich keine Daten (auch falsche oder widerrechtlich gespeicherte) löschen lassen. Ein Recht auf Vergessen ist nicht umsetzbar, die Daten wären für immer in der Blockchain gespeichert und damit auch öffentlich sichtbar.

Und begründet die Preisvergabe:

Aber die Verwendung von Blockchain-Technik für die Echtheitsprüfung von Zeugnissen ist vollkommen fehl am Platz: Anstelle von klassischen, digital signierten Dokumenten, die seit vielen Jahren technischer Standard sind und die von jeder/jedem mit einfacher Software geprüft werden können, drängt sich die Bundesdruckerei mit ihrem System als zentrale Prüfinstanz auf, die immer angefragt werden muss und an der kein Weg vorbei führt. Denn die Blockchain mit den Prüfsummen soll nicht öffentlich sein – es kann also niemand außer der Bundesdruckerei die Echtheit der Zeugnisse prüfen. Eigentlich genau das gegenteilige Konzept der Blockchain.

Kategorie Lebenswerk

Der BigBrotherAward in der Kategorie *Lebenswerk* wurde an die **Irische Datenschutzaufsichtsbehörde, vertreten durch Helen Dixon**, verliehen. Sie erhielt den Preis

für ihre dauerhafte Sabotage von Bemühungen, europäisches Datenschutzrecht durchzusetzen.

Laudatorin Rena Tangens – deren Laudatio in dieser Ausgabe der *FfF-Kommunikation* weiter unten abgedruckt ist – kritisiert:

Die irische Datenschutzaufsicht verhindert, dass geltendes Recht durchgesetzt wird – durch jahrelanges Verschleppen, de facto Nicht-Bearbeiten von Beschwerden, bürokratische Winkelzüge, abschreckende Kosten für Beschwerdeführer und mangelnde Kooperation mit den europäischen Kolleginnen und Kollegen. Behördenchefin Helen Dixon agiert erratisch und reagiert allergisch auf Kritik. Ihre Behörde lässt das europäische Datenschutzrecht ins Leere laufen – und das ausgerechnet gegenüber denen, die harte Kontrolle nötig hätten: Google, Facebook, Apple, Microsoft & Co.

Sie wirft der Behörde vor, Datenschutz auf vielfältige Weise zu verhindern:

- Beschwerden gegen die großen Digitalkonzerne liegen zu lassen,
- die Energie der Behörde in nichtssagende Public Relations und irreführende Tätigkeitsberichte zu investieren,
- Beschwerdeführer:innen mit unterschiedlichen Methoden von ihren eigenen Verfahren auszuschließen.

Abschließend kommt sie zu den Folgerungen, wie der irischen Behinderung des Datenschutzes begegnet werden kann:

1. *Das One-Stop-Shop-System ist offenbar dysfunktional und sollte grundlegend reformiert werden. Wenn ein einzelnes Land aus Eigennutz die Durchsetzung der Datenschutzgrundverordnung blockieren kann, dann ist das ein struktureller Fehler.*
2. *Sidestepping: Datenschutzbeschwerden einfach nicht mehr nach Irland weitergeben, sondern sich selbst für zuständig erklären und entscheiden. Einige EU-Länder, z. B. Frankreich, praktizieren das bereits hier und da. Betroffene werden auch zu Gerichten gehen statt zu Datenschutzbehörden, auch wenn das viel kosten kann.*
3. *Ein einheitliches europäisches Verfahrensrecht für grenzüberschreitende Datenschutzfälle einführen, das Fristen für die Bearbeitung und so weiter verbindlich festlegt .*
4. *Oder eine kompetente europäische Institution bestimmen, die die großen grenzüberschreitenden Fälle entscheidet. Das könnte zum Beispiel der Europäische Datenschutzausschuss (EDSA) sein, in dem die Datenschützer:innen der Mitgliedsländer seit langem vertrauensvoll zusammenarbeiten.*



Kategorie Arbeitswelt

Der BigBrotherAward in der Kategorie *Arbeitswelt* wurde von Peter Wedde angekündigt. Er wurde verliehen an **Lieferando** und den deutschen Betreiber dieses Angebots, die **yd.yourdelivery GmbH**

für die unzulässige Totalkontrolle ihrer beschäftigten „Rider“.

Peter Wedde begründet die Preisvergabe:

Lieferando erhält den BigBrotherAward 2022 für den Einsatz der Scoober-App, die eine umfassende Überwachung der für den Lieferdienst tätigen Fahrerinnen und Fahrer ermöglicht und die zugleich personenbezogene Daten an eine Reihe von Internet-Tracker weiterleitet.

Und weiter:

Lieferando ist aus unserer Sicht nur die Spitze eines Eisbergs von Firmen aus der „Plattform-“ oder „Gig-Economy“, die eine Tätigkeit davon abhängig machen, dass Beschäftigte ihnen vielfältige persönliche Daten zur Verfügung stellen. In Deutschland beziehen inzwischen fast sechs Prozent der Erwerbstätigen zumindest einen Teil ihres Einkommens aus sogenannter Plattformarbeit. Beschäftigte, die nicht aus Spaß für Firmen wie Lieferando arbeiten, sondern weil sie ihren Lebensunterhalt verdienen müssen.

Die App verarbeitet eine Reihe von Daten, die für den Prozess der Essenslieferung erforderlich sind – diese sind jedoch nicht der Grund für den BigBrotherAward. Der Preis wird verliehen, weil:

Nach einer Recherche des Bayerischen Rundfunks aus dem vergangenen Jahr werden pro Lieferung 39 Einzeldaten erhoben und verarbeitet. Dazu gehört neben Hinweisen zur verspäteten Ankunft in Restaurants oder bei Kunden insbesondere die permanente Erfassung der aktuellen Standorte der Fahrerinnen und Fahrer mittels GPS-Tracking. Sie erfolgt alle 15 bis 20 Sekunden. Die persönlichen Standortdaten erhält Lieferando.

Und:

Wir verleihen den Preis an Lieferando aber auch deshalb, weil diese Software eine Reihe von Internet-Tracker mit personenbezogenen Daten versorgt.

Laudator Peter Wedde abschließend:

Bleibt nur zu hoffen, dass der BigBrotherAward 2022 dazu beiträgt, dass die zuständigen Datenschutzaufsichtsbehörden sich mit der Verarbeitung von Beschäftigtendaten mit der Scoober-App schnell befassen.

Kategorie Gesundheit

Den letzten BigBrotherAward gab es an diesem Abend in der Kategorie *Verbraucherschutz*. Er ging an die **Klarna Bank AB** in Schweden,

die intransparent Daten und Macht als Shopping-Service, Zahlungsdienstleister, Preisvergleichsportal, persönlicher Finanzmanager, Bonitätskontrolleur und Bank bündelt.

Laudator padeluun beginnt mit dem Geschäftsmodell der Klarna Bank AB:

Klarna vermittelt als Dienstleister Zahlungen für Online-Stores. Sie kaufen den Läden Rechnungen ab – und nun muss die Kundin oder der Kunde den fälligen, ausstehenden Betrag an Klarna bezahlen (in der Welt des Handels heißt das „Factoring“).

Um das Zahlungsausfallrisiko zu vermindern, sammelt Klarna umfassend Daten über ihre Kunden:

Was könnte besser sein für die Bewertung der Bonität der willigen Käuferinnen und Käufer, als direkt oder indirekt auf ihre Kontobewegungen schauen zu können? Denn mal ganz ehrlich, Hand aufs Herz; wofür Menschen bereit sind Geld auszugeben und wie viel: das lässt in jeder Hinsicht tief blicken. Dann wirbelt man noch mit den üblichen Scoring-Algorithmen herum und teilt die Kaufwilligen in „gute“ und „schlechte“ Kunden (die Älteren unter Ihnen werden sich an den BigBrotherAward vor 21 Jahren erinnern, als wir diese menschenunwürdige Praxis von Beratungsunternehmen aufgedeckt haben) – und Simsalabim, Hokuspokus: das angeblich so schwierige Bezahlen im Internet lässt sich als „nun endlich ganz einfach“ anpreisen.

Dazu gibt es die Klarna-App, die sich als Browser darstellt, so auch unbemerkt zwischen Kund:in und Shop agiert und ihre Zahlungsdienstleistung anbietet. padeluun fordert statt solchen Dienstleistern, die für Zahlungen umfassend Daten sammeln, elektronisches Bargeld, mit dem anonym und für Händler:in und Kund:in sicher bezahlt werden kann.

Klarna dagegen hat ein unklares Geschäftsmodell, droht Kundinnen und Kunden mit überhöhten Mahngebühren und greift tief in die Trickkiste, um sich deren Privatsphäre zu bemächtigen. Personenbeziehbare Daten werden von Klarna nicht nur gesammelt, sondern auch gebündelt, aggregiert und großzügig weitergegeben.

Anmerkung

- 1 Weitere Informationen und Nachweise finden sich auf der Webseite der BigBrotherAwards, <http://www.bigbrotherawards.de>. Von dort stammen auch alle Zitate aus den Laudationes.



Kategorie *Behörden und Verwaltung* – Laudatio

Der BigBrotherAward in der Kategorie Behörden und Verwaltung 2022 geht an die deutsche Polizei, vertreten durch das Bundeskriminalamt,

für die Art, wie personenbezogene Daten in Dateien abgespeichert und genutzt werden. Entgegen der verfassungs- und europarechtlichen Vorgaben werden die Daten in Dateien nicht oder unzureichend gekennzeichnet. Dadurch besteht die Gefahr, dass Millionen Menschen von der Polizei oder anderen Behörden ungerechtfertigter Weise als Gefährder oder Straftäter behandelt werden.



Laudator Thilo Weichert – Foto: Matthias Hornung, CC BY-SA 4.0

Als im Juli 2017 in Hamburg der G 20-Gipfel stattfand, sollte eigentlich unabhängige Presse darüber berichten. Doch dazu kam es nur eingeschränkt. 32 Presseakkreditierungen wurden zurückgewiesen, weil „Straftäter“ – so die Bundesregierung – sich hierüber den Zutritt ins Innerste dieses Gipfels erschleichen wollten. Darunter ein 37-jähriger Fotograf, dessen Führungszeugnis zwar blütenweiß war, der aber 18 polizeiliche Dateneintragen vorwies, u. a. wegen „Herbeiführung einer Sprengstoffexplosion“ in der Kategorie „politisch motivierte Kriminalität“ – er hatte fotografiert, während in seiner Nähe ein Feuerwerkskörper explodiert war. Die Einträge lagen teilweise 14 Jahre zurück. Ein Online-Journalist wurde abgewiesen, weil gegen ihn die haltlose Anzeige eines vorbestraften Rechtsextremisten vorlag. Auch die anderen Ablehnungen von Journalist:innen erwiesen sich im Nachhinein als unbegründet.

Das Problem der Journalist:innen bestand darin, dass sie über politische Ereignisse berichtet hatten und dadurch in den Fokus der Polizei geraten waren. Die Polizei hatte sie vor Ort erfasst – z. B. ihren Presseausweis kontrolliert – und ihre Daten gespeichert. Kein harmloser Vorgang.

Denn schon das bloße Fotografieren von Polizist:innen kann unter dem Label „politisch motivierte Kriminalität“ oder „Verstoß gegen das Versammlungsgesetz“ abgelegt werden – ohne dass es je zu einer Anklage, geschweige denn zu einer Verurteilung gekommen ist. Diese Label befanden sich über viele Jahre in polizeilichen Datenbanken, die über das Informationssystem INPOL bundesweit miteinander verknüpft sind.¹

Der Sonnenstaat am Horizont

INPOL wird vom Bundeskriminalamt – dem BKA – betrieben und verbindet die Polizeibehörden des Bundes und der Länder elektronisch. Es geht zurück auf den legendären BKA-Präsidenten Horst Herold, der in den 1970er-Jahren dafür sorgte, dass das BKA massiv ausgebaut und mit moderner elektronischer Datenverarbeitung aufgerüstet wurde. Die Polizei wurde digitalisiert, zentralisiert und „intelligent“ gemacht. Herold hatte damals die Vision, dass die Polizei dank elektronischer Datenanalyse schon vor dem Straftäter am Tatort sein und so die Tat verhindern könne. Er sah am Horizont den „Sonnenstaat“, in dem Kriminalität weitgehend ausgemerzt sein würde. Das BKA galt weltweit als Vorreiter digitaler Polizeiermittlungen.²

Horst Herold wurde 1981 in den vorzeitigen Ruhestand geschickt. INPOL ist – nach einigen Überarbeitungen – bis heute in Betrieb. Inzwischen gibt es viele weitere Verbunddatenbanken beim BKA, u. a. die Anti-Terror-Datei und die Datei Rechtsextremismus – RED, an die auch die deutschen Geheimdienste angeschlossen sind. INPOL ist zwar in die Jahre gekommen, dient aber nach wie vor als Backbone für moderne Software. Die Analysemöglichkeiten digitaler Bestände sind unterdessen dank „Big Data“, „Data-Mining“ und sog. Künstlicher Intelligenz bei den Sicherheitsbehörden gewaltig gewachsen. Tatsächlich nutzt die Polizei inzwischen digitale Recherche- und Analysetools vom Feinsten. Und so ist es möglich, dass die Eintragung eines Polizeibeamten im digitalen Vorgangsverwaltungsprogramm in Coburg dazu beiträgt, dass die Bundesregierung einen Journalisten in Hamburg von der Berichterstattung beim G 20-Gipfel ausschließt.

Schmutzige Datenwäsche

Mit der Digitalisierung nicht mitgewachsen ist der Datenschutz – weder technisch noch rechtlich. Das hat immer wieder auch das Bundesverfassungsgericht festgestellt, u. a. mit seinem Urteil vom 20. April 2016.³ Darin attestiert das Gericht dem BKA-Gesetz von 2008, das den Datenverbund der Polizei regelt, in vielerlei Hinsicht nicht mit dem Grundgesetz im Einklang zu stehen. Ein Grund für die Verfassungswidrigkeit war, dass es keine Vorkehrungen gegen die Zweckänderung von Daten – ich nenne das Datenwäsche – gibt.

Datenwäsche funktioniert wie folgt: Bin ich z. B. von einer polizeilichen Telefonüberwachung betroffen – unabhängig davon, ob ich wirklich ein Krimineller bin – dann setzt das den Verdacht einer schweren Straftat voraus. Dank der digitalen Vernetzung sind aber die Daten, die bei der TKÜ (Telekommunikationsüberwachung) erlangt werden, theoretisch bei jeder Verkehrskontrolle vom Rechner im Polizeiauto abrufbar. Flugs stehe ich als potenzielle Schwerverbrecher:in vor den Streifenbeamten:innen.



Um diese Datenwäsche zu verhindern, muss der TKÜ-Datensatz als solcher markiert sein. Es muss klar erkennbar sein, als was ich in der Datenbank geführt werde: als Kontaktperson, als Täterin, als Opfer, als Zeugin. Wurde gegen mich ermittelt und – weil an dem Vorwurf nichts dran war – das Verfahren eingestellt, so muss in der Datenbank meine Unschuld dokumentiert und ein Zugriff auf den Vorgang verhindert werden. Meine Daten müssen gekennzeichnet werden, damit ich nicht fälschlich festgehalten, durchsucht oder verhaftet werde – oder als Journalist:in meine Akkreditierung entzogen bekomme.

Altsysteme und Neuprobleme

Die Rechtslage verlangt genau das.

Das BVerfG hat klar geregelt, wie Polizeidaten genutzt oder weitergegeben werden dürfen. In der Folge wurde auch im überarbeiteten BKA-Gesetz⁴ festgehalten, dass nicht gekennzeichnete Daten „so lange nicht weiterverarbeitet oder übermittelt werden (dürfen), bis eine Kennzeichnung ... erfolgt ist“.⁵

Auch die europäische Datenschutzrichtlinie für Polizei und Justiz, die seit 2018 in Kraft ist⁶, fordert, dass bei der polizeilichen Datenspeicherung klar ersichtlich sein muss, ob jemand Straftäter, Verdächtiger, Opfer, Zeuge, Hinweisgeber oder Kontaktperson ist. Ebenso muss erkennbar sein, ob eine Datenspeicherung nachgewiesene Fakten dokumentiert – oder vage Verdachte bzw. persönliche Einschätzungen.

Das Problem ist, dass in den veralteten Datenbanken der Polizei eine solche Kennzeichnung oder Markierung oft nicht vorgesehen war. Kurz vor der Verabschiedung des BKA-Gesetzes im Frühjahr 2017 ist das auch dem Bundestag aufgefallen, weshalb er als Übergangsbestimmung einen §91 einfügte. Demzufolge kann bei diesen Altsystemen auf die Datenmarkierung vorübergehend verzichtet werden. Ansonsten würde die „Funktionsfähigkeit der Polizei“ beeinträchtigt.⁷

Trotzdem hätte man vermuten dürfen, dass Anstrengungen unternommen werden, die Datenbanken der Polizei um entsprechende Kennzeichnungsmöglichkeiten zu erweitern. Dass nachträglich Markierungen vorgenommen werden, um den Anforderungen des BVerfG an den Datenschutz zu entsprechen. Doch das war und ist für die Polizei weiterhin viel zu viel Aufwand. Die alten Systeme wurden nicht nachgerüstet – stattdessen führt man neue Systeme ein, bei denen weiterhin die Gefahr besteht, dass z. B. Unschuldige als Gefährder behandelt werden.

Pfusch am Datenhausbau

Ein besonders abschreckendes Beispiel ist EASy GS bei der bayerischen Polizei, wo Stand Juni 2021 1.644 Fußballfans gespeichert sind, obwohl ihnen kein konkreter Vorwurf gemacht wird und ohne dass die meisten davon überhaupt etwas ahnen – allein auf Grund einer polizeilichen „Individualprognose“.⁸ Es genügt also schon der Verdacht, jemand könnte in Zukunft eine Straftat begehen, um in der Datei zu landen.

Das BKA vertröstet derweil auf seine Planungen für „Polizei 2020“, ein gemeinsames „Datenhaus“ von Bund und Ländern, das nur langsam vorankommt. 2018 wurde dazu ein 31-seitiges Whitepaper vom Bundesinnenministerium⁹ veröffentlicht, in dem die nötige Kennzeichnung nur einmal erwähnt ist. Wert wird allerdings darauf gelegt, „die permanente Verfügbarkeit der Altdaten im Transformationsprozess sicherzustellen“. Von Schutzmaßnahmen war schon damals keine Rede.

2020 beschwerte sich die Konferenz der Datenschutzbeauftragten, man müsse das „Datenhaus Polizei 2020“ „auch an datenschutzrechtlichen Kernforderungen“ ausrichten.¹⁰ Und ein Jahr später beklagte der Bundesbeauftragte für den Datenschutz in seinem Tätigkeitsbericht, dass sich der „fachliche Bebauungsplan“ für das „Datenhaus“ an polizeilichen Interessen ausrichte und die rechtlichen Rahmenbedingungen ignoriere.¹¹

Rendezvous mit VeRA

Heute, mehr als 6 Jahre nach dem Urteil des BVerfG, ist die Datenkennzeichnung immer noch nicht umgesetzt. Dass sich daran in näherer Zukunft etwas ändert, ist wenig wahrscheinlich. Im Gegenteil.

Im vergangenen Jahr wurde für das „Datenhaus“ vom bayerischen Landeskriminalamt eine Ausschreibung für ein „Verfahrensübergreifendes Recherche- und Analysesystem“ vorgenommen – abgekürzt mit dem schönen Namen „VeRA“:

„Die Kernkompetenz dieses Systems VeRA ist der direkte Zugriff, das Zusammenführen und Auswerten von Daten aus unterschiedlichen Quellen. Das System muss sowohl bereits vorhandene polizeiliche Datenbestände als auch externe Datenquellen verarbeiten können“.

„Externe Quellen“ – das kann praktisch alles sein. Bei Rasterfahndungen nach 9/11 wurden z. B. die Daten von Universitäten, dem Ausländerzentralregister und Meldebehörden zusammengeführt und abgeglichen, um sogenannte Schläfer zu identifizieren.

„Weitere Leistungsinhalte des ausgeschriebenen Systems umfassen insbesondere den Datenabgleich von internen und externen, strukturierten und unstrukturierten Datenbeständen“ – also Daten, die nach bestimmten Merkmalen wie Wohnort oder Beruf geordnet sind, genauso wie simple Texte – „zum Erkennen von Zusammenhängen innerhalb der Analysesoftware“.

Zusammenhänge? Nehmen wir als Beispiel Telefonverbindungsdaten, also wer hat wann mit wem telefoniert. Daraus können umfassende Kommunikations- und Beziehungsgeflechte erkannt werden, unabhängig davon, ob diese aus kriminellen oder aus ausschließlich persönlichen oder politischen Gründen bestehen. VeRA kommt auch meinen Amouren oder meinem politischen Netzwerk auf die Schliche.

Und schließlich zählt zum Anforderungskatalog „die Durchführung geografischer Auswertungen innerhalb des Systems und

die Visualisierung und der Export (mit Quellangaben) von Rechercheergebnissen sowie von Beziehungszusammenhängen zwischen Objekten.“

VeRA hat also auch graphisch aufbereitete Karten, Tortendiagramme oder Verbindungslinien im Repertoire, wie sie der Ermittler im Vorabendkrimi auf sein Flipchart malt.

Kurzum: die Wundersoftware VeRA kann praktisch alles, was sich mit Daten machen lässt.

Zum Datenschutz heißt es im umfangreichen Anforderungskatalog kurz und lapidar: „Das System VeRA muss den geltenden rechtlichen, insbesondere datenschutzrechtlichen Bestimmungen entsprechen.“¹²

Willkommen im Palantir-Reich

Am 7. März 2022 gab dann das Bayerische LKA stolz bekannt, dass den Zuschlag für VeRA die Firma Palantir Technologies GmbH bekommen hat, die Tochter des US-Unternehmens Palantir. Kein anderes Unternehmen habe die „sehr strengen Ausschreibungskriterien“ erfüllt. Als Beleg für „höchste Ansprüche an Datenschutz und Datensicherheit“ dient dabei, dass es bei VeRA von Palantir „keine Verbindung zum Internet“ geben werde. Von der umfassenden Kennzeichnung der Datensätze ist einmal mehr keine Rede.¹³

Bei Palantir dürfte Datenschutz eher unbekannt sein: Der US-Mutterkonzern war für US-Geheimdienste und das Pentagon tätig und wird von Bürgerrechtler:innen und Datenschützer:innen immer wieder kritisiert. Gegründet wurde Palantir von dem umstrittenen Tech-Milliardär Peter Thiel, der rechtsextreme Politiker:innen – darunter den ehemaligen US-Präsidenten Donald Trump – finanziell unterstützt.

In Hessen und Nordrhein-Westfalen hat die Polizei in der Vergangenheit schon Erfahrungen mit Palantir-Software gesammelt. Kritik an der Beschaffung und den Funktionen einer solchen hat Ende 2018 in Hessen für einen Untersuchungsausschuss gesorgt – und, nicht zu vergessen, dem CDU-Innenminister Peter Beuth einen Big Brother Award beschert.

So kommt es nicht von ungefähr, dass der bayerische Landesbeauftragte für Datenschutz, Thomas Petri, anlässlich der VeRA-Vergabe von einem massiven Eingriff in die Grundrechte von Millionen Menschen sprach.¹⁴

Gefahr ohne Verzug

Der BigBrotherAward in der Kategorie Verwaltung geht aber vorrangig nicht an das Bayerische LKA, sondern an das BKA. Der Grund dafür steht in §31 Abs. 1 BKAG: „Das Bundeskriminalamt hat als Zentralstelle für den polizeilichen Informationsverbund die Einhaltung der Regelungen zur Zusammenarbeit und zur Führung des Verbundsystems zu überwachen.“

Es ist das BKA, das die Gesamtverantwortung für „Polizei 2020“ trägt. Es hat die Gesamtverantwortung für INPOL sowie für viele weitere Systeme, bei denen die Kennzeichnungspflicht nicht umgesetzt ist – und letztlich auch für das Data-Mining-System VeRA.

Es ist dafür verantwortlich, dass im polizeilichen Datenverbund die verfassungsrechtlich geforderten Maßnahmen nicht umgesetzt sind und deswegen weiterhin die Gefahr besteht, dass Polizeidaten unberechtigterweise gegen uns Bürger:innen zum Einsatz kommen.

Daher herzlichen Glückwunsch für den BigBrotherAward 2022 in der Kategorie Verwaltung: Bundeskriminalamt.

Anmerkungen

- 1 Viele BKA-Datenspeicherungen rechtswidrig, *DatenschutzNachrichten (DANA) 4/2017*, 206 f.
- 2 *Transatlantik 11/1980*, 38; vgl. Weichert, *Informationelle Selbstbestimmung und strafrechtliche Ermittlung*, 1990, S. 6 ff.
- 3 *BVerfG U.v. 20.04.2016 – 1 BvR 966/09 u. 1 BvR 1140/09*, NJW 2016, 1781 ff. = NVwZ 2016, 839 ff. = DVBl 2016, 770 ff. = EuGRZ 2016, 149 ff. = K&R 2016, 395 ff. = CR 2016, 796 ff. = WM 2016, 1133 ff. = BB 2016, 1089 ff. = AnwBl 2016, 516 ff. = DÖV 2016, 530 ff.
- 4 BGAK v. 01.06.2017, BGBl. I S. 1354; zuletzt geändert durch G.v. 25.06.2021, BGBl. I S. 2099.
- 5 §14 Abs. 2 BKAG.
- 6 Richtlinie (EU) 2016/680 v. 27.04.2016, ABl. EU v. 04.05.2016, L 119/89.
- 7 BT-Drs. 18/12141, 6.
- 8 Patrick Kleinmann, *Geheime Fan-Datenbank in Bayern: 1644 Fragezeichen*, 18.08.2021, <https://www.kicker.de/geheime-fan-datenbank-in-bayern-1644-fragezeichen-868760/artikel>.
- 9 Bundesministerium des Innern, *Polizei 2020 – White Paper*, https://www.bmi.bund.de/SharedDocs/downloads/DE/veroeffentlichungen/2018/polizei-2020-white-paper.pdf?__blob=publicationFile&v=5.
- 10 DSB-Konferenz, *Entschließung v. 16.04.2020, Polizei 2020 – Risiken sehen, Chancen nutzen!*, https://www.bfdi.bund.de/SharedDocs/Downloads/DE/DSK/DSKEntschliessungen/99DSK-Polizei2020.pdf?__blob=publicationFile&v=2.
- 11 Bundesbeauftragter für den Datenschutz und die Informationsfreiheit, *29. Tätigkeitsbericht 2020, 2021*, Kap. 6.1 (S. 55 f.), https://www.bfdi.bund.de/SharedDocs/Downloads/DE/Taetigkeitsberichte/29TB_20.pdf?__blob=publicationFile&v=5.
- 12 *Germany-Munich: Database software package, 2021/S 011-023694*, Contract notice, 18.01.2021, <https://ted.europa.eu/udl?uri=TED:NOTICE:23694-2021:TEXT:EN:HTML&src=0>.
- 13 Bayerisches Landeskriminalamt, *Noch erfolgreichere Polizeiarbeit – Zuschlag für neues Recherche- und Analysesystem der Bayerischen Polizei: Höchste Ansprüche an Datensicherheit und Datenschutz*, PE v. 07.03.2022, <https://www.polizei.bayern.de/aktuelles/pressemitteilungen/025971/index.html>.
- 14 Knobloch, *Bayerns LKA will umstrittene Palantir-Software einsetzen*, 07.03.2022, [Kurzlink: https://heise.de/-6541763](https://heise.de/-6541763).

Kategorie *Lebenswerk* – Laudatio

Ein BigBrotherAward 2022 geht an die irische Datenschutzaufsicht, kurz DPC (Data Protection Commission), vertreten durch ihre Chefin Helen Dixon,

für ihre umfassende Sabotage des europäischen Datenschutzrechts. Und weil die irische Datenschutzaufsicht das so planvoll, seit so vielen Jahren und mit solch kafkaesker Phantasie betreibt, reicht dafür die Kategorie *Behörden und Verwaltung* nicht aus. Dafür gibt es den Preis fürs Lebenswerk.

Die irische Datenschutzaufsicht verhindert, dass geltendes Recht durchgesetzt wird – durch jahrelanges Verschleppen, de facto Nicht-Bearbeiten von Beschwerden, bürokratische Winkelzüge, abschreckende Kosten für Beschwerdeführer und mangelnde Kooperation mit den europäischen Kolleginnen und Kollegen. Behördenchefin Helen Dixon agiert erratisch und reagiert allergisch auf Kritik. Ihre Behörde lässt das europäische Datenschutzrecht ins Leere laufen – und das ausgerechnet gegenüber denen, die harte Kontrolle nötig hätten: Google, Facebook, Apple, Microsoft & Co.

Arbeitsverweigerung

Die Arbeitsverweigerung der irischen Datenschutzaufsicht betrifft dabei nicht nur Menschen, die in Irland leben, sondern sie gefährdet die Persönlichkeitsrechte von 450 Millionen EU-Bürgerinnen und Bürgern.

Rückblende: Mai 2018 – die europäische Datenschutzgrundverordnung tritt in Kraft. Sie stellt klar: 1. Egal, aus welchem Land ein Unternehmen kommt – sobald es persönliche Daten von EU-Bürger:innen verarbeitet, muss es sich an die in der EU geltenden Datenschutzregeln halten. Das ist das „Marktortprinzip“. 2. Bei Datenschutzvergehen drohen nun endlich empfindliche Bußgelder. „Empfindlich“ heißt: bis zu 4 % des weltweit erzielten Jahresumsatzes. Damit wurde Datenschutz bei den Konzernen zur Chefsache. Yay!

Doch leider haben wir uns zu früh gefreut: Denn ein europäisches Gesetz zu machen reicht nicht – seine Einhaltung muss auch in allen EU-Staaten durchgesetzt werden. Und da gibt es ein Problem. Es heißt: Irland.

Weshalb?

Jedes Unternehmen muss eine federführende Datenschutzaufsichtsbehörde festlegen, und zwar in dem Land, in dem das Unternehmen seine Hauptniederlassung in der EU hat. Diese federführende Aufsichtsbehörde ist fürderhin die Hauptansprechadresse für alle Beschwerden gegen dieses Unternehmen. Wenn sich Bürger:innen oder Organisationen über die Datenverarbeitung einer Firma beschweren wollen, tun sie das bei der Datenschutzaufsicht in ihrem eigenen Land. Diese leitet die Beschwerde dann an die zuständige Datenschutzaufsicht am Hauptsitz des beklagten Unternehmens weiter. Die Regelung, dass die Datenschutzaufsicht eines Unternehmens jeweils in den Händen nur eines EU-Landes liegt, heißt „One-Stop-Shop“.

Das klingt erst einmal praktisch – da wird Kompetenz gebündelt und die Aufsichtsbehörde kennt ihre Pappenheimer vor Ort und muss sich nicht bei jeder Beschwerde neu einarbeiten.

Doch es gibt einen Haken: Die großen Digitalkonzerne mit ihren immateriellen Geschäften können ihren „Hauptsitz“ relativ flexibel handhaben. Entsprechend suchen sie sich eine ihnen genehme Datenschutzaufsicht und eröffnen dann dort ihren vorgeblichen Hauptsitz.

Hier fällt eine gewissen Ballung ins Auge – Irland: Home of Google, Apple, Facebook und WhatsApp, Microsoft und LinkedIn, Adobe, Tiktok, Airbnb, Tinder, Twitter, Dropbox, Yahoo und so weiter.

Tja, was könnte wohl die Ortswahl der Digitalkonzerne beeinflussen?

Das grüne Gras, die weiten Wiesen, das dunkle Bier, ... oder liegt es ganz vielleicht doch an den Eigenheiten der Datenschutzaufsicht in Irland?

Schauen wir uns das mal genauer an:

Welche Bedeutung Irland dem Datenschutz gibt, das dokumentiert eindrucksvoll das Dienstgebäude der irischen Datenschutzaufsicht: Ein Büro im ersten Stock über einem kleinen Supermarkt in der Kleinstadt Portarlington – mehr als 70 Kilometer südwestlich von Dublin, jott we de. Das Foto der Datenschutzhilfe mit Supermarkt belustigt seit vielen Jahren die Presse in ganz Europa. Ja, inzwischen ist der Centra Supermarkt einem Spar gewichen und es ist immerhin einmal neu gestrichen worden.

2017 wurde ein zusätzlicher Behördensitz ganz repräsentativ mitten in der Hauptstadt Dublin eröffnet. Keineswegs aber, um der gestiegenen Bedeutung des Datenschutzes in Europa gerecht zu werden, sondern – Zitat eines Sprechers der Behörde – „um besseren Kontakt zu den multi-nationalen Unternehmen zu halten, die sich in Dublin angesiedelt haben“(!).¹

Kreative Behörde

Die irische Datenschutzaufsicht gibt sich jede erdenkliche Mühe, das von Big Tech in sie gesetzte Vertrauen nicht zu enttäuschen. Und dafür lassen sie sich ganz schön viel einfallen:

Wichtigste Taktik: Beschwerden gegen die großen Digitalkonzerne einfach liegen lassen

Fälle schlicht nicht bearbeiten.

Ein paar Zahlen zur Illustration: Seit Geltungsbeginn der DSGVO im Mai 2018 hat Ulrich Kelber, der deutsche Bundesbeauftragte

für den Datenschutz und die Informationsfreiheit (BfDI), etwa 50 Verfahren zuständigshalber an die irische Datenschutzaufsicht abgegeben – überwiegend Verfahren gegen WhatsApp. Von diesen 50 Verfahren ist nicht eines durch eine inhaltliche Entscheidung abgeschlossen worden.

Laut Gesetz müssen Beschwerden „unverzüglich“ bearbeitet werden. Die irische Datenschutzaufsicht argumentiert gerade vor dem irischen High Court, dass vier Jahre Bearbeitungszeit immer noch „unverzüglich“ seien ...

Den enormen Rückstau bei der Bearbeitung der Beschwerden erklärt die irische Datenschutzaufsicht damit, dass sie zu wenig Geld und zu wenig Personal hätte.

Doch die Behörden großer Mitgliedsstaaten wie Spanien oder Frankreich verfügen über ähnliche Budgets wie das kleine Irland. Die spanische Datenschutzaufsicht hat etwa das gleiche Budget und haut mehrere Entscheidungen pro Tag raus – die Iren schaffen nur ein paar pro Jahr. Es liegt also nicht am Geld, sondern an der Effizienz. Die Mitarbeiterzahl ist in Irland inzwischen von rund 30 (im Jahr 2014) auf 195 (im Jahr 2021) aufgestockt worden – leider ohne dass das die Arbeitsergebnisse verbessert hätte. Die Aufsichtsbehörden anderer Länder, u. a. Deutschland, haben den irischen Kolleg:innen schon vor Jahren Hilfe bei der Bearbeitung der Fälle angeboten. Doch Helen Dixon hat alle Hilfsangebote ausgeschlagen.

Zweite Taktik: Statistik-Bullshit verbreiten

Viel Energie der Behörde wird offenbar in Public Relations und hübsch gestaltete Tätigkeitsberichte gesteckt, um das Nichtstun gut aussehen zu lassen. Im Tätigkeitsbericht für das Jahr 2021² steht, dass 626 von 969 grenzüberschreitenden Beschwerden, die seit Mai 2018 eingegangen seien, abgeschlossen wurden.³ Aber Moment mal: Was bedeutet „abgeschlossen“? Offenbar nicht, dass die Fälle bearbeitet und entschieden wurden.

Ein weiterer Fall von Zahlen-Kosmetik: Zuvor hatte die irische Datenschutzaufsicht rund 10.000 „cases“ (Fälle) pro Jahr auf dem Tisch. Doch im Tätigkeitsbericht 2021 gibt es plötzlich nur noch rund 3.400 „complaints“ (Beschwerden), dazu aber rund 7.500 „inquiries“ (Anfragen). Der Trick: Alles, was bei der Einreichung nicht explizit „Beschwerde“ genannt wurde, fällt unter „Anfragen“. Anfragen landen im Papierkorb. Auch das ist natürlich eine Art, „sich damit befassen“ und „abzuschließen“: Ablage P. Erledigung durch Umbenennung.

Taktik Nummer drei: Ausschluss der Beschwerdeführer:innen vom Verfahren

Dafür gibt es eine ganze Reihe von Tricks, zum Beispiel:

a) Die Erpressungs-Methode:

Die Behörde fordert vom Kläger die Unterzeichnung einer Verschwiegenheitserklärung (englisch „Non Disclosure Agreement“, kurz NDA). Das heißt, er müsste Stillschweigen über die Verhandlung, die Verhandlungsergebnisse und die dort vorgelegten Informationen wahren. So geschehen mit

Johnny Ryan vom Irish Council for Civil Liberties bei seiner Klage gegen Googles Real Time Bidding Werbeauktionen. So geschehen mit Max Schrems von noyb (*none of your business*) bei seiner Klage gegen Facebook. Damit will die irische Datenschutzaufsicht dafür sorgen, dass alles hübsch intern und unterm Deckel gehalten wird. So etwas kann eine klagende Verbraucher- und Bürgerrechtsorganisation einfach nicht unterschreiben! Nach Nichtunterzeichnung wurde Max Schrems von seinem eigenen Verfahren ausgeschlossen. Was allen EU-Grundrechten widerspricht.

b) Die Umgehungsmethode:

Die irische Behörde leitet ein paralleles „amtsseitiges Verfahren“ zum selben Thema wie eine Beschwerde ein. Fortan wird nur noch dieses amtsseitige Verfahren bearbeitet – und solange lässt sie das Verfahren des Beschwerdeführers ruhen. Nach der Entscheidung des amtsseitigen Verfahrens wird dann das andere beendet, denn das Thema ist ja jetzt erledigt! Chapeau – eine wirklich elegante Art, die Bürger:innen von ihren eigenen Verfahren auszuschließen.

c) Die „Klag-doch-wenn-du-es-dir-leisten-kannst“-Methode:

Bei Untätigkeit der irischen Datenschutzbehörde den Rechtsweg einzuschlagen und zu klagen, ist eine so teure Angelegenheit, dass sie für Privatleute quasi nicht in Frage kommt. Dazu trägt u. a. die irische Regel bei, dass jedes Gesetz, auf das man sich bezieht, nicht nur als „nach Paragraph xy“ benannt werden, sondern vor Gericht vorgelesen werden muss. Das dauert nicht nur lang – sondern wird bei den bis zu 1000 Euro, die Anwaltskanzleien pro Stunde kassieren, auch sehr, sehr teuer für die Klagenden. Ein Beispiel: Im *Privacy-Shields-Fall* („Schrems II“) ging es um die Datenübermittlung von Facebook in die USA. Bei dem Fall gab es drei Parteien (Facebook, die Irische Datenschutzaufsicht und Max Schrems). Wer verliert, muss die Kosten von allen drei Parteien zahlen – die beliefen sich in diesem Fall auf insgesamt rund 10 Millionen Euro(!). Was für ein Glück, dass Max Schrems gewonnen und die Datenschutzbehörde verloren hat.

Soviel zur Kreativität der irischen Behörde – nun zu den Kopfnoten

Mangelnde Kollegialität, kein europäischer Geist, Geheimniskrämerei:

Zu all dem passt, wie Behördenchefin Dixon agiert. Sie nimmt an fast keiner gemeinsamen Sitzung der europäischen Datenschutzbeauftragten teil. Sie schickt meist einen Stellvertreter, der dann aber nichts sagen kann oder darf. Die Kommunikation auf der Leitungsebene ist damit schon mal tot. Das setzt sich auf der Sachbearbeitungsebene fort: Anfragen per E-Mail von deutschen oder österreichischen Kolleg:innen werden oft nicht beantwortet, Telefonanrufe nicht angenommen. Mitarbeiter:innen von anderen Datenschutzbehörden werden von den Iren gern „geghostet“, also komplett ignoriert, Akten nicht an die europäischen Kolleg:innen übermittelt. Die irische Behörde wirkt wie ein „schwarzes Loch“, in dem alles verschwindet.

Ihre Arbeitsverweigerung hat aber eben nicht nur für Menschen in Irland Folgen, sondern für 450 Millionen Menschen in der EU, deren Rechte von den großen Digitalkonzernen mit Füßen getreten werden. Das Gebaren der irischen Datenschutzaufsicht führt dazu, dass europaweit kleine und mittelständische Firmen bei Datenschutzvergehen von ihrer nationalen Aufsichtsbehörde sanktioniert werden; die großen Digitalkonzerne aber zeigen uns allen eine lange Nase: „Ätsch – beschwert euch doch – wir sind in Irland.“

Nun fragen wir uns: Warum tun die das?!

Durch all diese Winkelzüge und miesen Tricks macht die irische Datenschutzbehörde Irland zum Datenschutz-Freihafen, zum Schlupfloch für Verbrecher, zum Reservat für Datenkraken. All das fügt sich bestens in einen anderen Teil des inselgrünen Ökosystems ein:

Die Steueroase

Sie denken bei Steueroasen an die Cayman Inseln oder die Bahamas? Lenken Sie lieber Ihren Blick auf Irland. Offiziell gelten in Irland 12,5 % Unternehmenssteuern. Doch Irland hat es einigen ausländischen Konzernen ermöglicht, die effektiven Unternehmenssteuern für ihre globalen Gewinne auf 0 bis 2,5 % zu drücken. Die Steuertricks haben so phantasievolle Namen wie „Double Irish“, „Double Irish with a Dutch sandwich“ oder „Single Malt“. Über Irlands Steuersparmodelle werden mehr Gelder der Steuer (und damit der Allgemeinheit) entzogen als in der gesamten Karibik.⁴

Gesetzlosigkeit als Geschäftsprinzip?

Für Irland sind das lukrative Geschäfte. Es profitiert davon, dass es den Big-Tech-Konzernen ermöglicht, ihren Überwachungskapitalismus ohne Rücksicht auf Bürgerrechte durchzuziehen. Also persönliche Daten zu sammeln, zu Profilen zu verknüpfen, Kategorien zu bilden, Menschen zu manipulieren und ihnen durch den Verkauf von Prognosen viel Handlungsfreiheit für die Zukunft zu nehmen. Irland lebt von den Brosamen von Big Tech. Und es lebt gut davon, denn auch nur 2,5 % des globalen Umsatzes von Apple ist schon eine verdammt große Menge Geld.

Wir alle zahlen dafür mit unserer Freiheit.

Doch dieses Geschäftsgebaren hat eine Kehrseite: **Irland ist extrem abhängig von den Tech-Konzernen.** Ein paar Zahlen:

- 2016–17 haben ausländische Firmen 80 % der irischen Unternehmenssteuern gezahlt.
- 25 der Top-50-Unternehmen in Irland sind US-kontrolliert.
- 2018 machte Apple alleine ein Fünftel des irischen Bruttoinlandsproduktes aus.⁵

Ein Abgeordneter des irischen Parlaments, der 2017 auf das schändliche Treiben der Konzerne mit dem Steuervermeidungsmodell „Single Malt“ hinwies, bekam vom irischen Finanzminis-

ter den guten Rat „to put on the green jersey“ (das grüne Trikot anziehen) – mit anderen Worten: Zum Wohle Irlands die Klappe halten.⁶ Der Finanzminister machte sich also keine Sorgen wegen des Steuerbetrugs, sondern mehr wegen Irlands Reputation.

2023 kommt die **globale Mindeststeuer von 15 %** – und eine Digitalsteuer. Werden die Digitalkonzerne Irland verlassen, wenn der Steuervorteil versiegt ist? Oder reicht es ihnen, wenn auf der grünen Insel die Durchsetzung des Datenschutzes blockiert wird?

Darauf hofft offenbar der irische Premierminister Micheál Martin. Er lobte im Februar 2022 ausdrücklich „Ms. Dixons Kompetenz und Fähigkeiten“ und forderte, dass Irland die Arbeit ihrer Datenschutzbehörde „robuster gegen Kritik verteidigen“ solle.⁷

Wie lange wollen wir das noch zulassen?

Leider ist das grüne Trikot offenbar auch bei denen beliebt, die die Einhaltung der europäischen Gesetze durchsetzen sollen. Es wäre nämlich die Aufgabe der Europäischen Kommission – genau gesagt: der Generaldirektion Justiz und Verbraucher –, die irische Datenschutzaufsicht wirksam auf Vollzug des europäischen Datenschutzrechts zu kontrollieren. Tut sie aber leider seit Jahren nicht. Team Irland. Seriously?

Doch Obacht: Allmählich braut sich etwas zusammen in Europa. Denn irgendwann reicht es auch hier den diplomatischsten, freundlichsten und geduldigsten Datenschützer:innen und Politiker:innen.

Im Januar 2021 beantragt der LIBE-Ausschuss des Europäischen Parlaments (das ist der Ausschuss für bürgerliche Freiheiten, Justiz und Inneres), **ein EU-Vertragsverletzungsverfahren gegen Irland einzuleiten**, weil Irland die europäische Datenschutzgrundverordnung nicht ordnungsgemäß durchsetzt.⁸

Auftritt Helen Dixon: Sie fordert, dass sie dazu im LIBE-Ausschuss angehört wird. Der LIBE-Ausschuss setzt tatsächlich eine Anhörung für den 17. März 2021 an. Nun will Dixon sogar das Verfahren vor dem europäischen Parlament bestimmen – und nicht kommen, wenn auch Kritiker wie Max Schrems geladen werden. Damit allerdings blitzt sie bei den EU-Parlamentariern ab – der Ausschuss für bürgerliche Freiheiten lässt sich das Procedere seiner Anhörungen nicht diktieren. Daraufhin erscheint Helen Dixon nicht zu der Anhörung, die sie selbst verlangt hat. Großes Kino.

Kein Wunder, dass selbst in Irland manche mittlerweile Rot sehen. Wie der Justizausschuss des irischen Parlaments, der im Juli 2021 eine grundlegende Reform der irischen Datenschutzaufsicht gefordert hat. Wobei auch dieser Schritt zur Einsicht wohl nicht passiert wäre ohne die Initiative und das Engagement von Bürgerrechtsorganisationen, die unsere Rechte gegenüber den Digitalkonzernen durchzusetzen versuchen. Unser Dank gilt beharrlichen Menschen wie Max Schrems von noyb und Johnny Ryan vom Irish Council for Civil Liberties (ICCL).

Im März 2022 hat der High Court, der höchste Gerichtshof in Irland, die Klage des ICCL gegen die irische Datenschutzaufsicht wegen Untätigkeit zugelassen. Dixons Behörde hatte die Be-

schwerde des ICCL wegen Googles Werbeauktionen – das sogenannte *Real Time Bidding* – vier Jahre lang nicht bearbeitet. Stichwort: „unverzögliche Erledigung“ ... Zum Thema Googles Real Time Bidding empfehle ich meine Laudatio zum BigBrotherAward für Google von 2021. ☺

Das ICCL hat außerdem die EU-Kommission aufgefordert, ein Vertragsverletzungsverfahren einzuleiten. Im Februar 2022 reicht EU-Ombudsfrau Emily O'Reilly die Beschwerde wegen Untätigkeit der irischen Datenschutzbehörde an Ursula von der Leyen weiter und fordert Antwort bis 15. Mai 2022. Ausgang offen.

Was ist zu tun?

4. Das One-Stop-Shop-System ist offenbar dysfunktional und sollte grundlegend reformiert werden.⁹ Wenn ein einzelnes Land aus Eigennutz die Durchsetzung der Datenschutzgrundverordnung blockieren kann, dann ist das ein struktureller Fehler.
5. Sidestepping: Datenschutzbeschwerden einfach nicht mehr nach Irland weitergeben, sondern sich selbst für zuständig erklären und entscheiden. Einige EU-Länder, z.B. Frankreich, praktizieren das bereits hier und da. Betroffene werden auch zu Gerichten gehen statt zu Datenschutzbehörden, auch wenn das viel kosten kann.
6. Ein einheitliches europäisches Verfahrensrecht für grenzüberschreitende Datenschutzfälle einführen, das Fristen für die Bearbeitung und so weiter verbindlich festlegt.
7. Oder eine kompetente europäische Institution bestimmen, die die großen grenzüberschreitenden Fälle entscheidet. Das könnte zum Beispiel der Europäische Datenschutzausschuss (EDSA)¹⁰ sein, in dem die Datenschützer:innen der Mitgliedsländer seit langem vertrauensvoll zusammenarbeiten.

Klingt nach Arbeit – aber Politik ist nun mal das „starke langsame Bohren von harten Brettern“.¹¹

Was auch immer gemacht wird – eins ist klar: An den entscheidenden Stellen brauchen wir Menschen mit Motivation zum Ermitteln, der Beharrlichkeit, Fälle auszufeuchten und dem Willen, Datenschutz und die Persönlichkeitsrechte der Bürgerinnen und Bürger durchzusetzen.

Nicht die Bürokratie ist das Problem, sondern die Menschen, die sich hinter ihr verstecken.

Liebe Helen Dixon,

dies ist keine Anfrage, keine Beschwerde, keine Klageeinreichung, kein Zwischenbescheid – das ist ein BigBrotherAward.

Die BigBrotherAwards sind unabhängig. Wir bekommen kein Geld vom Staat, wir nehmen kein Sponsoring von Google, Facebook & Co. – die BigBrotherAwards leben durch die privaten Spenden von Tausenden von Menschen, die unsere Arbeit unterstützen. Wie wichtig unsere Unabhängigkeit ist, das zeigt uns genau dieser BigBrotherAward.

Herzlichen Glückwunsch Helen Dixon und der ganzen irischen Datenschutzaufsicht zum Preis fürs Lebenswerk.

Congratulations, Ms. Dixon and the Irish Data Protection Commission – the BigBrotherAward 2022 for Lifetime achievement is yours.

Anmerkungen

- 1 *The Journal*, 3.5.2014: *The Data Protection Commissioner is getting a new office, but keeping the one beside a convenience store in Laois*
<https://www.thejournal.ie/data-protection-commissioner-new-office-1488473-May2014/>
- 2 *Aktivitätsbericht 2021*: https://www.dataprotection.ie/sites/default/files/uploads/2022-02/Data%20Protection%20Commission%20AR%202021%20English%20FINAL_0.pdf
- 3 Zitiert vom *Irish Council for Civil Liberties* in einem Brief an die Data Protection Commission:
<https://www.iccl.ie/wp-content/uploads/2022/03/letter-to-Helen-Dixon-from-Liam-Herrick-2-March-2022.pdf>
- 4 Zitat Wikipedia: *Ireland's base erosion and profit shifting (BEPS) tools give some foreign corporates Effective tax rates of 0% to 2.5% on global profits re-routed to Ireland via their tax treaty network. (...) Ireland's BEPS tools are the world's largest BEPS flows, exceed the entire Caribbean system,*
Quelle: https://en.wikipedia.org/wiki/Ireland_as_a_tax_haven
- 5 Quelle Wikipedia: https://en.wikipedia.org/wiki/Corporation_tax_in_the_Republic_of_Ireland
- 6 Zitat aus der Parlamentsdebatte: *It was interesting that when Matt Carthy put that to the Minister's predecessor, his response was that this was very unpatriotic and he should wear the green jersey. That was the former Minister's response to the fact there is a major loophole, whether intentional or unintentional, in our tax code that has allowed large companies to continue to use the double Irish. The Minister's predecessor has acknowledged the reputational damage this has done to Ireland. He was not really concerned about losing tax revenue and all the rest, but about the reputational damage. Let there be no doubt that, as we close one loophole and create another door, or do not close the door, this reputational damage is going to continue.*
Quelle: *Dáil Éireann debate – Thursday, 23 Nov 2017, Vol. 962 No. 2.*
<https://www.oireachtas.ie/en/debates/debate/dail/2017-11-23/18/>
- 7 *The Irish Times*, 22.2.2022: *Taoiseach defends Irish data protection commissioner in Germany – Martin doesn't 'readily agree' with many criticisms of Helen Dixon's record,*
<https://www.irishtimes.com/business/technology/taoiseach-defends-irish-data-protection-commissioner-in-germany-1.4809442>
- 8 LIBE Ausschuss – Antrag auf Vertragsverletzungsverfahren vom 13.1.2021: *Draft Motion for a resolution – B9-0000/2021 European Parliament Resolution on the ruling of the ECJ of 16 July 2020 – Data Protection Commissioner v Facebook Ireland Limited, Maximilian Schrems ("Schrems II") – Case C-311/18 (2020/2789(RSP))*
- 9 *Im gerade fertiggestellten Digital Services Act (DSA) hat man die Lehren aus der DSGVO gezogen und die Aufsicht anders geregelt.*
- 10 engl. *European Data Protection Board, EDPB – das ist die Organisation, die früher „Artikel-29-Arbeitsgruppe“ hieß.*
- 11 Zitat Max Weber: *„Die Politik bedeutet ein starkes langsames Bohren von harten Brettern mit Leidenschaft und Augenmaß zugleich.“*
https://de.wikipedia.org/wiki/Max_Weber#cite_note-169

Im FIF 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 FIF 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 FIF-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 FIF regelmäßig an weiteren Veranstaltungen, Publikationen, vermittelt bei Presse- oder Vortragsanfragen ExpertInnen, führt Studien durch und gibt Stellungnahmen ab etc. Das FIF kooperiert mit zahlreichen Initiativen und Organisationen im In- und Ausland.

FIF-Mailinglisten

FIF-Mailingliste

An- und Abmeldungen an:

<http://lists.fiff.de/mailman/listinfo/fiff-L>

Beiträge an: fiff-L@lists.fiff.de

FIF-Mitgliederliste

An- und Abmeldungen an:

<http://lists.fiff.de/mailman/listinfo/mitglieder>

Mailingliste Videoüberwachung:

An- und Abmeldungen an:

<http://lists.fiff.de/mailman/listinfo/cctv-L>

Beiträge an: cctv-L@lists.fiff.de

FIF online

Das ganze FIF

www.fiff.de

Twitter FIF e.V. – [@Fiff_de](https://twitter.com/Fiff_de)

Cyberpeace

cyberpeace.fiff.de

Twitter Cyberpeace – [@Fiff_AK_RUIN](https://twitter.com/Fiff_AK_RUIN)

Faire Computer

blog.faire-computer.de

Twitter Faire Computer – [@FaireComputer](https://twitter.com/FaireComputer)

Mitglieder-Wiki

<https://wiki.fiff.de>

FIF-Beirat

Ute Bernhardt (Berlin); **Peter Bittner** (Bad Homburg); **Dagmar Boedicker** (München); Dr. **Phillip W. Brunst** (Köln); 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)

FIF-Vorstand

Stefan Hügel (Vorsitzender) – Frankfurt am Main
Rainer Rehak (stellv. Vorsitzender) – Berlin
Michael Ahlmann – Kiel / Blumenthal
Maximilian Hagner – Jena
Alexander Heim – Berlin
Sylvia Johnigk – München
Benjamin Kees – Berlin
Prof. Dr. **Hans-Jörg Kreowski** – Bremen
Kai Nothdurft – München
Prof. Dr. **Britta Schinzel** – Freiburg im Breisgau
Ingrid Schlagheck – Bremen
Anne Schnerrer – Berlin
Dr. **Friedrich Strauß** – München
Prof. Dr. **Werner Winzerling** – Fulda

FIF-Geschäftsstelle

Ingrid Schlagheck (Geschäftsführung) – Bremen

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	1 300 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, Stefan Hügel
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	Friedrich Dürrenmatt, Labyrinth II : Der verängstigte Minotaurus, 1974, Sammlung Centre Dürrenmatt Neuchâtel, Quelle: Bibliothèque nationale suisse, CC BY-SA 4.0
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

3/2022 Politik, Staat und Verwaltung

Jörg Pohle, Stefan Hügel

Redaktionsschluss: 5. August 2022

4/2022 100 Jahre Joseph Weizenbaum

Stefan Ullrich, Andrea Knaut

Redaktionsschluss: 4. November 2022

1/2023 FIfFKon 2022: FIfF Berlin

make install peace – Impulse für den Frieden

FIfF Berlin

Redaktionsschluss: 4. Februar 2023

Zuletzt erschienen:

1/2021 FIfFKon 2020 – Datenschutz, Usability, Barrierefreiheit und Sicherheit – Teil 1

2+3/2021 FIfFKon 2020 Teil 2

4/2021 Künstliche Intelligenz zieht in den Krieg

1/2022 Selbstbestimmung in digitalen Räumen

W&F – Wissenschaft & Frieden

3/21 Frieden lernen, aber wie? – Aktuelle Fragen der Friedenspädagogik

4/21 Chinas Welt? – Zwischen Konflikt und Kooperation

1/22 Täter:innen

2/22 Kriegerische Verhältnisse

vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik

#231/232 Zwei Jahre Datenschutz-Grundverordnung

#233 Kirchliches Sonderarbeitsrecht

#234 Strafvollzug in der Pandemie

#235 Zwei Jahre Corona – und wie weiter?

DANA – Datenschutz-Nachrichten

3/21 E-Government

4/21 Datenschutz-Visionen

1/22 Ampelpolitik

2/22 Social Media

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

Schluss F.I.f.F.

**Wie kommt es, dass Länder,
die wir als „stark“ bezeichnen,
so mächtig sind, wenn es darum geht,
Krieg zu führen,
aber so schwach, wenn es darum geht,
Frieden zu schaffen?**

**Wie kommt es, dass es so leicht ist,
Waffen zu beschaffen, aber so schwer,
Bücher zu liefern?**

**Warum ist es so leicht, Panzer zu bauen,
aber so schwer, Schulen zu errichten?**

Malala Yousafzai

Nutzungshinweise

Die vorliegende Datei wird im Rahmen der Mitgliedschaft des FfF e. V. oder eines Abonnements der FfF-Kommunikation zur Verfügung gestellt.

Die Einspeisung in Datenbanksysteme, Listen, Blogs oder die Bereitstellung der Datei zum Download durch Dritte wird ausdrücklich untersagt – die Datei dient ausschließlich dem privaten unbegrenzten Gebrauch durch die Mitglieder und die Abonnent:innen.