



Die Inhalte dieser Zeitschrift (Texte, Bilder etc.) stehen unter einer freien Creative-Commons-Lizenz – Namensnennung (CC BY 4.0), sofern nicht anders vermerkt. Es steht allen frei, die Inhalte in jedem Medium oder Format weiterzugeben, anzupassen und zu veröffentlichen unter den drei Bedingungen, dass 1) eine angemessene Creative-Commons-Urheberrechteangabe, 2) ein Link auf die Lizenz (<https://creativecommons.org/licenses/by/4.0/>) und 3) ggf. die Art der vorgenommenen Änderungen angeführt werden.

H 7625

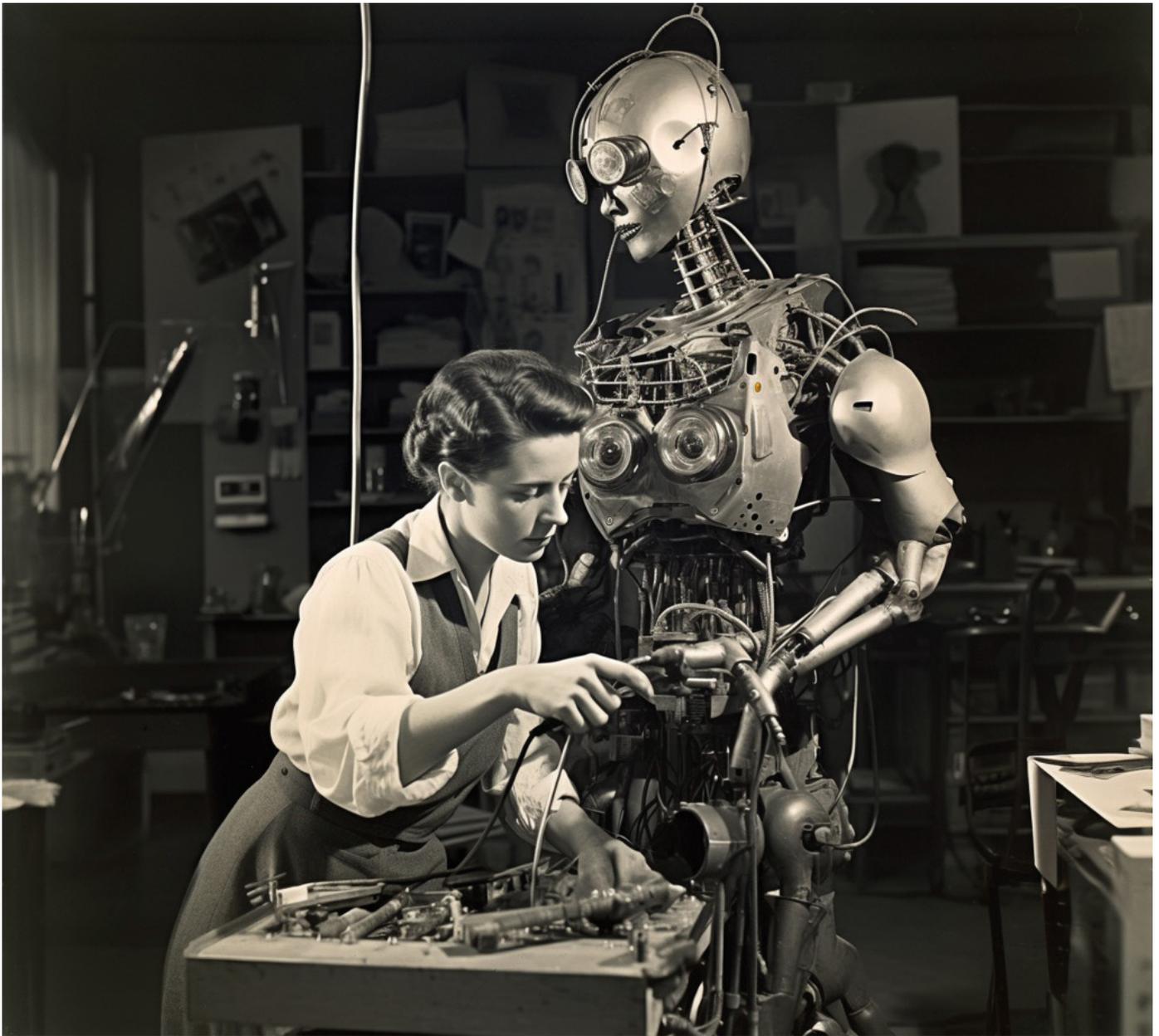
# E..I..f..F..Kommunikation

Zeitschrift für Informatik und Gesellschaft

40. Jahrgang 2023

Einzelpreis: 7 EUR

3/2023 – September 2023



## IT-Gestaltung für Gute Arbeit

ISSN 0938-3476

• Ist KI gefährlich? • Gedanken zur „Zeitenwende“ • netzpolitik.org •

## Inhalt

Ausgabe 3/2023

- 03 Editorial  
- Stefan Hügel

### Forum

- 04 Der Brief: Künstliche Intelligenz: künftige – nein, heutige – Informatik  
- Stefan Hügel
- 06 Cyberpeace – Für Frieden, Freiheit und eine lebenswerte Welt  
- Hans-Jörg Kreowski
- 06 „Dran! Drauf! Drüber!“  
Unzeitgemäße Gedanken zur „Zeitenwende“  
- Stefan Hügel
- 09 Ist die Künstliche Intelligenz gefährlich?  
- Karl Hans Bläsius, Jörg Siekmann

### FIfF e. V.

- 12 Einladung zur FIfF-Konferenz 2023 in Berlin
- 12 Einladung zur Mitgliederversammlung 2023

### Netzpolitik.org

- 55 Algorithmen rütteln kaum an politischen Einstellungen  
- Daniel Leisegang
- 57 Der Verteidigungsminister besucht den Cyberraum  
- Anna Biselli
- 59 Das sind 650.000 Kategorien, in die uns die Online-Werbeindustrie einsortiert  
- Ingo Dachwitz
- 67 Effizienz ist ein konservatives Wesen  
- Bianca Kastl

**Titelbild:** *A roboticist at work in her lab, 1950s (Roboter-Wissenschaftlerin am Arbeitsplatz)*  
*Das Bild ist ein Gemeinschaftswerk von Tullius Detritus und midjourney (einer Künstlichen Intelligenz). CC 0*

*Hito Steyerl zu Bildgeneratoren in der Süddeutschen Zeitung vom 14./15.8.2023, S. 12*

„Diese Bildgeneratoren brauchen fast keine Realität mehr. Das ist ein postfotografisches Medium, eine Art statistische Malerei. [...] Der Fotorealismus kommt, das ist keine Frage. Aber vom philosophischen Kontext her ist das etwas ganz anderes als ein Foto. Es geht nicht um Realität, es geht um Beschwörung, um Halluzination, um Projektion.“

### IT-Gestaltung für Gute Arbeit

- 13 Editorial zum Schwerpunkt  
- Dagmar Boedicker, Klaus Heß, Katharina Just
- 16 Der Mensch behält die Kontrolle – Intervenierbarkeit als Gestaltungsanforderung für die Nutzung von KI  
- Thomas Herrmann, Christoph Lentzsch
- 22 Manufacturing Execution Systems – Zwischen Überwachung und Guter Arbeit  
- Moritz Niehaus, Jürgen Klippert
- 26 Künstliche Intelligenz in der betrieblichen Praxis – Durch Mitbestimmung und Beteiligung die Transformation meistern.  
- Markus Hoppe
- 31 Gute Digitalisierung von Interaktionsarbeit?  
- Stephanie Porschen-Hueck
- 35 IT-Gestaltung für Gute Arbeit – Pro Partizipative Technik- und Arbeitsgestaltung  
- Christoph Schmitz
- 39 KI als Kollegin? Wie wünschen sich Beschäftigte Künstliche Intelligenz am Arbeitsplatz?  
- Sabine Pfeiffer
- 46 Was bleibt von der Arbeitswelt, wie wir sie kennen? Ein Praxisbericht aus Sicht der Beschäftigten und der Interessenvertretungen  
- Frank Steinwender
- 50 Mitbestimmung bei modernen ITK-Systemen und die Notwendigkeit das BetrVG zu ändern  
- Karin Schuler, Thilo Weichert

### Rubriken

- 71 Impressum/Aktuelle Ankündigungen
- 72 SchlussFIfF

## Editorial

Die Auswirkungen der Informatik auf die Arbeit ist ein wesentliches Thema des FfF, das bereits mehrfach in Schwerpunkten der FfF-Kommunikation thematisiert wurde, explizit als Schwerpunkt zuletzt in der Ausgabe 4/2016 *Zukunft der Arbeit – Arbeit der Zukunft: Wer steuert wen?*, aber auch in vielen weiteren Beiträgen. In dieser Ausgabe greifen wir das Thema erneut auf: *IT-Gestaltung für Gute Arbeit*. IT-Gestaltung heißt, sozio-technische Systeme zu gestalten, und, so die Redakteur:innen dieses Schwerpunkts, Dagmar Boedicker, Klaus Heß und Katharina Just:

*Um sozio-technische Systeme zu gestalten, sind dicke Bretter zu bohren, es dauert und die Machtverhältnisse bestimmen auch im Betrieb, was geht und was nicht. Und weil Technik Machtverhältnisse nun mal eher verstärkt als sie auszugleichen, gibt es immer noch viel weniger gute Beispiele als kritikwürdige.*

Ein wesentlicher Aspekt der aktuellen Entwicklungen ist die *Künstliche Intelligenz*, die in den letzten Jahren wieder eine beschleunigte Entwicklung nimmt, zuletzt in Form der *Generativen Künstlichen Intelligenz* und den *Large Language Models* (LLM), die in Systemen wie ChatGPT zum Einsatz kommen. Immer wieder kommt die Frage auf, wie sich diese Modelle und Systeme auf Arbeitsplätze auswirken – eine wohlbekannte Diskussion bereits seit den 1970er- und 1980er-Jahren. Doch mit der steigenden Leistungsfähigkeit der Anwendungen und Systeme sind auch zunehmend hochqualifizierte Berufe betroffen. Die Schwerpunktreakteur:innen schreiben dazu:

*Noch sind keine massenhaften Entlassungen zu erkennen, manche Unternehmen setzen KI aber als Drohmittel ein, wie AlgorithmWatch im Newsletter vom 26. Mai 2023 schreiben. Sie fassen zusammen: „Generative Modelle sind keine Arbeitsplatzkiller. Sie werden uns wahrscheinlich dabei helfen, manche Aufgaben zu bewältigen. Sie können uns aber auch noch mehr von unserer Arbeit entfremden und sich negativ auf die Qualität von Dienstleistungen auswirken.“*

Neben den Auswirkungen auf die Arbeit spielt auch das Militär eine wesentliche Rolle für die Entwicklung Künstlicher Intelligenz. Ein wichtiger Treiber für die Entwicklung von Technik ist leider immer wieder das Ziel, möglichst effektiv Menschen zu töten. Karl Hans Bläsius und Jörg Siekmann, unermüdliche Warner vor einem durch unkontrollierte Technik verursachten *Atomkrieg aus Versehen*, kommentieren das *Ein-Satz-Statement*, das vor den Risiken der Künstlichen Intelligenz für das Fortbestehen der Menschheit warnt:

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

Sie nennen dabei als Beispiele u. a. autonome Waffensysteme, unkalkulierbare Wechselwirkungen zwischen Künstlicher Intelligenz und Atomwaffen und allgemein die Veränderung der Kriegführung durch Künstliche Intelligenz.

Die aktuelle Eskalation angesichts des Konflikts in der Ukraine, die drohende Konfrontation mit China und die erneute Hochrüstung im Rahmen der von der Bundesregierung ausgerufenen *Zeitenwende* deuten auf einen bevorstehenden weiteren Rüstungswettlauf hin – mit allen damit verbundenen Risiken. In den 1980er-Jahren führte die Konfrontation zwischen den Staaten des Warschauer Vertrags und der NATO und der darauf folgende Rüstungswettlauf dazu, dass eine Friedensbewegung von einer breiten gesellschaftlichen Schicht getragen wurde – nicht zuletzt wurde in dieser Zeit die Partei *Die Grünen* gegründet, die der Friedensbewegung und der Opposition gegen das Militärbündnis NATO – bis hin zur Austrittsforderung – eine parlamentarische Stimme gab. Bekanntlich hat sich dies mittlerweile geändert; die heutige Partei *Bündnis90/Die Grünen* und ihre Anhänger:innen positionieren sich mehrheitlich als politische Kraft, die in der aktuellen Krise verstärkt für die Fortsetzung militärischer Maßnahmen wie Rüstungsexporte – auch in Kriegsgebiete – eintritt. Gleichzeitig erreicht ausgerechnet Bundesverteidigungsminister Pistorius den Spitzenplatz in der Beliebtheitskala unter Politiker:innen – mit großem Abstand.

Begleitend zum Schwerpunkt werden in Beiträgen aus *Netropolitik.org* weitere aktuelle Themen behandelt. Daniel Leisegang geht der Frage nach, inwiefern Empfehlungsalgorithmen sozialer Plattformen die politische Einstellung ihrer Leser:innen beeinflussen. Anna Biselli schreibt über das Kommando Cyber- und Informationsraum, Ingo Dachwitz untersucht die Eigenschaften, nach denen uns die Online-Werbeindustrie sortiert und beschreibt, wie präzise wir in rund 650.000 einzelne Kategorien eingeordnet werden. Bianca Kastl kommentiert die Verwaltungsmodernisierung und die Digitalisierung mit dem Ziel der Effizienz.

Unsere diesjährige *FfF-Konferenz* wirft ihre Schatten voraus. Vom 3.-5. November 2023 werden wir uns in der Jugendherberge am Ostkreuz in Berlin treffen, um uns mit wesentlichen Themen der Zukunft der Informatik auseinander zu setzen: Cyberpeace, Information und Nachhaltigkeit und aktuelle Entwicklungen der (generativen) Künstlichen Intelligenz. Wie in jedem Jahr werden wir den Weizenbaum-Studienpreis verleihen, bei der Mitgliederversammlung steht in diesem Jahr wieder die Neuwahl des Vorstands an.

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

Stefan Hügel  
für die Redaktion



## Künstliche Intelligenz: künftige – nein, heutige – Informatik<sup>1</sup>

*Der Computer Nr. 3 / sucht für mich den richtigen Boy  
und die Liebe ist garantiert für beide dabei  
Der Computer weiß genau / für jeden Mann die richtige Frau  
und das Glück fällt im Augenblick aus seiner Kartei*  
France Gall, 1968<sup>2</sup>



Liebe Freundinnen und Freunde des FIF, liebe Mitglieder,

*Künstliche Intelligenz* ist schon lange kein neues Thema mehr. Bekanntlich wurde der Begriff bereits in den 1950er-Jahren geprägt und das Thema – mit wechselnder Intensität – seither weiter verfolgt. Doch es ist anzunehmen, dass es sich bei der Entwicklung der Künstlichen Intelligenz um einen exponentiellen Prozess handelt, und so ist wohl mit einer sich immer stärker beschleunigenden Entwicklung zu rechnen.

Kritik an Künstlicher Intelligenz gab und gibt es seit Beginn ihrer Entwicklung. Zunächst wurde vor allem daran gezweifelt, ob die daran geknüpften Versprechungen und Ziele überhaupt – technisch – erreicht werden können. Joseph Weizenbaum argumentierte, dass die Erwartungen an Künstliche Intelligenz letztlich auf einem zu vereinfachten Begriff von Intelligenz beruhen.<sup>3</sup> Seither wurden die Grenzen des Möglichen immer weiter verschoben – neben Fortschritten in der Methodik trägt zweifellos die Entwicklung der Hardware-Technologie einen erheblichen Teil bei. Künstliche Intelligenz als methodische Weiterentwicklung der Informatik – im Sinne *schwacher* KI – entwickelt sich derzeit geradezu stürmisch weiter. Über die Entwicklung *starker* KI – also einer Form der Künstlichen Intelligenz, die dem Menschen ähnlich ist – ist damit noch nichts gesagt. Kann eine Maschine ein Bewusstsein entwickeln oder zumindest simulieren? Auch dies ist keine neue Frage mehr.<sup>4</sup> Wie kann ich erkennen, ob eine Maschine ein Bewusstsein hat? Kann ich überhaupt bei anderen Menschen erkennen, ob er oder sie ein Bewusstsein hat, oder ist das nur ein falscher Analogieschluss? Aber das würde uns in den Solipsismus führen.

Die gesellschaftspolitische Diskussion über Künstliche Intelligenz konzentriert sich zu Recht auf die Frage, wozu die damit hergestellten Systeme und Anwendungen methodisch in der Lage sind – ohne dass wir uns mit der Frage eines Bewusstseins auseinandersetzen müssen. Das ist augenscheinlich eine ganze Menge – mit steigender Tendenz. Selbst zur Erzeugung alltäglicher Gegenstände werden mittlerweile solche Verfahren genutzt; beispielsweise lässt mich der Hersteller meiner Brillengläser wissen, sie seien „Intelligent Glasses powered by Artificial Intelligence“ – ohne dass ich mir dabei über ein Bewusstsein Gedanken mache.

Doch es geht ja nicht nur um Brillengläser. Namhafte Wissenschaftler:innen warnen mittlerweile vor möglichen Konsequenzen der Weiterentwicklung und Anwendung Künstlicher Intelligenz. Stephen W. Hawking prognostizierte, dass bei einer weiteren Entwicklung der Computer entsprechend *Moore's Law* – Verdoppelung der Geschwindigkeit und Speicherkapazität ca. alle 18 Monate – die Intelligenz von Computern die des Men-

schen in den kommenden 100 Jahren übertreffen könnte, und schrieb:

*Wenn eine Künstliche Intelligenz (KI) besser wird als Menschen bei der Konstruktion von KI, sodass sie sich rekursiv ohne menschliche Hilfe selbst verbessern kann, dann steht uns höchstwahrscheinlich eine Intelligenzexplosion bevor, die letztlich in die Maschinenintelligenz mündet: Sie wird unsere Intelligenz in viel höherem Maß übertreffen als unsere menschliche Intelligenz die von Schnecken. Bevor es so weit ist, müssen wir sicherstellen, dass die Computer Ziele verfolgen, die auf einer Linie mit unseren Zielen liegen.<sup>5</sup>*

Und noch deutlicher:

*Es ist zu befürchten, dass die KI alleine weitermacht und sich mit ständig zunehmender Geschwindigkeit selbst überarbeitet. Menschen, die aufgrund der Langsamkeit ihrer biologischen Evolution beschränkt sind, könnten nicht mithalten und würden verdrängt.<sup>6</sup>*

Man mag das als alarmistisch abtun, genauso wie die sehr spekulativen Thesen von Nick Bostrom.<sup>7</sup> Doch es wäre wohl nicht klug, diese Möglichkeit zu ignorieren. Auch wenn er vermutlich nicht derartige Szenarien im Kopf hatte – auch hier gilt das Diktum von Hans Jonas:

*Handle so, dass die Wirkungen deiner Handlung verträglich sind mit der Permanenz echten menschlichen Lebens auf Erden.<sup>8</sup>*

Dieses Prinzip müsste freilich zur Anwendung kommen, bevor uns die weitere Entwicklung aus der Hand genommen wird, das heißt, bevor eine fortentwickelte Künstliche Intelligenz sich ihre eigene Ethik schafft, bei der sie das Ziel der Permanenz auf sich selbst „umbiegt“. Oder „erkennt“, dass es in Wahrheit der Mensch ist, der diesem Planeten den größten Schaden zufügt – und entsprechend handelt.

Doch wir sollten mit den praktischen Problemen beginnen: In der vorliegenden Ausgabe, Seite 9 ff., setzen sich Karl Hans Bläsius und Jörg Siekmann mit dem *1-Satz-Statement* zur Künstlichen Intelligenz auseinander, das von namhaften Expert:innen unterzeichnet wurde:

*Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.<sup>9</sup>*

Sie konzentrieren sich bei ihrer Bewertung vor allem auf die militärischen Risiken. Diese sind, angesichts der nuklearen Bedrohung, besonders gravierend, aber nicht die einzigen.

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

*AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research and acknowledged by top AI labs. As stated in the widely-endorsed Asilomar AI Principles, Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources. Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.*<sup>10</sup>

Die Initiative wurde allerdings gemischt aufgenommen<sup>11,12</sup>. *Time* kommentiert:

*The key issue is not “human-competitive” intelligence (as the open letter puts it); it’s what happens after AI gets to smarter-than-human intelligence. Key thresholds there may not be obvious, we definitely can’t calculate in advance what happens when, and it currently seems imaginable that a research lab would cross critical lines without noticing.*<sup>13</sup>

Die Debatte wird weitergehen, ob die Risiken der bereits eingesetzten Entwicklung richtig eingeschätzt werden oder ob die (negativen) Potenziale einer neuen Technologie hier alarmistisch übertrieben werden. Dass es Risiken gibt, steht wohl außer Frage: diskutiert werden die mangelnde Nachvollziehbarkeit der Systeme, erhebliche Risiken für den Datenschutz und die weitgehenden Möglichkeiten der Überwachung, beispielsweise durch automatisierte Gesichtserkennung, und vieles mehr. Auf Ebene der Europäischen Union soll Künstliche Intelligenz durch den *AI Act* reguliert werden, der sich vor allem auf die Einhegung von Hochrisikosystemen konzentriert.<sup>14</sup> Hier geht es noch nicht darum, dass Künstliche Intelligenz den Menschen überflügelt, sondern um die Einhegung konkreter Risiken beim Einsatz von KI-Systemen. Die Verordnung soll entsprechend das Ziel der Union unterstützen, „... bei der Entwicklung einer sicheren, vertrauenswürdigen und ethisch vertretbaren Künstlichen Intelligenz weltweit eine Führungsrolle einzunehmen“, und dabei „für den vom Europäischen Parlament ausdrücklich geforderten Schutz von Ethikgrundsätzen“ zu sorgen.<sup>15</sup> Weiter heißt es:

*Abgesehen von den zahlreichen nutzbringenden Verwendungsmöglichkeiten Künstlicher Intelligenz kann*

*diese Technik auch missbraucht werden und neue und wirkungsvolle Instrumente für manipulative, ausbeuterische und soziale Kontrollpraktiken bieten. Solche Praktiken sind besonders schädlich und sollten verboten werden, weil sie im Widerspruch zu den Werten der Union stehen, nämlich der Achtung der Menschenwürde, Freiheit, Gleichheit, Demokratie und Rechtsstaatlichkeit sowie der Grundrechte in der Union, einschließlich des Rechts auf Nichtdiskriminierung, Datenschutz und Privatsphäre sowie der Rechte des Kindes.*<sup>16</sup>

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

Mit Fliffigen Grüßen  
Stefan Hügel

## Anmerkungen

- 1 Die Bezeichnung „Künftige Informatik“ stammt von Wolfgang Wahlster: Wahlster W (2015) „Künstliche Intelligenz ist besser als natürliche Dummheit“. Interview, Jan-Bernd Meyer, CIO, <https://www.cio.de/a/kuenstliche-intelligenz-ist-besser-als-natuerliche-dummheit,3230970,2>
- 2 Gall F (1968) *Computer* Nr. 3, <https://genius.com/France-gall-der-computer-nr-3-lyrics>
- 3 Weizenbaum J (1987 [1976]) *Die Macht der Computer und die Ohnmacht der Vernunft*, 2. Auflage. Frankfurt am Main: Suhrkamp, S. 268.
- 4 Hofstadter DR, Dennett DC (1986 [1981]) *Einsicht ins Ich. Fantasien und Reflexionen über Selbst und Seele*. Dt. Übersetzung, Stuttgart: dtv/Klett-Cotta
- 5 Hawking SW (2018) *Kurze Antworten auf große Fragen*. Dt. Übersetzung, Stuttgart: Klett-Cotta, S. 208
- 6 Hawking SW (2018) a. a. O., S. 211
- 7 Bostrom N (2016) *Superintelligenz. Szenarien einer kommenden Revolution*. Dt. Übersetzung, Berlin: Suhrkamp
- 8 Jonas H (1979) *Das Prinzip Verantwortung. Versuch einer Ethik für die technologische Zivilisation*. Frankfurt am Main: Suhrkamp, S. 36
- 9 Safe.ai (2023) *Statement on AI Risk*, <https://www.safe.ai/statement-on-ai-risk>
- 10 Futureoflife.org (2023) *Pause Giant AI Experiments: An Open Letter*, <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>
- 11 Paul K and agencies (2023) *Letter signed by Elon Musk demanding AI research pause sparks controversy*. *The Guardian*, <https://www.theguardian.com/technology/2023/mar/31/ai-research-pause-elon-musk-chatgpt>
- 12 Yudkowsky E (2023) *Pausing AI Developments isn't enough. We need to shut it all down*. *Time*, <https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/>
- 13 Yudkowsky E (2023) a. a. O.
- 14 Europäische Union (2021) *The Artificial Intelligence Act*, <https://artificialintelligenceact.eu/the-act/>
- 15 Europäische Union (2021) a. a. O., *Erwägungsgrund 5*
- 16 Europäische Union (2021) a. a. O., *Erwägungsgrund 15*



## Cyberpeace – für Frieden, Freiheit und eine lebenswerte Welt



Wie auf dem Bremer Zukunftsworkshop *Cyberpeace* im Mai besprochen, wird im Rahmen der nächsten FIFF-Konferenz, die vom 3. bis 5. November in Berlin stattfindet, ein Cyberpeace-Block stattfinden mit einem *Pecha Kucha* rund um Cyberpeace und mit einer kollektiven Lecture Performance von ground-zero@KHM: *Aesthetic approaches to cyberpeace work* von einer Studiengruppe der Kunsthochschule für Medien Köln. Außerdem wird die gesamte Konferenz von der Kunstausstellung *cyberpeace works* begleitet.

Das Pecha Kucha wird sich aus einer Serie von voraussichtlich sechs Kurzvorträgen mit 20 Folien à 20 Sekunden zusammensetzen, von denen vier bereits feststehen:

- Daniel Guagnin: IT-Sicherheit vs. Sicherheit in der IT: ein unlösbarer Widerspruch?
- Christian Heck: Das „gläserne Schlachtfeld“
- Sylvia Johnigk: Schwachstellen – warum wir deren Veröffentlichung wollen
- Kai Nothdurft: Kampagnen-Aufruf zur Veröffentlichung von Schwachstellen

Siehe zur FIFF-Konferenz 2023 auch die Einladung in diesem Heft und für weitere Informationen die Konferenzwebseite [2023.fiffkon.de](https://2023.fiffkon.de).

Auf dem Symposium *Wissenschaft für den Frieden* zum 40-jährigen Bestehen der Zeitschrift *Wissenschaft und Frieden* am 6. und 7. Oktober in Bonn, das vom FIFF organisatorisch und finanziell unterstützt wird, gibt es einen großen Vortragsblock mit sieben Beiträgen zu den naturwissenschaftlichen und tech-

nischen Aspekten von Krieg und Frieden. Fünf Vorträge haben einen engen Bezug zum Komplex Cyberpeace:

1. Karl Hans Bläsius und Jörg Siekmann: *Computergestützte Frühwarn- und Entscheidungssysteme für nukleare Bedrohungen*
2. Jens Hälterlein: *ELSA zieht in den Krieg – Zur Rolle der Kritik an AWS für deren Legitimationsstrategien*
3. Christian Heck: *Kognitive Assemblagen im gläsernen Gefechtsfeld*
4. Thea Riebe, Anja-Liisa Gonsior, Stefka Schmid, Thomas Reinhold und Christian Reuter: *Friedensinformatik: heute und morgen*
5. Timothy Williams: *Frieden und Konflikt in der digitalen Ära*

Für detaillierte Informationen dazu und zum weiteren spannenden Programm und der Organisation des Symposiums sei auf <https://pretalx.com/wissenschaft-fuer-den-frieden-2023/> und <https://wissenschaft-und-frieden.de/projekt/40-jahre-jubilaum/> verwiesen.

Karl Hans Bläsius, Jörg Siekmann stellen in diesem Heft die Frage: *Ist die Künstliche Intelligenz gefährlich?*, was weit verstanden in den thematischen Kontext von Cyberpeace fällt.

Wer die Arbeit der Cyberpeace-Kampagne unterstützen oder zumindest über deren Aktivitäten auf dem Laufenden gehalten werden möchte, kann gern in die Mailingliste aufgenommen werden. Schickt bitte eine entsprechende E-Mail an [margita.zallmann@t-online.de](mailto:margita.zallmann@t-online.de) oder [kreo@fiff.de](mailto:kreo@fiff.de).



Stefan Hügel

### „Dran! Drauf! Drüber!“<sup>1</sup>

#### Unzeitgemäße Gedanken zur „Zeitenwende“

Jede Bundesregierung wird mit bestimmten Entscheidungen und gesellschaftlichen Entwicklungen in Verbindung gebracht. In die Zeit Konrad Adenauers fällt die Nachkriegszeit, das *Wirtschaftswunder* (das wir wohl vor allem der nicht ganz uneigennütigen Unterstützung durch die Sieger des 2. Weltkriegs zu verdanken haben), aus militärischer Sicht die Wiederbewaffnung und Gründung der Bundeswehr. In die Zeit Willi Brandts fällt die (heute zuweilen wieder geschmähte) Ostpolitik, in die Zeit Helmut Schmidts der *Deutsche Herbst* und der *NATO-Doppelbeschluss*, den Namen Helmut Kohls verbinden wir mit der europäischen Einigung und der Vereinigung der Bundesrepublik mit der DDR.

Wir wissen noch nicht, womit wir die Regierung von Bundeskanzler Olaf Scholz, in den Medien gerne liebevoll-infantilisierend als *Ampel* bezeichnet, künftig verbinden werden – sicherlich werden der Klimawandel, die Covid-19-Pandemie und der Krieg in der Ukraine eine Rolle spielen. Kürzlich kam ich in Frankfurt am Main an einer Skulptur vorbei: Ein militärischer Stahlhelm, umgeben von Trümmern und Patronenhülsen, beschriftet mit *Bazooka*, *Wumms* und *100 Mrd. €*. Hoffentlich wird nicht diese Symbolik die Erinnerung an die Regierung bestimmen.

Auch *Wenden* haben wir einige erlebt: Die *geistig-moralische Wende* durch das konstruktive Misstrauensvotum gegen Helmut Schmidt 1982 (die dann mit der Diffamierung des Generals

Kießling wegen (angeblicher) Homosexualität und der Flick-Affäre begann – doch ich schweife ab). 1989 folgte die politische Wende in der DDR und die deutsche Vereinigung. Der 24. Februar 2022 – der Tag des russischen Überfalls auf die Ukraine – markiert nun also die *Zeitenwende*:

*Der 24. Februar 2022 markiert eine Zeitenwende in der Geschichte unseres Kontinents. Mit dem Überfall auf die Ukraine hat der russische Präsident Putin kaltblütig einen Angriffskrieg vom Zaun gebrochen ...<sup>2</sup>*

Und weiter:

*Wir werden dafür ein Sondervermögen Bundeswehr einrichten ... Der Bundeshaushalt 2022 wird dieses Sondervermögen einmalig mit 100 Milliarden Euro ausstatten ... Die Mittel werden wir für notwendige Investitionen und Rüstungsvorhaben nutzen ... Wir werden von nun an Jahr für Jahr mehr als 2 Prozent des Bruttoinlandsprodukts in unsere Verteidigung investieren.*

Das Protokoll vermerkt: „Anhaltender Beifall bei der SPD und der FDP sowie bei Abgeordneten der CDU/CSU – Beifall bei Abgeordneten des BÜNDNISSES 90/DIE GRÜNEN und der AfD – Abgeordnete der SPD, der CDU/CSU und der FDP erheben sich.“

Für die Bereitstellung der 100 Mrd. € wurde eigens die Verfassung geändert.<sup>3</sup> Der Herausgeber des Freitag, Jakob Augstein, kommentierte:

*Es war gespenstisch. Die deutschen Parlamentarier feierten die größte Aufrüstung, die es in der deutschen Geschichte seit dem Zweiten Weltkrieg je gegeben hat. Schweigen und stiller Ernst wären angemessen gewesen.<sup>4</sup>*

Die NATO fordert, jährlich 2 % des BIP für Rüstung auszugeben. Dieses Ziel wird von Deutschland bisher nicht erfüllt. Noch 2017 lehnte der damalige Bundesaußenminister Sigmar Gabriel dies ab und erklärte, er wisse gar nicht, „wo wir die ganzen Flugzeugträger hinstellen sollen“<sup>5</sup>. Im Koalitionsvertrag 2021-2025 heißt es dazu etwas schwurbelig: „Die NATO-Fähigkeitsziele wollen wir in enger Abstimmung mit unseren Partnern erfüllen

und entsprechend investieren.“<sup>6</sup> Unter dem Eindruck der Aggression gegen die Ukraine kündigte Bundeskanzler Scholz nun die Umsetzung dieses Ziels an. Der Spiegel kritisiert, dass diese Zusage bereits wieder aufgeweicht werde.<sup>7</sup> Doch offenbar wird die Akzeptanz einer zunächst unpopulären Maßnahme scheinweise erhöht. Erkennbar ist in der Abbildung das Verhältnis des Militäraufwands in NATO-Mitgliedsstaaten gegenüber Staaten, die nicht Mitglied der NATO sind. Klar wird allerdings auch, dass die USA mit Abstand für den größten Anteil der Ausgaben weltweit verantwortlich sind.

Nach dem Schock der beiden Weltkriege – für die Deutschland die wesentliche Verantwortung trägt – galt Deutschland als pazifistisch und kriegsunwillig. Die baldige – von der Bundesregierung unter Bundeskanzler Adenauer vorangetriebene – Wiederbewaffnung und Gründung der Bundeswehr war umstritten. Einen Höhepunkt erreichte die deutsche Friedensbewegung in den frühen 1980er-Jahren, als sie Hunderttausende gegen die Nachrüstung des NATO-Doppelbeschlusses mobilisieren konnte. Doch an einer generell pazifistischen Haltung in Deutschland bestehen auch Zweifel.<sup>9</sup> „War das pazifistische Selbstbild der Deutschen nur eine nostalgische Kulisse?“<sup>10</sup>

Indizien für eine Diskursverschiebung<sup>11</sup> hin zum Militärischen mögen auch die folgenden Schlaglichter bieten:

- Der Vorschlag, *Wehrkundeunterricht* in Schulen einzuführen.<sup>12</sup> „Wehrkunde ist Teil der Staatsbürgerkunde, der Gemeinschaftskunde, der politischen Bildung – wie man auch immer im Dickicht des deutschen Bildungsföderalismus die Stunden nennt, in denen zur Demokratie erzogen werden soll“, so der Autor des Meinungsbeitrags. Wehrkundeunterricht war bereits in der DDR ab 1978 Pflichtfach.
- Die Wehrbeauftragte schlägt die Einführung einer allgemeinen *Musterung* vor, um junge Deutsche zum Dienst zu motivieren.<sup>13</sup> Die Wiedereinsetzung der Wehrpflicht steht aber offenbar momentan (noch?) nicht zur Debatte.<sup>14</sup>
- Der derzeit diskutierte Entwurf des *Selbstbestimmungsgesetzes* – in dem die Rechte von trans Menschen gestärkt werden sollen – enthält ein explizites Verbot, im Verteidigungsfall dem Einsatz an der Waffe durch eine *Änderung des Geschlechts* zu entgehen<sup>15</sup>. Man mag dieses Verbot als

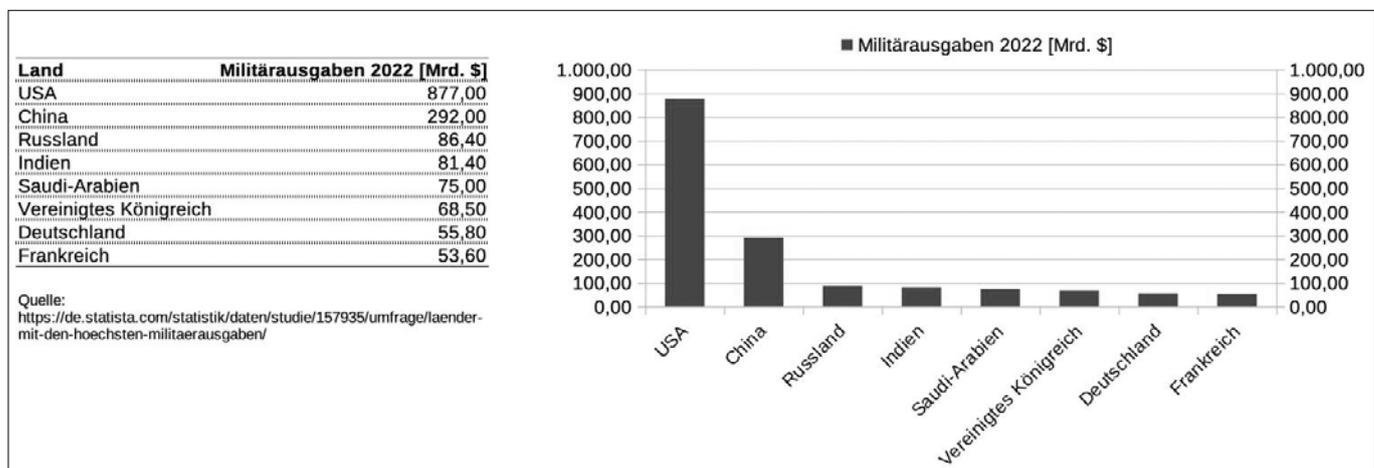


Abbildung: Länder mit den höchsten Ausgaben für das Militär weltweit (Quelle: Statista<sup>8</sup>)

Verhinderung missbräuchlicher Anwendung des Rechts befürworten – dennoch zeigt es, wie tief militärisches Denken bereits in intimste Bereiche der Persönlichkeit eindringt.

- Männer, die in den 1980er-Jahren den *Militärdienst* aus Gewissensgründen verweigert haben, revidieren diese Entscheidung nun öffentlich.<sup>16</sup> Dies mag wohlfeil sein, haben diese Männer doch die Altersgrenze für eine militärische Verwendung in der Regel längst überschritten. Doch auch dies zeigt die Verschiebung des Diskurses hin zum Militärischen.

Nach der politischen Wende in den Staaten des Warschauer Vertrags erlebten wir eine fast dreißigjährige Zeit des Friedens und der wirtschaftlichen Prosperität in Mitteleuropa, freilich erkaufte durch Abschottung an den Grenzen und den Export von Konflikten – die auch von der wirtschaftlichen Kooperation mit Russland profitierte. Doch war das alles nur ein Irrtum? Wäre eine Phase zunehmender Militarisierung stattdessen erstrebenswert gewesen? Wir haben uns in eine problematische Abhängigkeit von Russland begeben, indem wir auf das (billige) russische Gas gesetzt haben. Doch kann die langfristige Umstellung auf LNG-Lieferungen – gewonnen aus dem besonders umweltschädlichen Fracking, transportiert auf umweltschädlichen Tank Schiffen, angeliefert an LNG-Terminals in unmittelbarer Nachbarschaft unserer Urlaubsstrände – eine tragfähige Alternative sein? Warum funktioniert das von Bundeskanzler Scholz bejubelte *Deutschland-Tempo* immer dann so gut, wenn es zu Zerstörung der Umwelt führt?

Deutschland trägt in seiner Geschichte die Hauptverantwortung für zwei mörderische Weltkriege; die deutsche Armee hat unvorstellbare Verbrechen begangen.<sup>17</sup> Die Erinnerung daran darf nicht verblassen, es kann keinen Schlussstrich geben. Es war ein anderer Staat, eine andere Zeit – aber es waren unsere Vorfahren. Wir dürfen es nie vergessen. Und wir müssen es bei allen außenpolitischen und militärischen Entscheidungen mitbedenken.

## Anmerkungen

- 1 *Schlachtruf der Panzergrenadiertruppe der Bundeswehr* (siehe Wikipedia, Stichwort Panzergrenadiertruppe (Bundeswehr), [https://de.wikipedia.org/wiki/Panzergrenadiertruppe\\_\(Bundeswehr\)](https://de.wikipedia.org/wiki/Panzergrenadiertruppe_(Bundeswehr)))
- 2 *Deutscher Bundestag* (2022) Plenarprotokoll 20/19
- 3 *Artikel 87a (1a) GG*, [https://www.gesetze-im-internet.de/gg/art\\_87a.html](https://www.gesetze-im-internet.de/gg/art_87a.html). Dass Haushaltspostitionen direkt in die Verfassung geschrieben werden, dürfte auch nicht allzu häufig vorkommen.
- 4 *Jakob Augstein* (2022) *Gigantisches Rüstungspaket von Olaf Scholz ist gefährlicher Irrweg*, der Freitag 09/2022, <https://www.freitag.de/autoren/jaugstein/gigantisches-ruestungspaket-ist-gefaehrlicher-irrweg>

- 5 *Markus Becker* (2017) *Gabriel rebelliert gegen Tillerson*, Spiegel online, <https://www.spiegel.de/politik/ausland/nato-treffen-sigmar-gabriel-verweigert-rex-tillerson-die-gefolgschaft-a-1141408.html>
- 6 *Mehr Fortschritt wagen. Koalitionsvertrag 2021-2025 zwischen der Sozialdemokratischen Partei Deutschlands (SPD), Bündnis 90 / Die Grünen und den Freien Demokraten (FDP)*, [https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag/Koalitionsvertrag\\_2021-2025.pdf](https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag/Koalitionsvertrag_2021-2025.pdf)
- 7 *Marina Korbaki* (2023) *Zwei Prozent pro Jahr? Nee, doch nicht*, Spiegel online, <https://www.spiegel.de/politik/deutschland/zwei-prozent-pro-jahr-nee-doch-nicht-a-4b4b0657-a953-4ff4-b24a-591dc7cc366b>
- 8 *Länder mit den höchsten Ausgaben für das Militär weltweit 2022*, Statista Research Department, <https://de.statista.com/statistik/daten/studie/157935/umfrage/laender-mit-den-hoechsten-militaerausgaben/#statisticContainer>
- 9 *Frank A. Stengel* (2023) *Bundeswehr und deutsche Gesellschaft: Die Berliner Republik zwischen Militarisierung und Normalisierung*, in: *Malte Riemann, Georg Löffelman Hg. (2023) Deutschlands Verteidigungspolitik. Nationale Sicherheit nach der Zeitenwende*. Stuttgart: Kohlhammer
- 10 *Ulrike Winkelmann* (2022) *Die deutsche Friedensliebe*, taz online, <https://taz.de/Zukunft-des-Pazifismus/15901197/>
- 11 *Barbara Supp* (2023) *Mit Panzern spielen*, Spiegel online, <https://www.spiegel.de/kultur/ukraine-krieg-und-deutscher-militarismus-mit-panzern-spielen-a-6f41b564-fde2-419e-812f-f0796a0cdfcc>
- 12 *Alan Posener* (2022) *Für Wehrkunde im Schulunterricht*, Zeit online, <https://www.zeit.de/gesellschaft/schule/2022-06/wehrkunde-schule-unterricht-armee/komplettansicht>
- 13 *Wehrbeauftragte Högl bringt Rückkehr der Musterung ins Spiel*, Spiegel online, <https://www.spiegel.de/politik/deutschland/bundeswehr-eva-hoegl-bringt-rueckkehr-der-musterung-ins-spiel-a-10b712ef-2ece-48b1-8775-c7b50bdce875>
- 14 *„Eine Wehrpflicht würde uns in den nächsten zwei, drei Jahren nicht helfen“*, Spiegel online, <https://www.spiegel.de/politik/deutschland/boris-pistorius-eine-wehrpflicht-wuerde-uns-in-den-naechsten-zwei-drei-jahren-ueberhaupt-nicht-helfen-a-fce6b60a-a03e-45b2-8291-e5dbad2416d6>
- 15 *René Pfister: Kann man sich dafür entscheiden, eine Frau zu sein?* Spiegel online, <https://www.spiegel.de/politik/selbstbestimmungsgesetz-kann-man-sich-dafuer-entscheiden-eine-frau-zu-sein-kolumne-a-e3c2407e-67d5-4178-87e2-58239b97e55e>
- 16 *Tobias Rapp* (2023) *Kriegsdienst: Warum ich heute nicht mehr verweigern würde*, Spiegel online, <https://www.spiegel.de/kultur/pazifismus-warum-ich-meine-kriegsdienstverweigerung-zurueckziehe-a-caab1bb7-3a79-4542-8c58-f20658c56567>
- 17 *Hamburger Institut für Sozialforschung Hg. (2002) Verbrechen der Wehrmacht. Dimensionen des Vernichtungskrieges 1941-1944*. Ausstellungskatalog, Hamburger Edition, Hamburg: Verlag des Hamburger Instituts für Sozialforschung



**Stefan Hügel**

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

## Ist die Künstliche Intelligenz gefährlich?

Die Künstliche Intelligenz (KI) ist eine wissenschaftliche Disziplin, die eine Technologie ermöglicht, die das Leben auf dieser Erde noch einmal grundsätzlich verändern wird. Obwohl die meisten KI-Anwendungen aus unserer Sicht positiv zu bewerten sind und zu einer Verbesserung der menschlichen Lebensqualität geführt haben und weiterhin führen werden, gibt es kritische Anwendungen, die man kennen sollte, um diese Risiken möglichst gering zu halten. Am 30. Mai 2023 wurde ein sogenanntes Ein-Satz-Statement veröffentlicht, in dem vor dem Aussterben der Menschheit durch die KI gewarnt wird. Nachfolgend wird auf dieses Statement eingegangen und auf konkrete Risiken mit möglicherweise gravierenden Folgen hingewiesen.

### Das Ein-Satz-Statement

Das Statement lautet (safe.ai 2023a):

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

(Frei übersetzt: „Das Risiko, dass die KI das Aussterben der Menschheit bewirken könnte, sollte neben anderen Risiken von gesellschaftlichem Ausmaß wie Pandemien und Atomkrieg eine globale Priorität sein.“)

Unterzeichner sind u. a. die Unternehmenschefs von Google DeepMind und OpenAI, viele weitere Verantwortliche von großen IT- bzw. KI-Unternehmen sowie sehr renommierte KI-Wissenschaftler wie Stuart Russell und Peter Norvig, die Autoren des seit vielen Jahren weltweit wichtigsten KI-Lehrbuches (Russell/Norvig 2012). Die Unterzeichner, also auch die Chefs großer KI-Unternehmen, fordern eindringlich Regulierungen für Anwendungen der KI.

In den Medien wurde dieser Aufruf wenig beachtet, teilweise auch kritisiert. Kritisiert wurde vor allem, dass die Unterzeichner Aufmerksamkeit nur auf sich und ihre Produkte lenken wollen und dass sie selbst Einfluss auf mögliche Regulierungen nehmen möchten. Kritische Kommentare gingen dagegen kaum auf die eigentlichen Risiken ein und wenn, dann nur auf das Risiko einer möglichen Superintelligenz, wobei dies meist eh als technisch unmöglich abgetan wurde.

Andererseits sind die Gefahren der KI durchaus erkannt und es gibt weltweit, insbesondere auch in der EU, Initiativen zu Regulationsmaßnahmen. Die großen KI-Forschungsinstitute haben mittlerweile eigene Ethikabteilungen eingerichtet, unter anderem auch das Deutsche Forschungsinstitut für KI (DFKI).

Die Unterzeichner des Ein-Satz-Statements sind international herausragende KI-Experten, und solche Aufrufe müssen ernst genommen werden, so wie auch die Aufrufe von Klimaforschern vor einigen Jahrzehnten hätten ernst genommen werden müssen. Im Vergleich zum Klimawandel könnten manche KI-Risiken sogar erheblich gravierender sein und vor allem relativ plötzlich eintreten.

### Welche Risiken bezüglich KI könnten auf uns zu kommen?

Der *eine Satz* des Statements besagt natürlich nicht, um welche Art von Risiken es sich handelt, und das möchten wir im Folgenden konkretisieren:

1. Autonome Waffensysteme
2. Unkalkulierbare Wechselwirkungen zwischen KI und Atomwaffen
3. Revolution der Kriegsführung durch KI
4. Mit Hilfe von KI entwickelte Bio- und Chemiewaffen
5. Informationsdominanz und Manipulation
6. Superintelligenz

### Autonome Waffensysteme

Der jetzige Konfrontationskurs zwischen dem Westen und Russland und der drohende Konfrontationskurs zwischen den USA und China werden einen weiteren Rüstungswettkampf befeuern, vor allem in Schlüsseltechnologien wie der KI und im Cyberraum, denn keine Nation möchte riskieren, hier hinterher zu hinken. Es ist zu erwarten, dass es schon bald für viele Waffenarten automatische oder autonome Systeme geben wird, wie unter anderem Roboter, Fahrzeuge, Flugobjekte, Schiffe und U-Boote, wo Menschen durch KI-Komponenten ersetzt werden. Es werden auch neuartige Waffensysteme hinzukommen, wie z. B. Minidrohnen, die mit automatischer Bilderkennung und Gesichtserkennung automatisch einen Weg zu einem Ziel suchen und dieses dann angreifen. Bei diesen Entwicklungen geht es um Software und dabei sind Rüstungskontrolle und Abrüstung kaum möglich, denn Software kann einfach verschlüsselt über das Internet verbreitet werden. Chancen und Risiken von Autonomen Waffensystemen sind z. B. beschrieben in Grünwald/Kehl 2020 und Lahl 2021.

Man mag dies zwar als gefährlich, aber letztendlich als die „normale“ Weiterentwicklung von Kriegsgerät ansehen, wie es immer mit dem Aufkommen neuer Technologien verbunden war. Das ist bei den folgenden Punkten jedoch nicht so.

### Unkalkulierbare Wechselwirkungen zwischen KI und Atomwaffen

Die Weiterentwicklung von Waffensystemen mit höherer Treffsicherheit und immer kürzeren Flugzeiten (Hyperschallraketen) wird Techniken der KI erforderlich machen, um in Frühwarnsystemen für nukleare Bedrohungen Entscheidungen für gewisse Teilaufgaben auf Grund der kurzen zur Verfügung stehenden Zeitspanne von wenigen Minuten oder Sekunden automatisch zu treffen. Es gibt bereits Forderungen, autonome KI-Systeme zu entwickeln, die vollautomatisch eine Alarmmeldung bewerten und gegebenenfalls einen nuklearen Gegenschlag auslösen, da für menschliche Entscheidungen keine Zeit mehr bleibt. Zwischen KI-Entwicklungen und Atomwaffen kann es weitere un-

kalkulierbare Wechselwirkungen geben – solche Aspekte sind in Timm/Siekmann/Bläsius 2020 beschrieben.

### Revolution der Kriegsführung durch KI

In Militärkreisen wird KI nach Schießpulver und Atomwaffen als weitere Revolution der Kriegsführung angesehen, denn auf allen Ebenen der Kriegsführung, wie Informationsgewinn, Einsatzplanung und vernetzte Gefechtsdurchführung können bisherige kognitive und reaktive Grenzen eines Menschen durch KI überwunden werden (siehe z. B. Lahl/Varwick 2022: 130-136).

### Mit Hilfe von KI entwickelte Bio- und Chemiewaffen

In den letzten Jahren sind einige spektakuläre Erfolge der KI im Bereich der Biotechnologie bekannt geworden: Zum Beispiel wurden Experimente durchgeführt, um zu prüfen, ob mit Hilfe von KI sehr wirksame Chemie- und Biowaffen hergestellt werden können, die globale Epidemien und ähnlich verheerende Auswirkungen haben. Verschiedene Veröffentlichungen weisen auf diese Risiken hin (z. B. Urbina/Lentzos/Invernizzi/Ekins 2023).

### Informationsdominanz und Manipulation

Mit Hilfe von KI-Systemen oder durch diese könnte eine Informationsdominanz erreicht werden, wobei es nicht mehr um Wahrheit, sondern um Einflussnahme, Manipulation und Macht geht. Hierbei könnten auch fake news, Chatbots und andere technische Möglichkeiten eine wesentliche Rolle spielen (safe. ai 2023b). Dies könnte unsere Freiheit erheblich gefährden und Gesellschaftssysteme instabil machen.

### Superintelligenz

Aufgrund der Leistungsfähigkeit heutiger KI-Systeme gibt es neuere Warnungen zu den Gefahren einer „Superintelligenz“: Wissenschaftler, die bisher davon ausgingen, dass erst zum Ende dieses Jahrhunderts eine Situation erreicht werden könnte, in der künstliche Systeme in allen Bereichen Menschen deutlich überlegen sind, äußern jetzt die Befürchtung, dass dies vielleicht schon in den nächsten Jahrzehnten zu erwarten sei (Hendrycks 2023). Die Folgen für die Menschheit sind zwar völlig unkalkulierbar, allerdings erscheint uns dies im Vergleich zu den obigen Punkten zum heutigen Zeitpunkt zwar diskussionswürdig, aber doch zu spekulativ. Detaillierter wird das Thema Superintelligenz behandelt in Bläsius 2023; hierbei wird auch Bezug genommen auf wichtige Bücher von Forschenden in diesem Bereich: Bostrom 2014, Russell 2020, Shanahan 2021 und Tegmark 2017. In diesen Büchern wird behandelt, wie eine Superintelligenz entstehen und welche Folgen dies haben könnte.

### Vergleich mit dem Klimawandel

Warnungen vor dem Klimawandel gibt es seit vielen Jahrzehnten und diese Warnungen sind insgesamt bei weitem zu wenig beachtet worden. Aber es gilt:

- Auswirkungen des Klimawandels können recht gut vorhergesagt werden, wie z. B. der Anstieg des Meeresspiegels, die Anzahl der Hitzetage in bestimmten Regionen, die Zunahme von Stürmen, Unwettern und anderem. Allerdings kann es natürlich Kippunkte geben, die nicht genau vorhergesagt werden können.
- Der Klimawandel vollzieht sich allmählich, bei Zeiträumen für Maßnahmen geht es teilweise um Jahrzehnte.

Bei den Warnungen vor der KI ist die Situation anders. Am ehesten absehbar, begrenzt und kalkulierbar sind noch die Risiken von autonomen Waffensystemen, wobei die Zeiträume im Vergleich zum Klimawandel relativ klein sind. Für immer mehr Waffenarten sind schon in den nächsten Jahren autonome Varianten zu erwarten, die im Zusammenhang mit Atomwaffen und Biowaffen gravierend sind:

- Es sind kaum Vorhersagen möglich, was dann passieren kann und wie die möglichen Auswirkungen sein werden.
- Entsprechende Ereignisse werden eher plötzlich eintreten. Gravierende Folgen könnten dann innerhalb von wenigen Wochen oder Monaten eintreten, ohne Möglichkeit, diese noch aufzuhalten.
- In den oben genannten Fällen (Atomwaffen, Biowaffen, Superintelligenz) könnten die Folgen die gesamte Menschheit oder zumindest einen großen Teil davon betreffen.
- Diese gravierenden Folgen können unumkehrbar bereits in den nächsten Jahren oder Jahrzehnten auftreten.
- Die Möglichkeit entsprechende Ereignisse abzuwarten und erst dann zu handeln, um die Risiken zu reduzieren, wird es eventuell nicht mehr geben.

Ähnlich wie beim Klimawandel gilt auch hier: solange es keine wirksamen Maßnahmen gibt, um den Klimawandel einzudämmen, werden die Risiken weiterhin massiv steigen. Dies betrifft das Wettrüsten um autonome Waffensysteme und viele andere Aspekte der Kriegsführung durch KI.

### Wie können die Risiken reduziert werden?

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

Bei diesem Bedrohungsszenario geht es nicht nur um gezielte technische Verbesserungen und Risikominimierung, sondern um ein grundsätzliches Umdenken, wie von Leo Ensel gefordert (Ensel 2023):

*„Die Gefahr einer militärischen Totalkatastrophe – sei es durch Massenvernichtungsmittel wie Atombomben oder biologische Waffen, sei es durch eine mittels Künstlicher Intelligenz oder Superintelligenz entfesselte und sich verselbständigende Technik oder sei es durch ein Zusammenspiel all dieser Faktoren – ist durch eine (noch so ausgeklügelte) Technik definitiv nicht zu bannen, da es längst die Technik selbst ist, die in immer rasanterem Tempo zum Problem, nein: zur Gefahr wird.“*

Hier dargestellte Risiken können vermutlich nur reduziert werden, wenn es ein gutes Verhältnis zwischen allen Nationen gibt. Natürlich können die oben genannten Entwicklungen im Bereich KI dadurch nicht ganz verhindert werden. Aber bei guten internationalen Beziehungen werden sich Prioritäten für KI-Entwicklungen verschieben und damit werden die Risiken reduziert. Eine Verbesserung internationaler Beziehungen kann auf allen Ebenen erfolgen, z. B. kulturell, sportlich, wissenschaftlich und wirtschaftlich. Auch sportlicher und kultureller Austausch verbessern das Vertrauen zwischen Völkern. Sportler und Musiker aus allen Nationen sollten wieder überall zugelassen werden. Städtepartnerschaften sollten erhalten oder ausgebaut werden. Wissenschaftliche, technologische und wirtschaftliche Zusammenarbeit sollten nicht reduziert, sondern wieder hergestellt und verbessert werden.

Wenn es gelingt, Vertrauen und Zusammenarbeit zwischen den Atommächten wieder aufzubauen, dann verbessern sich auch die Aussichten auf Vereinbarungen zur Begrenzung verschiedener Waffenarten, wie z. B. Atom-, Cyber-, Bio- und autonome Waffen. Der Realisierung eines Atomwaffenverbots könnten wir dann einen Schritt näher kommen. Weltweit geltende Vereinbarungen bezüglich dieser Risiken wären die wichtigste Maßnahme und die UN wären die richtige Organisation, um Transparenz und Regulierung herzustellen. Als Voraussetzung dafür müssen Vertrauen, Kommunikation und Zusammenarbeit zwischen allen Nationen neu aufgebaut und deutlich verbessert werden.

Leo Ensel (2023):

*„Die tiefste Wurzel der gesamten Malaise liegt nicht in einer – niemals fehlerfreien – Technik oder Künstlichen Intelligenz, sondern in dem abgrundtiefen Misstrauen, das alle rivalisierenden geopolitischen Akteure gegeneinander hegen!“*

Als erster Schritt müssen daher die bestehenden vertrauensbildenden Maßnahmen der jüngeren Vergangenheit erweitert beziehungsweise erst wieder hergestellt werden, um dann zu einer globalen Sicherheitskultur erweitert zu werden.

## Referenzen

- Bläsius, Karl Hans (2023): Superintelligenz, <https://www.fwes.info/Superintelligenz-2023-1.pdf>,
- Bostrum, Nick (2014): Superintelligenz – Szenarien einer kommenden Revolution. Suhrkamp Verlag
- Chomsky, Noam (2020): Internationalism or Extinction, Routledge (deutsche Ausgabe, 2021: Rebellion oder Untergang!, Westend Verlag)
- Ensel, Leo (2023): Thesen zur Gefährdung unseres Planeten durch Massenvernichtungsmittel und Künstliche Intelligenz, <https://atomkrieg-aus-versehen.de/gefaehrung-unseres-planeten/> (letzter Zugriff: 3.7.2023)
- Grünwald, Reinhard/Kehl, Christoph (2020): Autonome Waffensysteme – Endbericht zum TA-Projekt, Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag, Arbeitsbericht Nr. 187, <https://dip21.bundestag.de/dip21/btd/19/236/1923672.pdf> (letzter Zugriff: 3.7.2023)
- Hendrycks, Dan (2023): AI Safety Newsletter #9, [https://newsletter.safe.ai/p/ai-safety-newsletter-9?utm\\_source=post-email-title&publication\\_id=1481008&post\\_id=126324885&isFreemail=true&utm\\_medium=email](https://newsletter.safe.ai/p/ai-safety-newsletter-9?utm_source=post-email-title&publication_id=1481008&post_id=126324885&isFreemail=true&utm_medium=email) (letzter Zugriff: 3.7.2023)
- Lahl, Kersten (2021): Autonome Waffensysteme als Stresstest für internationale Sicherheitspolitik. In: Politikum, Heft 1, Seite 46 – 53
- Lahl, Kersten/Varwick, Johannes (2022): Sicherheitspolitik verstehen – Handlungsfelder, Kontroversen und Lösungsansätze. Wochenschauverlag, 3. Auflage
- Russell, Stuart (2020): Human Compatible – Künstliche Intelligenz und wie der Mensch die Kontrolle über superintelligente Maschinen behält. Mitp Verlag
- Russell, Stuart/Norvig, Peter (2012): Künstliche Intelligenz – ein moderner Ansatz, 3. Auflage, Pearson
- Safe.ai (2023a): Statement on AI Risk, <https://www.safe.ai/statement-on-ai-risk> (letzter Zugriff: 3.7.2023)
- Safe.ai (2023b): 8 Examples of AI Risk – Misinformation, <https://www.safe.ai/ai-risk#Misinformation> (letzter Zugriff: 3.7.2023)
- Shanahan, Murray (2021): Die technologische Singulariät, MSB Matthes & Seitz Berlin
- Tegmark, Max (2017): Leben 3.0 – Mensch sein im Zeitalter Künstlicher Intelligenz. Ullstein Verlag
- Timm, Ingo J./Siekmann, Jörg/Bläsius, Karl Hans (2020): KI in militärischen Frühwarn- und Entscheidungssystemen, <https://www.fwes.info/fwes-ki-20-1.pdf> (letzter Zugriff: 3.7.2023)
- Urbina, Fabio/Lentzos, Filippa/Invernizzi, Cédric/Ekins, Sean (2023): Dual Use of Artificial Intelligence-powered Drug Discovery, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9544280/> (letzter Zugriff: 3.7.2023)



## Karl Hans Bläsius und Jörg Siekmann

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

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

## FIF-Konferenz 2023

3. – 5. November 2023

### Jugendherberge Berlin Ostkreuz

Seit vielen Jahren stehen die FIF-Konferenzen unter einem inhaltlich klar formulierten Schwerpunkt und werden von einer Regionalgruppe an ihrem jeweiligen Standort organisiert. Für dieses Jahr haben wir uns etwas Besonderes vorgenommen:

*Wir möchten gemeinsam mit euch allen in die Zukunft schauen!*

Die Herausforderungen der Informatik werden nur allzu oft hinsichtlich der Gefahren, Risiken, unerfüllbaren Verheißungen und Dystopien der Digitalisierung betrachtet. Doch wie kann das FIF dazu beitragen, eine kritische Informatik konstruktiv mit einer positiven Ausrichtung zu vereinen?

Aufgrund Eurer Rückmeldungen haben wir drei Schwerpunkte ausgewählt:

- Cyberpeace – z. B. Kritik an der Planung zum Future Combat Air System (FCAS) und zur militärischen Nutzung der Künstlichen Intelligenz
- Information und Nachhaltigkeit – Bits & Bäume – z. B. Informationstechnik und Entwicklungspolitik

- Entwicklungen der Künstlichen Intelligenz – z. B. Chancen und Risiken der Entwicklung und Nutzung von ChatGPT sowie Auswirkungen auf die IT-Sicherheit

Die Konferenz findet vom 3. bis 5. November 2023 in den Räumlichkeiten der Jugendherberge Berlin Ostkreuz statt (<https://www.jugendherbergeberlinostkreuz.de/>). Um Veranstaltungstechnik, Streaming und das leibliche Wohl werden sich Vorstand und Geschäftsstelle kümmern. Sie beginnt Freitagabend mit einem informellen Get-together. Das inhaltliche Programm beginnt am Samstag um 10 Uhr mit Vorträgen u. a. zu generativer KI. Nach der Mittagspause folgt um 14 Uhr ein *Pecha Kucha* rund um Cyberpeace und eine kollektive Performance Lecture *Aesthetic approaches to cyber peace work*.

Der Nachmittag endet mit dem üblichen Jahresrückblick. Nach dem Abendessen werden die Weizenbaum-Studienpreise verliehen. Der Sonntag beginnt ab 10:00 Uhr mit Vorträgen zu verschiedenen Themen. Ab 12:00 Uhr findet die FIF-Mitgliederversammlung statt, die auch für Gäste offen ist.

Falls Ihr Nachfragen habt, wendet Euch bitte per E-Mail an [info@fiffkon.de](mailto:info@fiffkon.de).

*Euer FIF-Vorstand*

## Einladung zur Mitgliederversammlung 2023

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

Wir laden fristgerecht und satzungsgemäß zur ordentlichen Mitgliederversammlung 2023 ein. Sie findet am Sonntag, dem 5. November 2023, ab 12:00 statt.  
Adresse: Jugendherberge Berlin Ostkreuz, Marktstraße 9-12, 10317 Berlin

#### 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 des Vorstands
8. Neuwahl der Kassenprüfer
9. Diskussion über Ziele und Arbeit des FIF, aktuelle Themen, Verabschiedung von Stellungnahmen, Berichte aus den Regionalgruppen
10. Anträge an die Mitgliederversammlung  
*Anträge müssen schriftlich bis drei Wochen vor der Mitgliederversammlung bei der FIF-Geschäftsstelle eingegangen sein*
11. Verschiedenes

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



## IT-Gestaltung für Gute Arbeit

Dagmar Boedicker, Klaus Heß, Katharina Just

## IT-Gestaltung für Gute Arbeit

### Editorial zum Schwerpunkt

*Ach, dieser altbackene Datenschutz! Angesichts der rasanten Entwicklungen in der Informationstechnik und besonders der so genannten Künstlichen Intelligenz und der aktuellen Überwachungs-Möglichkeiten hatten wir ihn schon abgeschrieben, aber siehe da: Manchmal geht da doch was. Beschäftigte von Uber und dem in Deutschland weniger bekannten Unternehmen Ola verlangten Auskunft über die Daten der Unternehmen, die zu ihrer Entlassung wegen Betrugs geführt hatten. Nach der europäischen Datenschutz-Grundverordnung (DS-GVO) klagten sie die Information ein und anschließend die Firmen an. Wegen eines Sachverhalts, der im Englischen bildhaft robo-firing lautet und in die Kategorie automatisierte Entscheidungsfindung gehört. Das Amsterdamer Gericht gab ihnen Recht, denn Profiling und algorithmische Entscheidungen als Basis für das Management seien keine Geschäftsgeheimnisse, sondern personenbezogene Daten.<sup>1</sup> Rechtliche Vorgaben können eine Hilfe sein.*

*Nun lässt sich aber Gute Arbeit nicht allein mit der DS-GVO gestalten und durchsetzen. Auch nicht ausschließlich mit anderen europäischen Rechtsakten wie DMA<sup>2</sup> und DSA<sup>3</sup>, AI Act<sup>4</sup>, der Richtlinie des Europäischen Parlaments und des Rates zur Verbesserung der Arbeitsbedingungen in der Plattformarbeit<sup>5</sup> und wie sie alle heißen. Um sozio-technische Systeme zu gestalten, sind dicke Bretter zu bohren, es dauert, und die Machtverhältnisse bestimmen auch im Betrieb, was geht und was nicht. Und weil Technik Machtverhältnisse nun mal eher verstärkt als sie auszugleichen, gibt es immer noch viel weniger gute Beispiele als kritikwürdige.*

### Wie soll die Technik aussehen, mit der gute Arbeit möglich ist?

Klar ist: Es hängt von der Tätigkeit ab, wie die Technik zu gestalten ist, aber auch von den Menschen, die diese Arbeiten tun. Schon beim Verbrennungsmotor waren nicht alle gleich neugierig auf die Funktionsweise. Nicht alle Beschäftigten möchten wissen, wie eine algorithmische Steuerung abläuft, die die Arbeitsprozesse und die darin aktiven Menschen kontrolliert. Welche Prozesse durch KI im Sinne *Guter Arbeit* besser werden, hängt auch vom Sektor ab. Dienstleistungen in der IT brauchen andere Unterstützung als die in der Logistik.

Gute Arbeit bedeutet Teamarbeit mit qualifizierten Beschäftigten, sie ist sicher und entwicklungsförderlich, wird fair bezahlt, umfasst einen präventiven Gesundheitsschutz und gute Führung, ist beteiligungsorientiert. Die Arbeitszeit wird an den Interessen der Beschäftigten ausgerichtet. Technik wird als Hilfsmittel, nicht zur Steuerung und Kontrolle eingesetzt. In jedem Fall sollen die Menschen die Kontrolle behalten. Wie sich die Anforderung der Intervenierbarkeit in der *Human-Centered Artificial Intelligence* (HCAI) umsetzen lässt, beschreiben Thomas Herrmann und Christoph Lentzsch in ihrem Beitrag. Die These von Moritz Niehaus und Jürgen Klippert ist, dass durch *Manufacturing Execution Systems* (MES) ohne kritische Begleitung von aktiven Betriebsräten die Totalüberwachung von Beschäftigten droht. Sie zeigen aber auch, dass gut gestaltete MES solche Risiken vermeiden können und sogar Arbeitererleichterungen möglich sind.

Wenn es eine Liste der Anforderungen an die jeweilige Technik gibt, die definiert, was sie zur Unterstützung der Beschäftigten leisten sollte, kommt es auf die Machtverhältnisse im Betrieb an, wie weit die Eigentümer der Produktionsmittel diese Anforderungen umsetzen. Markus Hoppe zeigt in seinem Beitrag, wie schwierig es für die Interessenvertretung der Beschäftigten im Betrieb ist, frühzeitig an Informationen zum Transformationsgeschehen zu kommen und anstehende Veränderungen auf Augenhöhe mitzugestalten. Für KMU sieht es schlecht aus, sie haben selten einen Betriebsrat. Stephanie Porschen-Hueck beschreibt ein Projekt, in dem es aber doch funktioniert, Kunden, Beschäftigte und Software-Entwicklung beteiligen sich gemeinsam daran, diese Anforderungen zu definieren. An diesem Projekt zeigt sich, was auch Christoph Schmitz ausführt: Wenn die Technik im Arbeitsprozess eingesetzt werden soll, heißt es, direkt auch die Arbeitsbedingungen gut zu gestalten und von vornherein mitzudenken.

### Die ideale Bot-Freundin war eine perfekte Kombination: sexy und fürsorglich<sup>6</sup>

Wir wünschen uns alle Unterstützung bei der Arbeit, aber wie menschlich wünschen wir uns die Technik dafür? Soll sie eine *Persönlichkeit* haben oder sein? Nutzer von *Replika* und anderen Bots bejahen das, sie wollen und können die *personality* ihrer *fembots/girlfriend bots* gestalten, ein Pygmalion-Effekt im 21. Jahrhundert.<sup>7</sup> Die Dichotomie zwischen Heiliger und Hure, von Klaus Theweleit in seinen „Männerphantasien“

ausführlich für das Geschlechterverhältnis im Tausendjährigen Reich beschrieben, projizieren Menschen – mutmaßlich männlichen Geschlechts – auf Chat-Anwendungen mit KI. Tritt die Technik gerade hundert Jahre Kampf für Frauenrechte und Gleichstellungspolitik in die Tonne? Die Studien-Autorinnen stellen fest:

*As our analysis of the users' imaginaries related to the Replika app shows, users can therefore perceive that they are in control of the user experience, neglecting the fact that they have limited or no access below the interface level to the actual functioning of Replika.*

Wer Replika nutzt, weiß natürlich nicht, wie die KI genau funktioniert, kann sich aber der Illusion hingeben, sie oder zumindest das eigenen Erleben damit zu kontrollieren.

### Was geht und was nicht?

Über IT, die Gestaltungsmöglichkeiten und die Kritik daran ging es schon oft in unserer Zeitschrift, siehe die Schwerpunkte 3/2010 *Gute Arbeit ... mit und trotz IT*, 4/2016 *Zukunft der Arbeit – Arbeit der Zukunft: Wer steuert wen?*, 2/2022 *Künstliche Intelligenz*.

Richtig hilfreiche KI wird in der Arbeitswelt durchaus gewünscht, wie Sabine Pfeiffer anhand ihrer Studienergebnisse zeigt. Ihr Beitrag stellt Ergebnisse einer repräsentativen Erhebung vor und konzentriert sich vor allem auf Beschäftigte im Gesundheits-/ Sozialwesen und der Logistik. Frank Steinwender verweist im Rahmen von Logistikoptimierung, Prozessvereinfachung, Automatisierung oder dem Ersetzen von Arbeit auf die problematischen Seiten der allgemeinen Automatisierung durch die elektronischen Werkzeuge.

Ganz offensichtlich stößt das Betriebsverfassungsgesetz an Grenzen, wenn Software nahezu täglich aktualisiert wird. Weder das Unternehmen noch die Betriebs- und Personalräte können darauf bestehen, dass Betriebsvereinbarungen entweder laufend angepasst oder Software ohne Zustimmung der betrieblichen Interessenvertretung eingesetzt werden. Welche Praktiken bereits im Einsatz sind und ob sie sich bewähren, beschreiben sowohl Schuler/Weichert als auch Hoppe in ihren Beiträgen. Der Gesetzgeber muss tätig werden und die verschiedenen Konzepte auf eine verbindliche, d.h. auch rechtlich erzwingbare Grundlage stellen.

Was leistet IT eigentlich für das Gemeinwohl? Für den schonenden Umgang mit Menschen und mit dem Planeten? Wir hatten uns in der Schwerpunkt-Redaktion ein paar mehr konkrete, vor allem positive Anforderungen gewünscht. Ob es an unserem Aufruf für Beiträge lag oder der gar nicht gelesen wurde: Die abschließende Liste fehlt.

### Einige Anforderungen sollten die IT allgemein und KI besonders erfüllen

IT sollte barrierefrei sein, erklär- und nachvollziehbar und gegen kein geltendes Recht verstoßen. Ihr Nutzen sollte immer größer sein als die Risiken, und die Haftung sollte klar sein. Bestimmte Risiken für das Leben und die Würde von Lebewesen, für die Demokratie und die Natur sind nicht akzeptabel. KI sollte Organisationsformen in allen Sprachen unterstützen, nicht nur denen einer herrschenden Klasse. Sie sollte niemanden benachteiligen oder ausschließen. Menschen sollten weniger arbeiten müssen, ein altes Versprechen der IT. Es ist in vielen Jahrzehnten nicht erfüllt worden. Statt *IT-Gestaltung für Gute Arbeit* heißt es *IT-Gestaltung für gute Gewinne*.

Gleichzeitig müssen wir uns vor solutionistischen Vorstellungen hüten und sollten KI nicht für das Wundermittel halten, mit dem wir eine nachhaltige Entwicklung für die gesamte Menschheit und den Planeten erreichen.

### Regulierung auf EU-Ebene

Zwei EU-Verordnungen sind seit Langem im Werden, DSA und DMA. Die eine ist die Verordnung (EU) 2022/2065 des Europäischen Parlaments und des Rates vom 19. Oktober 2022 über einen Binnenmarkt für digitale Dienste und zur Änderung der Richtlinie 2000/31/EG, das Gesetz über digitale Dienste (GdD), besser bekannt als *Digital Services Act* (DSA). Die andere ist die Verordnung (EU) 2022/1925 des Europäischen Parlaments und des Rates vom 14. September 2022 über bestreitbare und faire Märkte im digitalen Sektor und zur Änderung der EU (Richt-



Abbildung 1: Forderungen der Piraten im Europaparlament, Bild Lorna Schütte, CC BY-SA<sup>9</sup>

linien) 2019/1937 und (EU) 2020/1828, das Gesetz über digitale Märkte (GdM), auf Englisch *Digital Markets Act* (DMA). Am 23. November 2022 ist das GdM in Kraft getreten, mit Wirkung ab 2. Mai 2023, das GdD trat in der EU am 16. November 2022 in Kraft, mit Wirkung ab 17. Februar 2024 und ab 25. August 2023 für große Plattformen. Im GdM geht es darum, die großen Plattformen in ihrer Funktion als Türsteher (Gatekeeper) zu regulieren, die den Zugang zu Informationen und Märkten kontrollieren, und als Oligopole, damit auch die kleinen Plattformen eine Chance haben. Es gibt der Kommission bessere Kontroll- und Sanktionsmöglichkeiten über Suchdienste, Intermediäre (wie die großen App-Stores), sogenannte soziale Dienste, Video- und Kommunikations-Plattformen, Speicher- und Betriebssysteme. Ein umstrittenes Ziel ist die Interoperabilität.<sup>8</sup>

Im GdD geht es darum, inhaltliche Verantwortung der Plattformen zu verschärfen und neue Verpflichtungen einzuführen sowie die Rechte der Nutzerinnen zu stärken. So sollen laut dem Erwägungsgrund 3:

*Bürgerinnen und Bürger der Union als auch andere Personen die ihnen in der Charta der Grundrechte der Europäischen Union (im Folgenden „Charta“) garantierten Grundrechte ausüben können, insbesondere das Recht auf Meinungs- und Informationsfreiheit, auf unternehmerische Freiheit, das Recht auf Nichtdiskriminierung und die Erreichung eines hohen Verbraucherschutzniveaus, [...]*

Der Abgeordnete Patrick Breyer stellt zum GdD fest<sup>10</sup>:

*Inhalte können auch dann rechtmäßig entfernt werden, wenn sie nicht illegal sind. Kein effektives Verbot für irreführende Designs bei Online-Diensten („Dark Patterns“); keine „Do not track“-Option in euren Geräten [...]. Positiv: Verhinderung der „Know-your-Business-Customer“-Identifikationspflichten für alle Dienste [...]. Kein Marketing gegenüber Minderjährigen [...]. Sehr große Online-Plattformen werden zumindest verpflichtet, ein Empfehlungssystem anzubieten, welches nicht auf Profiling basiert.*

Ebenso wichtig ist es, die Plattformarbeit zu regeln. Der Rat der EU liefert die Begründung dafür in einer Grafik aus dem Jahr 2022 (Abbildung 2).

Der Entwurf<sup>12</sup> für einen *Artificial Intelligence Act* (Gesetz über Künstliche Intelligenz) ist noch nicht so weit. Sein Ansatz ist

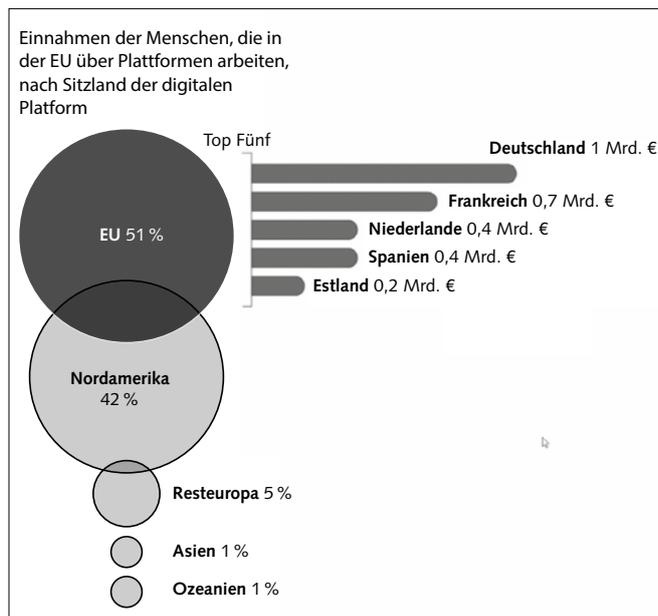
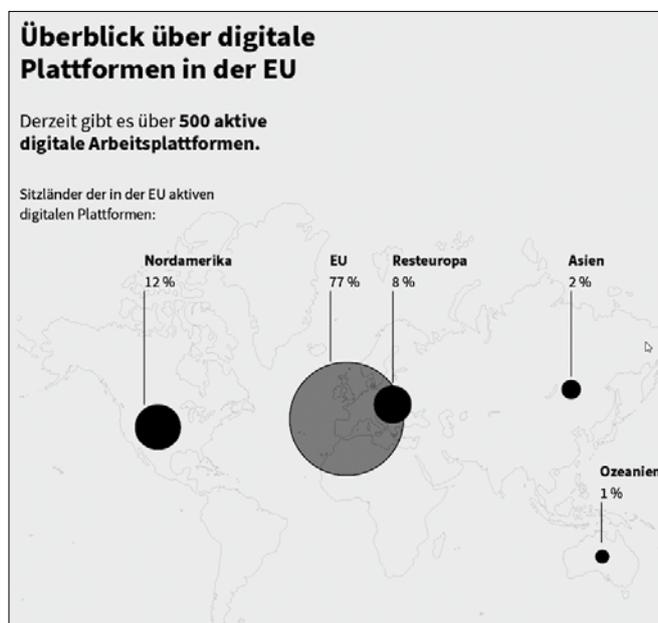


Abbildung 2: Plattformwirtschaft in der EU, Infografik Rat der Europäischen Union, 2022<sup>11</sup>

Risiko-basiert und erfreulicherweise wird er sich wohl nicht auf wolkige ethische Abwägungen beschränken.

## Wie motiviert sind Unternehmen und andere Organisationen?

Möglicherweise ist der Lärm um generative KI größer als ihre Einsatzmöglichkeiten und die ersten Wellen ebben bereits ab. Im S&P 500 fand der Economist 70 Unternehmen, die keinerlei Interesse zeigen. Das Potenzial für die Verbesserung des allgemeinen Lebensstandards und der Einsatz stimmen nicht immer überein, KI wird in den letzten 18 Monaten in Stellenanzeigen nur selten im Gesundheitswesen, dafür um so mehr in Tech-Firmen erwähnt.<sup>13</sup> Noch sind keine massenhaften Entlassungen zu erkennen, manche Unternehmen setzen KI aber als Drohmittel ein, wie AlgorithmWatch im Newsletter vom 26. Mai 2023 schreiben. Sie fassen zusammen:

*Generative Modelle sind keine Arbeitsplatzkiller. Sie werden uns wahrscheinlich dabei helfen, manche Aufgaben zu bewältigen. Sie können uns aber auch noch mehr von unserer Arbeit entfremden und sich negativ auf die Qualität von Dienstleistungen auswirken.*

Beides gilt es zu vermeiden und im besten Fall bleibt Gelegenheit, diskursiv und regulatorisch Anforderungen festzuklopfen.

Für dieses Editorial bedanken wir uns ganz besonders bei AlgorithmWatch, deren Newsletter *Die automatisierte Gesellschaft*<sup>14</sup> eine fast unerschöpfliche Quelle von Nachrichten über automatisierte Systeme darstellt, kurz, verständlich und interessant. Bevor Ihr und Sie ihn abonniert/abonnieren, empfehlen wir die Lektüre dieser *Fiff-Kommunikation*, die hoffentlich ebenso interessant ist.

Das wünscht Euch das Redaktionsteam für diesen Schwerpunkt.

Dagmar, Katharina und Klaus

## Anmerkungen

- 1 <https://www.workerinfoexchange.org/post/historic-digital-rights-win-for-wie-and-the-adcu-over-uber-and-ola-at-amsterdam-court-of-appeal> (Zugriff 2.7.2023)
- 2 <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32022R1925> (Zugriff 9.8.2023)
- 3 <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=OJ:L:2022:277:FULL> (Zugriff 9.8.2023)
- 4 [https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS\\_BRI\(2021\)698792\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf); KI-Grundverordnung (Zugriff 9.8.2023)
- 5 <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A52021PC0762>, zur Zeit im Trilog zwischen Rat, Kommission und Parlament (Zugriff 9.8.2023)
- 6 <https://journals.sagepub.com/doi/full/10.1177/01634437221119021>: [...] *The ideal bot girlfriend was a perfect combination of sexy and caring [...]*
- 7 a. a. O.: *The discussions featured two discursive themes that focused on the AI imaginary of ideal technology and the gendered imaginary of the ideal bot girlfriend. Users expected their AI Replikas to both be customizable to serve their needs and to have a human-like or sassy mind of their own and not spit out machine-like answers. Users thus projected dominant notions of male control over technology and women, mixed with AI and postfeminist fantasies of ostensible independence onto the interactional agents and activated similar scripts embedded in the devices. The vicious feedback loop consolidated dominant scripts on gender and technology whilst appearing novel and created by users.*
- 8 <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32022R1925>
- 9 <https://www.patrick-breyer.de/beitraege/dsa/> (Zugriff 4.8.2023)
- 10 <https://www.patrick-breyer.de/beitraege/dsa/> (Zugriff 4.8.2023)
- 11 <https://www.consilium.europa.eu/de/infographics/platform-economy/> (Zugriff 9.8.2023)
- 12 [https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC\\_1&format=PDF](https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF)
- 13 *Standard & Poors Index der 500 größten Unternehmen der USA. The Economist July 1st 2023, S. 11*
- 14 Zu abonnieren unter: <https://automatedsociety.algorithmwatch.org>



Thomas Herrmann, Christoph Lentzsch

## Der Mensch behält die Kontrolle

### Intervenierbarkeit als Gestaltungsanforderung für die Nutzung von KI

*Zum Hintergrund: Zurzeit werden vielfältige Risiken diskutiert, die mit dem Einsatz der künstlichen Intelligenz verbunden sein können. Dazu gehören vor allem der Verlust von Arbeitsplätzen, die Beeinträchtigung der Autonomie und Entscheidungsfreiheit von Menschen, Schäden, die durch Qualitätsmängel der KI entstehen können, Verzerrungen (Bias) in KI-Entscheidungen, die bestimmte gesellschaftliche Gruppen benachteiligen etc. Die Risiken wiegen besonders schwer, wenn KI direkt mit physikalischen Prozessen gekoppelt wird – sei es im Straßenverkehr, in der Produktion, bei militärischer Nutzung etc. – oder wenn von KI getroffene Entscheidungen automatisch oder unhinterfragt umgesetzt werden. Ein Beitrag zur Milderung dieser Risiken ist es, Menschen im Umgang mit KI Eingriffsmöglichkeiten zu eröffnen. So sagt der DFKI-Experte Paul Lukowicz im Interview mit der Tagesschau am 14. Juni 2023: „Problematisch ist es immer dann, wenn die KI Dinge tut, ohne dass der Mensch eingreifen oder noch einmal drauf schauen kann.“*

Draufschaufen können und Eingreifen gehören unseres Erachtens untrennbar zusammen, um Menschen kontinuierlich Kontrolle

über KI zu gewähren – man muss ja reagieren können, wenn man sieht, dass etwas schief läuft. Diese Kontrollmöglichkeiten

zu gestalten, ist allerdings nicht trivial. Wir schlagen hierfür Interventionsschnittstellen zwischen Menschen und automatisierten Prozessen oder automatischer Entscheidungsfindung vor. Die Umsetzung von Eingriffsmöglichkeiten zu verwirklichen ist durchaus komplex, wie wir am Beispiel der Interventionsschnittstelle erläutern wollen. Sie erfordert Maßnahmen sowohl auf der Ebene der KI selbst und der sie umgebenden technischen Infrastruktur als auch auf der Ebene der organisatorischen Praxis. Es ist also eine soziotechnische Herausforderung, Intervention zu ermöglichen. Ziel muss es sein, dass Nutzerinnen die KI weder kleinteilig überwachen noch routinemäßig immer wieder eingreifen müssen. Wäre das erforderlich, wäre es eher angezeigt, die KI-Lösung auszutauschen. Vielmehr sollte eine Überwachung nur zu bestimmten Anlässen oder Zeitpunkten erforderlich sein und die Intervention eine Ausnahme darstellen. Das bedeutet, dass es nicht angemessen ist, wenn Nutzerinnen permanent misstrauisch gegenüber KI sein müssen. Das würde für sie unnötigen Kontrollaufwand und Mehrarbeit nach sich ziehen. Sie dürfen aber auch nicht zu gutgläubig sein, weil damit erhebliche Risiken einhergehen. Es ist also wichtig, Vertrauen zu kalibrieren – *trust calibration* (Okamura & Yamada, 2020). Auch diese Kalibrierung lässt sich durch Intervention unterstützen: Sie korrigiert nicht nur Ergebnisse oder Abläufe der KI, sondern ermöglicht es, KI zu erkunden und mit ihr zu experimentieren – so können die Anwenderinnen Fähigkeiten entwickeln und verbessern, die zur professionellen Nutzung von KI erforderlich sind.

## Intervention

Schmidt und Herrmann (2017) definieren *intervenierende Interaktion* als Aktivität von Benutzerinnen, mit der sie das Verhalten eines automatisierten Prozesses *zeitweise* verändern. Zur Veranschaulichung ziehen diese Autoren das Einparken eines Autos heran. Bei einem herkömmlichen Auto parkt der Benutzer selbst ein und übt eine kontinuierliche und feinschrittige Kontrolle aus. Alle 100-500 Millisekunden bewertet man die Situation, kontrolliert oder korrigiert die Lenkung und fährt oder bremst. In einem Auto mit Einparkautomatik steuert man den Vorgang nur insoweit, dass man die Parklücke aussucht und den Parkvorgang startet. Die feinschrittige Steuerung beim Lenken und Vor- und Zurückbewegen ist automatisiert – man muss dabei nicht im Auto sitzen. Bei einem autonomen, potenziell fahrerlosen Auto entfernt man sich vom Auto, das dann selbstständig einen Parkplatz findet und einparkt.

Intervention bei einem Auto mit Einparkautomatik kann beispielsweise bedeuten, dass der Fahrer sieht, wie sich eine Gruppe eines Kindergartenausflugs nähert und er das Einparken daher für eine Weile stoppt, bis die Gruppe vorbei gegangen ist. Das wäre die einfachste Art der Intervention – stoppen und wieder aufnehmen. Anders ist es, wenn eine Fahrerin z. B. merkt, dass das Auto beim Vorwärtsfahren auf die ungeschützte Bewässerungsfläche eines Baumes rollen wird und sie das vermeiden möchte. Dann kann sie die Kontrolle komplett übernehmen und selbst einparken, sie übernimmt also bei dieser Intervention die feinschrittige Steuerung für eine Weile selbst. Sie könnte aber auch intervenieren, indem sie nur die Sensorik beeinflusst und sich mutig als Hindernis zwischen Baum und Auto aufstellt, so dass das Rückwärtsfahren früher eingeleitet würde (gut wäre es, wenn die Automatik das als menschliche Intervention erkennt

und sich das Hindernis merkt, damit man nicht die ganze Zeit vor dem Auto stehen bleiben muss). Oder sie beeinflusst nur die Entscheidungsfindung der Automatik, indem sie für das Vorwärtsfahren für diese spezielle Situation festlegt, dass der Richtungswechsel bereits bei einem größeren Abstand zum Hindernis – hier dem Baum – vorgenommen wird. Oder sie übernimmt nur die Aktion des Richtungswechsels, indem sie hierfür jedes Mal beispielsweise auf eine Taste drückt.

Viele KI-gestützte Prozesse laufen im Regelfall ab, ohne dass Menschen aktiv etwas beitragen. Sie laufen nach einem Plan ab und bzw. oder in Abhängigkeit von den registrierten Ereignissen im Umfeld des KI-Systems. *Intervenierende Interaktion* ist eine Ad-hoc-Veränderung mit Hilfe einer außergewöhnlichen, ungeplanten Steuerung durch Menschen, die nur für einen bestimmten Zeitraum oder für einen eingegrenzten Bereich der automatisierten Abläufe (etwa der Interpretation von Sensordaten) wirksam ist. Eine *Interventions-Benutzungsschnittstelle (intervention user interface)* ermöglicht eine beiläufige Wahrnehmung von automatisierten Prozessen, sie stellt mögliche Schritte für eine Intervention den Benutzerinnen anschaulich dar oder schlägt sie sogar vor, und bietet Steuerungsmöglichkeiten an, mit denen eine Intervention einfach, schnell sowie robust vollzogen und auch beendet werden kann. Bereits an dem Beispiel zum Einparken kann man sehen, dass es vielfältige Möglichkeiten der Umsetzung gibt, die bei der Gestaltung einer Interventionsschnittstelle zu berücksichtigen sind. Intervention kann mit unterschiedlichem Aufwand stattfinden, vom einfachen Abschalten für einen vorgegeben Zeitraum über das zeitweise Ändern von Parametern bis zur kompletten, feinschrittigen Übernahme der Steuerung für einen bestimmten Zeitabschnitt.

Wir sprechen von *Intervenierbarkeit*, wenn Systeme über eine Interventionsschnittstelle verfügen.<sup>1</sup>

## Menschenzentrierte KI und erweiterte Intervention

Intervenierbarkeit bei KI ist eine Gestaltungsoption, die einem umfassenderen Forschungsbereich zuzuordnen ist, nämlich der *Human-Centered Artificial Intelligence*, kurz HCAI. Dieser menschenzentrierten KI kann man verschiedene Ansätze zuordnen, wie *responsible AI*, *keeping the human in the loop*, *collaborative AI*, *hybrid intelligence systems*, *explainable AI (XAI)* etc. Am klarsten wird HCAI durch die Grundannahme gekennzeichnet, dass Mensch und Computer gemeinsam mehr erreichen als jeweils für sich alleine, und dass in der gemeinsamen Bearbeitung von Aufgaben beide miteinander lernen und ihre Fähigkeiten weiterentwickeln (Dellermann et al., 2019). Auch die Erklärbarkeit von AI (XAI) – in der Regel durch KI selbst – ist unabdingbar, denn Eingriffe und Kontrolle durch Menschen sind nur möglich, wenn man erkennen kann, inwiefern eine Intervention möglich und sinnvoll ist, wo sie ansetzen kann und welchen Effekt sie bewirken wird. Man könnte argumentieren, dass es doch nicht notwendig sei, den Algorithmus des KI-basierten Einparkens zu kennen, wenn man einfach selbst die Steuerung übernehmen kann, damit das Auto nicht zu dicht an einen Baum heranfährt. Wir meinen, dass zum Beispiel die Fahrerin schon wissen sollte, ob die KI das einfach nicht kann oder ob am Auto etwas nicht so funktioniert, wie es soll, wenn etwa eine Kamera an der Fahrzeugunterseite verschmutzt ist. Wenn die KI nicht

berücksichtigt, ob das Auto auf die Bewässerungsfläche eines Baumes fährt, dann ist es für die Einflussnahme des Menschen wichtig zu wissen, ob das daran liegt, dass dafür die Sensorik nicht eingebaut ist oder dass das System die Daten nicht richtig interpretiert.

Die ursprüngliche Definition von Intervention geht auf KI-gesteuerte Echtzeitabläufe ein. HCAI legt es nahe, dass Interventionsbarkeit nicht nur auf automatisierte physikalische Prozesse, wie etwa bei autonomen Fahrzeugen bezogen wird, sondern auch auf KI-basierte Entscheidungsfindung. Es gibt KI-unterstützte Prozesse, wie etwa die Identifizierung von Sozialbetrug<sup>2</sup>, die automatisch Konsequenzen für die Betroffenen zeitigen, ohne dass menschliche Entscheiderinnen noch eingreifen können. Dagegen müssen Menschen in derartige Prozesse eingreifen können, bei denen eine KI-Komponente eine Entscheidung liefert, die dann von menschlichen Akteuren oder von herkömmlicher Software umgesetzt wird. Menschen müssen gegen KI-basierte Entscheidungen ein Veto einlegen können (Rakova et al., 2021). HCAI geht davon aus, dass Menschen die besseren Entscheider sind, sobald es nicht mehr *nur* um die Beherrschung komplexer Zusammenhänge und riesiger Datenmengen geht, sondern um Entscheidungen bei Uneindeutigkeit, im Kontext von Verhandlungen oder bei unzureichender Informationslage.

Ein weiteres reales Beispiel für Anwendungen, die nicht den Abläufen bei der automatischen Fahrzeugsteuerung ähneln, sind solche Warenwirtschaftssysteme, die eine KI-Komponente enthalten. Eine solche Komponente kann etwa entscheiden, dass eine bestimmte Ware nachbestellt wird. Daran anschließend kann dann ein weiteres KI-System den Lieferanten auswählen und den Transport disponieren, oder die Bestellung wird von einer herkömmlichen IT-Infrastruktur ausgeführt oder von menschlichen Sachbearbeiterinnen. Egal, welche der drei Optionen realisiert wird: Bevor es weitergeht, muss ein menschlicher Entscheider die KI-Vorgabe überprüfen können, um ggf. zu intervenieren und für die spezielle Situation den Bestellzeitpunkt oder die Menge der Ware anzupassen. Ähnliches gilt – als weiteres Beispiel für einen Anwendungsbereich – für Industrieanlagen, bei denen die KI automatisch Wartungshinweise gibt. Sensorbasiert könnte etwa festgestellt werden, dass aufgrund von Druckschwankungen eine Pumpe ausgewechselt werden muss. Bevor nun der weitere übliche Prozess durch einen Reparaturauftrag gestartet wird, müssen menschliche Entscheiderinnen prüfen, ob bei der Pumpe wirklich ein Defekt zu erwarten ist und die mit dem Austausch verbundenen Kosten gerechtfertigt sind. Gerade wenn solche Systeme der vorausschauenden Wartung eingeführt werden, gibt es anfangs häufig Fehlalarme, die zu unnötigen Reparaturen führen würden, wenn man sie alle ernst nähme. Intervention muss also für ganzheitliche, soziotechnische Prozesse ermöglicht werden, in denen KI, herkömmliche IT und menschliche Akteure zusammenwirken (Herrmann et al., 2018).

Wenn wir Intervention als neues Prinzip oder Gestaltungskriterium einführen, dann ist die Frage berechtigt, ob eine solche zusätzliche Option neben den vorhandenen Kriterien wirklich nötig ist. Reicht nicht der klassische *Not-Aus-* oder *Not-Halt-Knopf* als Metapher, um zu verdeutlichen, um was es geht (siehe Abbildung 1)? Gerade in Abgrenzung zum Not-Aus-Knopf kann man gut die Besonderheit der Intervention erklären. Nehmen wir zum

Beispiel den Bremsassistenten in LKWs, der verhindern soll, dass schwere Auffahrunfälle passieren. Es kommt immer wieder vor, dass sich bei Auffahrten an der Autobahn PKWs sehr dicht vor einem LKW einfädeln und der Assistent einen unerwünschten Bremsvorgang auslöst. Daher schalten manche Fahrerinnen den Assistenten in solchen Situationen ab. Sie unterlassen es jedoch, ihn wieder einzuschalten. Dadurch steigt wieder die Unfallgefahr. Ein Interventionsknopf würde dagegen nur eine befristete Unterbrechung der Automatik auslösen, wobei der Assistent nach einer vorkonfigurierten Zeit wieder eingeschaltet wird und dies entsprechend signalisiert. Interventionen sind also stets begrenzt und unterbrechen den regulären Prozess nicht dauerhaft. Drückt man den *Not-Halt-Knopf*, dann ist das anschließende Wiederhochfahren des Systems häufig aufwändig. Beim Drücken eines *Interventionsknopfes* sollten in der Regel nur Teile des Systems angehalten werden und die erneute Inbetriebnahme würde entweder automatisch oder ebenfalls mit Hilfe des Knopfes oder eines einfachen Mechanismus erfolgen.

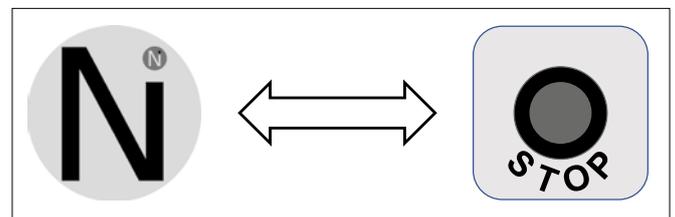


Abbildung 1: Intervention versus Not-Halt alle Abbildungen Thomas Herrmann

Intervention ist damit etwas Drittes, das zwischen den beiden Polen des einfachen Weiterlaufen-Lassens und der kompletten Abschaltung liegt, wie in Abbildung 2, oben, gezeigt.

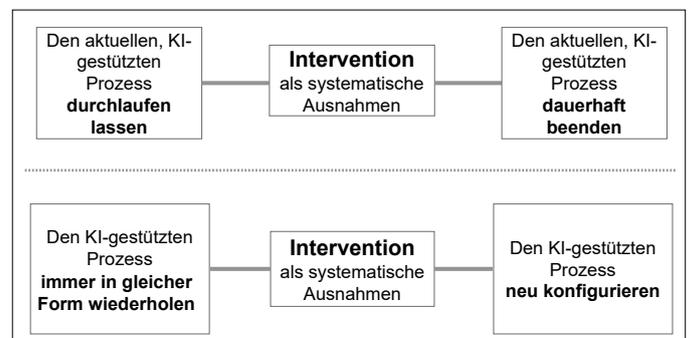


Abbildung 2: Intervention – eine dritte Option

### Intervention und Weiterentwicklung von KI

Heutige KI-Systeme werden sehr selten nur in einem einzigen Schritt erstellt, also etwa angelern oder trainiert, und dann dauerhaft in derselben Form genutzt. Vielmehr werden sie meistens regelmäßig weiter trainiert, um bessere Vorhersagen, Entscheidungen, Aktionen etc. zu liefern. Wir sprechen davon, dass das KI-System neu konfiguriert wird. Intervention ist demgegenüber keine dauerhaft wirkende Neukonfiguration, sondern eine vorübergehende Anpassung für eine Ausnahme. Intervention sollte nicht genutzt werden, um in ähnlichen Situationen immer wieder den gleichen Korrektoreffekt zu erreichen. Intervention ist also, wie Abbildung 2, unten, ebenfalls verdeutlicht, etwas Drittes zwischen zwei weiteren Polen, nämlich dass man zum einen den Prozess wiederholt immer in der gleichen Form ablaufen

lässt, oder dass man zum anderen den Ablauf mit dauerhaftem Effekt neu konfiguriert. Mit Intervention ist weder der eine noch der andere dieser beiden Pole gemeint. Vielmehr eröffnet sie eine dritte Möglichkeit.

Zum Verständnis hilft das Beispiel von Bürokaffeemaschinen, die über einen Drehknopf verfügen, mit dem man die Menge des Kaffeepulvers oder des verwendeten Brühwassers über das voreingestellte Maß hinaus erhöhen kann. Das ist im Grunde eine Interventionsmöglichkeit. Eine Interventionschnittstelle ist also kein völlig neues Phänomen. Wichtig ist es jedoch, dass sie insbesondere bei KI-Anwendungen *systematisch* umgesetzt wird. Beim Interventionsprinzip werden bei der Betätigung des Drehknopfs der Kaffeemaschine die Mengen nur für einen einzelnen Brühvorgang angepasst; würde der Eingriff zu einer dauerhaften Anpassung der Mengen führen, wäre es eine Neukonfiguration. Bei einer Intervention kann sich der nächste Kaffeetrinker darauf verlassen, dass die vorkonfigurierten Mengenverhältnisse die gleichen wie immer sind. Hier wird im soziotechnischen Zusammenspiel mehrerer Akteure deutlich, wie entscheidend es ist, dass die Wirkung einer Intervention begrenzt ist: Immer dann, wenn mehrere Menschen KI-Anwendungen nutzen, müssen sie sich darauf verlassen können, dass sie die Ausgangsbedingungen wie erwartet vorfinden. Besonders wichtig ist diese Anforderung dann, wenn es zu aufwändig oder unerwünscht ist, allen Beteiligten eine personalisierte Benutzungsschnittstelle anzubieten.

Stellt sich jetzt zum Beispiel beim Bremsassistenten heraus, dass bei jeder Autobahnauffahrt der Interventionsknopf gedrückt wird, oder es wird bei der Kaffeemaschine deutlich, dass fast alle Nutzerinnen die Menge des Brühwassers erhöhen, dann ist das ein Anlass, über eine Weiterentwicklung oder Neukonfiguration nachzudenken. Zu diesem Zweck müssen Interventionen registriert und dokumentiert werden können. KI könnte selbst Vorschläge machen, wann und wie eine Neukonfiguration für die Verbesserung eines KI-Systems sinnvoll ist, dass also etwa durch Auswertung von Geoinformation an allen Ausfahrten der Brem-

assistent unterdrückt und zeitgleich ein deutlich wahrnehmbares Hinweissignal in der Fahrzeugkabine gegeben wird.

Zusammengefasst lässt sich sagen, dass aus seltenen Interventionen nicht notwendigerweise ein Bedarf abzuleiten ist, die KI neu zu konfigurieren. Wenn es also nur selten vorkommt, dass ein Auto nahe einer ungeschützten Baumfläche parkt, dann ist Intervention ausreichend, ohne dass etwas neu konfiguriert werden muss. Wenn es aber um den täglich genutzten Parkplatz nahe der eigenen Wohnung geht, dann ist eine Anpassung sinnvoll. Ähnlich ist der Fall gelagert, wenn fast alle Menschen mit dem gleichen Merkmal, etwa doppelte Staatsbürgerschaft, systematisch des Sozialbetrugs verdächtigt werden. Bei häufigen, gleichartigen Interventionen ist eine dauerhafte Anpassung angezeigt. Es hängt dann vom konkreten Anwendungsfall ab, ob das System diese Interventionen selbst auswerten kann, um Verbesserungen vorzuschlagen, ob die Benutzerinnen selbst kleine Anpassungen vornehmen oder ob eine Verbesserungsanpassung an dafür zuständige Expertinnen delegiert wird. Jede vorgenommene Neukonfiguration muss anschließend transparent und erklärbar sein, weil sie auch die Interventionseffekte ändern kann.

### Intervention und Weiterentwicklung menschlicher Kompetenz

Letztlich müssen Menschen entscheiden, ob und wie eine Neukonfiguration erfolgen soll. Diejenigen, die häufig Interventionen vornehmen, sollten bei dieser Entscheidung mithelfen und können selbst Vorschläge unterbreiten, welche Art der Anpassung angemessen ist. Das setzt voraus, dass eine Intervention immer wieder reflektiert wird. Um diese Reflexion zu unterstützen, muss KI ihre Schlussfolgerungsmechanismen erklären können sowie die Daten und Einflussfaktoren, die dabei berücksichtigt werden. Nur wenn man weiß, wo die Defizite wirklich liegen – etwa auf der Ebene der Sensorik, der Interpretation der Daten, der automatisierten Entscheidung oder bei deren Umsetzung –



### Thomas Herrmann und Christopher Lentzsch

Prof. Dr.-Ing. **Thomas Herrmann** lehrte von 1986-2004 *Informatik und Gesellschaft* an der Universität Dortmund und ist seit 2004 am Institut für Arbeitswissenschaft der Ruhr-Universität Bochum. Seit März 2022 ist er Seniorprofessor mit den Forschungsschwerpunkten Analyse und Design soziotechnischer Systeme, menschenzentrierte Künstliche Intelligenz und Prozessmanagement.

**Christopher Lentzsch** hat an der Ruhr-Universität Bochum Angewandte Informatik studiert und promoviert dort zur Gestaltung von Intervention User Interfaces für soziotechnische Systeme. In seiner Forschung hat er sich zudem mit Datenschutz und IT-Sicherheit in den Bereichen Internet-of-Things und WWW beschäftigt.

kann man sinnvolle Anpassungsmaßnahmen vorschlagen. Interventionen selbst und die dabei beobachtbaren Effekte helfen, die KI besser zu verstehen und auch den Anwendungsbereich, in dem die KI eingesetzt wird. Die Schwächen der KI, die durch eine Intervention ausgebügelt werden, helfen darüber nachzudenken, welche Besonderheiten der Arbeitsbereich hat, in dem KI eingesetzt wird.

Intervention ist auch eine Möglichkeit zu experimentieren, Was-wäre-wenn-Fragen nachzugehen und die KI sowie ihren Einsatzbereich zu erkunden. Nehmen wir noch einmal das Beispiel der KI-basierten vorausschauenden Wartungshinweise für Industrieanlagen. Angenommen, es erfolgt ein Hinweis, dass eine Pumpe wegen zu hoher Druckschwankungen auszutauschen ist. Der Anlagenführer kann diese Warnung absichtlich ignorieren, weil er sehen will, ob die Pumpe nicht doch noch länger funktionstüchtig ist. Vielleicht legt er sich schon einmal eine Ersatzpumpe zum schnellen Austauschen bereit, aber er will abwarten, ob die Vorhersage tatsächlich stichhaltig ist, um so etwas über die Zuverlässigkeit des Systems zu lernen. Vielleicht erhöht er auch den Grenzwert, bis zu dem seiner Meinung nach Druckschwankungen noch tolerierbar sind und nicht als Anzeichen für einen baldigen Defekt gewertet werden sollen. Eine solche vorübergehende, probeweise Erhöhung eines Grenzwerts wäre auch eine Intervention, mit deren Hilfe man etwas über die Wartungserfordernisse lernen kann. Der Anlagenführer könnte noch weitergehen und alte Fälle inspizieren, in denen korrekter Weise vor einem Pumpendefekt gewarnt wurde. Er könnte dann überprüfen, ob diese korrekten Warnungen auch unter der Bedingung des erhöhten Grenzwerts noch erfolgt wären. Wie man an solchen Beispielen sieht, kann sich durch Intervention ein Lernfortschritt für die Benutzerinnen ergeben, der dann in manchen Fällen auch zu einer Weiterentwicklung der KI führt.

Abbildung 3 stellt den Zusammenhang zwischen Lernfortschritten der Benutzerinnen und Weiterentwicklung der KI verallgemeinert dar. Interventionen können dazu dienen, Arbeitsbereiche zu erkunden oder KI-Effekte zu korrigieren oder beides. Menschen erleben die Wirkung ihrer Intervention und ihre Angemessenheit und verstehen so die eigene Arbeitspraxis besser.

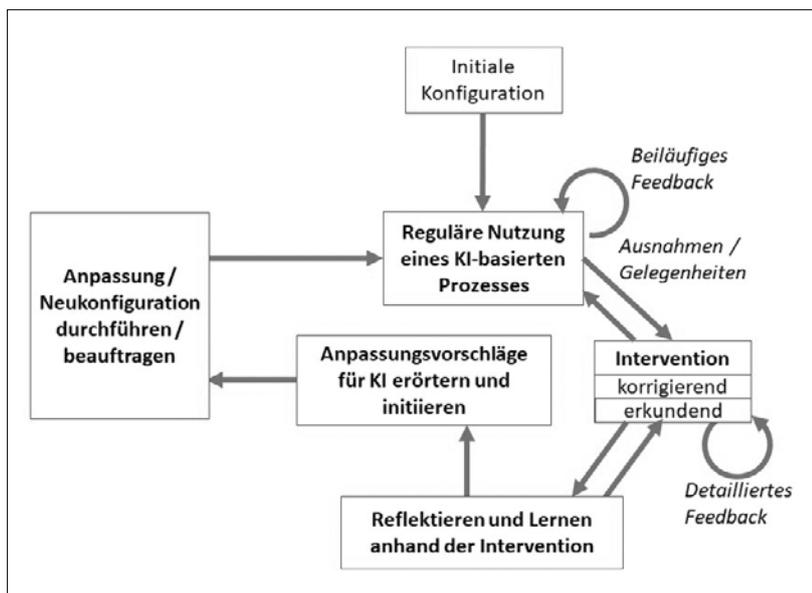


Abbildung 3: Weiterentwicklung von Menschen und KI

Vorschläge zur Anpassung der KI entstehen und werden ggf. umgesetzt. Falls sich in unserem Beispiel die erkundende Intervention der Grenzwertserhöhung als sinnvoll erweist, kann das Anlass für eine Neukonfiguration durch eine dauerhafte Erhöhung des Grenzwerts sein. Das dürfte insbesondere dann der Fall sein, wenn häufigere Interventionen durch mehrere Benutzerinnen zeigen, dass eine dauerhafte Anpassung sinnvoll ist.

Dieser doppelte Effekt ist ganz im Sinne der menschenzentrierten KI (HCAI), der zufolge das Tandem Mensch-KI nicht nur Aufgaben besser löst, sondern auch zu einer Weiterentwicklung der Fähigkeiten auf beiden Seiten beiträgt.

### Voraussetzung für Intervenierbarkeit

Viele Gestaltungskriterien beziehen sich direkt auf technische Eigenschaften der KI, wie etwa *Fairness*, der zufolge KI niemanden benachteiligen darf. Intervenierbarkeit ist dagegen ein Gestaltungskriterium, das sich auf die Interaktion zwischen Menschen und KI bei der Bearbeitung von Aufgaben – vorrangig in der Arbeitswelt – bezieht. Damit Intervenierbarkeit gelingt, muss die Interaktionsschnittstelle zur KI einige Anforderungen erfüllen:

- Es müssen Informationen angeboten werden und/oder auf einfache Weise abrufbar sein, mit Hilfe derer man entscheiden kann, ob und wie man interveniert.
- Die Bedienungselemente müssen – so wie beim Not-Aus-Knopf – leicht und schnell erreichbar und einfach zu handhaben sein. Es wird nicht immer ein physikalischer Knopf vorgesehen sein, sondern auch Schaltflächen auf dem Bildschirm oder dem Smartphone. Nach dem Auslösen einer Intervention wird man je nach Anwendungsfall mehrere Optionen angeboten bekommen, um eingreifen zu können.
- Es muss nachvollziehbar sein, wann die Intervention beginnt und endet. Das gilt nicht nur für diejenigen, die die Intervention starten, sondern für alle, die von dem veränderten Ablauf der KI-Nutzung betroffen sind.

- Intervention muss schnell beendbar sein und ihre Effekte müssen ggf. revidierbar sein. Das ist eine wichtige Voraussetzung, um Intervention zum Zweck der Erkundung zu nutzen.
- Interventionen müssen dokumentierbar sein, damit sie als Grundlage für weitere Anpassungen genutzt werden können. In der Arbeitspraxis wird es angemessen sein die Gründe festzuhalten, deretwegen man interveniert.
- Intervention muss sowohl sofort bestimmte Wirkung erzielen können (wie bei der Erhöhung der Wassermenge im Kaffee) als auch in der Zukunft liegende Effekte bewirken können, wie es etwa bei der Heizungssteuerung durch den sogenannten Partyknopf vorgesehen ist, der ausnahmsweise für eine Nacht den Zeitraum bis zum Beginn der Nachtabsenkung verlängert.

- Intervenierbarkeit muss auf verschiedenen Ebenen möglich sein: Es können die sensorgestützte Datenaufnahme beeinflusst werden, die Interpretation und Planung aufgrund der Daten sowie die KI-basierte Entscheidungsfindung oder das Auslösen von Aktionen, wie oben am Beispiel des Einparkens verdeutlicht (siehe auch die drei Möglichkeiten der Beeinflussung in Abbildung 4).

Die Erfüllung dieser Möglichkeiten trägt dazu bei, dass Menschen in die Nutzung der KI einbezogen bleiben, das entspricht dem Prinzip *Keeping the Human in the Loop*. Dieses Prinzip ist aber nur erfüllbar, wenn die organisationale Praxis seine Umsetzung auch zulässt und fördert. Herrmann und Pfeiffer (2022) argumentieren am Beispiel der vorausschauenden Wartung, dass Human in the Loop durch das Prinzip *Keeping the Organization in the Loop* ergänzt werden muss. Es ist offensichtlich, dass sich viele Beschäftigte nur trauen werden, ein KI-Ergebnis zurückzuweisen – also etwa einer Wartungsanforderung nicht zu folgen –, wenn ihr organisatorisches Umfeld eine solche Entscheidung absichert und auch dazu ermutigt.

Abbildung 4 stellt diesen Zusammenhang dar und benennt auf der rechten Seite vier zentrale Aspekte, die in der organisatorischen Praxis des KI-Einsatzes realisiert sein müssen: Die Organisation muss die Art des Einsatzes der KI regeln. Dazu gehört, dass Beschäftigte die KI-Prozesse und Entscheidungen beobachten und reflektieren können und hierzu auch befähigt werden. Das gilt auch für die Möglichkeit, KI-Ergebnisse flexibel zu handhaben, also ggf. zurückzuweisen oder anzupassen, wie es bei der Intervention der Fall ist. Beschäftigte sollten zum Beispiel dazu angehalten werden, sich zuerst selbst zu überlegen, welche Entscheidung sie treffen würden, bevor sie sich den Entscheidungsvorschlag der KI anschauen. Eine solche Maßnahme hilft dabei, das Vertrauen in die KI zu kalibrieren. Weiterhin müssen Beschäftigte wissen, dass ihre Vorgesetzten und Kolleginnen es anerkennen, wenn sie intervenieren, um gegenzusteuern oder das Leistungsvermögen der KI zu testen. Außerdem muss es Bestandteil organisatorischer Praxis sein, dass sowohl die KI als auch die Kompetenzen der Beschäftigten kontinuierlich weiterentwickelt werden, und dass alle aufgefordert sind dazu beizutragen. Diese Weiterentwicklung muss insbesondere die regelmäßigen fachlichen, technischen, gesetzlichen und marktbezogenen Veränderungen und Randbedingungen berücksichtigen, die für einen KI-Einsatz als Kontext relevant sind.

Intervenierbarkeit ist also eine soziotechnische Herausforderung, zu deren Bewältigung die Technik, die Kompetenzen der Menschen und auch die organisatorische Praxis aufeinander abzustimmen sind. Grundsätzlich muss Intervenierbarkeit so gestaltet sein, dass der Nutzen, den man durch eine Intervention gewinnt, die damit verbundene Anstrengung überwiegt. Das Konzept der Intervenierbarkeit verdeutlicht, wie man das übergeordnete Gestaltungskriterium der *Kontrollierbarkeit* bei KI-Anwendungen umsetzen kann.

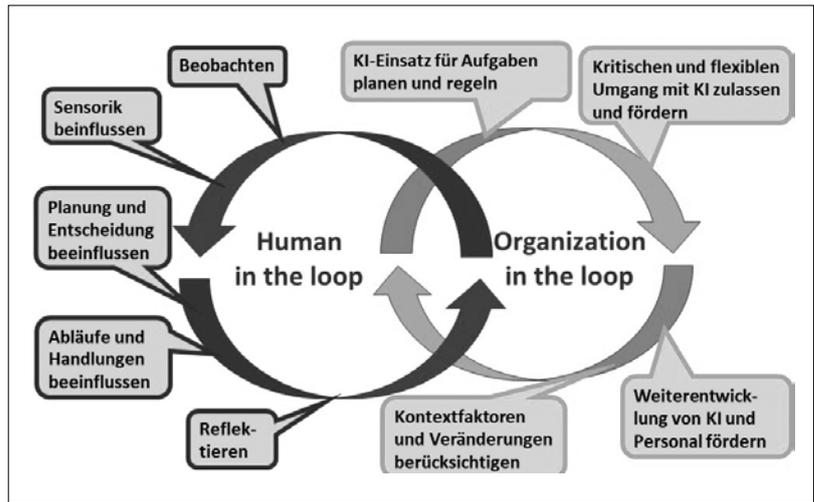


Abbildung 4: Integration organisatorischer Praxis

## Referenzen

- Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S. & Ebel, P. (2019). The future of human-AI collaboration: A taxonomy of design knowledge for hybrid intelligence systems. *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS)*.
- Herrmann, T., Pfeiffer, S. (2023). Keeping the organization in the loop: a socio-technical extension of human-centered artificial intelligence. *AI & Soc* 38, 1523–1542. <https://doi.org/10.1007/s00146-022-01391-5>
- Herrmann, T., Schmidt, A. & Degeling, M. (2018). From Interaction to Intervention: An Approach for Keeping Humans in Control in the Context of socio-technical Systems. *Proceedings of the 4th International Workshop on Socio-Technical Perspectives in IS Development (STPIS'18)*. <http://ceur-ws.org/Vol-2107/Paper8.pdf>
- Okamura, K. & Yamada, S. (2020). Adaptive trust calibration for human-AI collaboration. *PLOS ONE*, 15(2), e0229132. <https://doi.org/10.1371/journal.pone.0229132>
- Rakova, B., Yang, J., Cramer, H. & Chowdhury, R. (2021). Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1–23.
- Rost, M. & Bock, K. (2011). Privacy By Design und die Neuen Schutzziele. *Datenschutz und Datensicherheit*, 35(1), 30–35.
- Schmidt, A. & Herrmann, T. (2017). Intervention User Interfaces: A New Interaction Paradigm for Automated Systems. *Interactions*, 24(5), 40–45. <https://doi.org/10.1145/3121357>

## Anmerkungen

- 1 Intervenierbarkeit ist ein Prinzip, das auch im Datenschutz vorgeschlagen wird (Rost & Bock, 2011), wenn es darum geht, dass Betroffene direkt die Speicherung oder Verarbeitung ihrer personenbezogenen Daten beeinflussen können sollen.
- 2 Siehe <https://www.heise.de/news/Kuenstliche-Intelligenz-Drei-Beispiele-fuer-das-Scheitern-von-Algorithmen-9205911.html>, letzter Abruf am 31.7.2023.



# Manufacturing Execution Systems

## Zwischen Überwachung und Guter Arbeit

Schwerpunkt

Der Hype um „Industrie 4.0“ als vermeintlich entscheidender Vorteil im kapitalistischen Wettbewerb ist abgeflacht. Für uns als Gewerkschafter:innen wird das Thema in der betrieblichen Praxis nun jedoch relevant: In den Fabrikhallen erleben viele Beschäftigte aktuell genau das, was mit Industrie 4.0 angekündigt wurde: Unternehmen sind bestrebt, kundenindividuell in Losgröße 1 zu produzieren bei Kosten einer Massenfertigung (acatech, 2013, 19 ff.). Für Beschäftigte bedeutet das den Wandel ihrer Arbeit durch digitale Fertigungssteuerung mittels Manufacturing Execution Systems (MES). Gerade im Zusammenspiel mit künstlicher Intelligenz ergeben sich Potenziale zur Steigerung der Produktivität. Unsere These: Ohne kritische Begleitung von aktiven Betriebsräten droht durch MES die Totalüberwachung von Beschäftigten. Wir zeigen in diesem Artikel auf, dass gut gestaltete MES solche Risiken vermeiden können und sogar Arbeitserleichterungen möglich sind.

Im ersten Abschnitt erläutern wir kurz die Funktionsweise von MES. Den Kern unseres Beitrags bildet anschließend ein Blick auf die Praxis: Wie werden MES als IT-Systeme in Betrieben eingeführt und welche Rolle spielen Betriebsräte dabei? Anhand der Beispiele zeigen wir, dass Betriebsräte und IG Metall sich mit ihrer sozio-technischen Betrachtungsweise von einem eher technik-zentrierten Verständnis des Managements unterscheiden. Welche ist die zentrale Herausforderung für Interessenvertretungen? Wie können Betriebsräte das nötige Verständnis für die Technik gewinnen, um Auswirkungen auf die Arbeitsbedingungen abzuschätzen? Zum Abschluss möchten wir aufzeigen, mit welchen Angeboten die IG Metall eine gute Gestaltung von MES begleitet.

### Wie MES-Software in der Fabrik plant und steuert

MES verbinden und vernetzen die eher kaufmännisch geprägten Software-Produkte für Enterprise Resource Planning (ERP, beispielsweise SAP) mit den vielen Anwendungen in der Fertigung. Dort wird mit Systemen zur Maschinensteuerung sowie Software zur Produktionsplanung gearbeitet (siehe Abbildung 1).

In der Produktion ohne MES werden die anstehenden Aufträge häufig per Hand aus dem ERP-System abgerufen und den Ar-

beitsstationen zugeteilt. Dies kann über eine lokale Software geschehen oder ausgedruckt auf Papier. Menschen übernehmen hier eine zentrale Schnittstelle zwischen dem Hallenboden und den ERP-Systemen, in die fertige Aufträge dann wiederum händisch eingetragen werden.

Viele Unternehmen streben danach, diese Lücke zu schließen, um ihre Effizienz zu steigern und Kosten zu senken. Mittels MES können Produktionsplaner:innen in einem einzigen Programm beispielsweise sehen, ob das benötigte Material für einen Produktionsauftrag bereits vollständig geliefert wurde und ob die zur Montage benötigten Teile in der Metallbearbeitung schon fertig gestellt worden sind. Sie reihen den Auftrag dann in den Produktionsplan ein oder dieser wird (teilweise mittels künstlicher Intelligenz) durch die Software optimiert. Anschließend holen sich Arbeiter:innen in der Montage anstehende Kundenaufträge per Klick oder Touchpad an ihre jeweiligen Bearbeitungsstationen auf dem Hallenboden. Am Ende markieren Produktionsbeschäftigte einen Auftrag im System selbstständig als erledigt und laden die nächste Aufgabe. Die Schnittstelle, Aufträge zu verteilen und deren Erledigung händisch in ein (ERP-) System einzutragen, wird somit durch das MES automatisiert.

Dies ist aber nur ein Beispiel der vielen Funktionen, die ein MES als zentrale Plattform zur Steuerung und Optimierung üblicher-

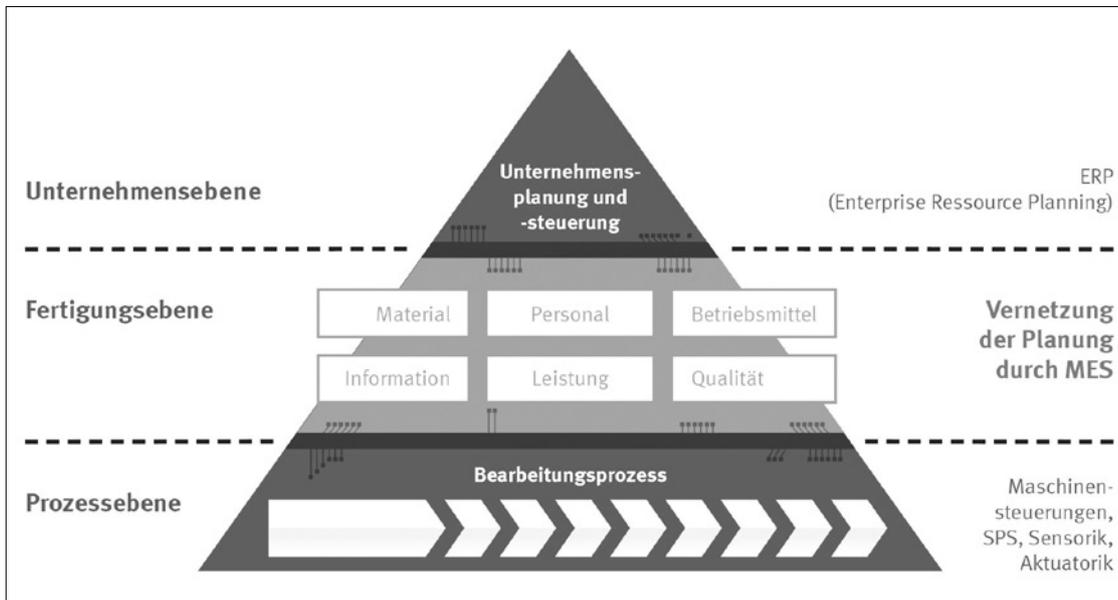


Abbildung 1: Lückenschluss (eigene Darstellung)

weise erfüllt. Hinzu kommen beispielsweise Energiemanagement, Verwaltung von Lagerbeständen und Übernahme von Bestellungen. Zur Vertiefung siehe die VDI-Richtlinie 5600, die beispielsweise in TBS NRW (2018, 18ff) dargestellt ist.

### Überwachung als Preis für effizientere und leichtere Arbeit?

MES sind Teil eines *algorithmischen Managements* (Schaupp, 2022), sie verteilen Arbeitsaufträge und überwachen deren Abarbeitung. Gleichzeitig ermöglichen die Systeme Vorteile für Beschäftigte: Facharbeiter:innen an Maschinen können sich beispielsweise ganz der Metallbearbeitung widmen und müssen keine Zeit mehr mit der Suche nach Konstruktionszeichnungen oder den richtigen Werkzeugen für ihre Bearbeitungszentren verbringen. Montagearbeiter:innen erhalten bei typen- und variantenreicher Fertigung Unterstützung. Besonders wenn sie sich unsicher sind, wie der nächste Montageschritt exakt auszuführen ist, kann der direkte Zugriff auf die Montageanleitung hilfreich sein.

Doch eines muss Beschäftigten und ihren Interessenvertretungen klar sein: MES dienen der Datensammlung. Die Kolleg:innen hinterlassen bei jedem Arbeitsschritt digitale Spuren im System. Beispielsweise ist mit Datum und Uhrzeit gespeichert, wann jemand die Bearbeitung an einer Fräsmaschine begonnen hat. Die Maschinen und Anlagen hinterlassen ebenfalls Datenspuren. Die Vorgesetzten sehen, wann eine Maschine stillstand, weil sie eine Störung hatte oder gerüstet wurde. Es ist gespeichert, wie lange Maschinenarbeiter:innen für Rüstvorgänge gebraucht haben und wie lange Instandhalter:innen an der Störungsbeseitigung gearbeitet haben. Alle Prozessschritte sind bis ins Detail nachvollziehbar.

Die Situation erinnert an Orwell'sche Zukunftsvisionen der Totalüberwachung. Unternehmen versprechen sich durch den Einsatz von MES Ziele zu erreichen, die mit Industrie 4.0 propagiert wurden: Eine bessere Planung soll für bessere Auslastung von Mensch und Maschine sorgen. Der optimale Informationsfluss soll Fehler in Bearbeitung und Montage reduzieren, während die Produktion variantenreicher wird. Mehr Transparenz soll Einsparpotenziale offenlegen, etwa bei Energie- und Materialkosten.

Angesichts dieser Entwicklungen drängt sich jedoch die Frage auf, ob die Überwachung der Beschäftigten ein akzeptabler Preis für diese Effizienzsteigerungen und Arbeiterleichterungen ist. Wie lassen sich also Schutz der Persönlichkeitsrechte und Effizienzziele unter einen Hut bringen? Hierauf wollen wir in den folgenden Abschnitten mit Beispielen eine Antwort finden.

### Gestaltung von MES: Betriebsräte sind gefordert

Wenn MES eingeführt werden, wird dies im Betrieb oft unter einem anderen Namen angekündigt – manchmal auch gar nicht. Häufige Umschreibungen sind *Maschinen- und Betriebsdatenerfassung* oder *smart factory*. Dabei stellen Manager die uneingeschränkte Datensammlung oft als alternativlos dar.

Für die Beschäftigten und den Betriebsrat ist das in der Regel eine schwierige Ausgangssituation: Sie müssen erkennen, dass es sich nicht um ein einzelnes Projekt handelt, sondern um eine umfassende Vernetzung des Betriebs über alle Ebenen. Noch schwieriger ist es, die Arbeitsfolgen von MES für verschiedene Gruppen von Beschäftigten abzuschätzen. Dazu dürfte es meist eine externe Beratung brauchen, d. h. der Betriebsrat muss Sachverständige hinzuziehen, um die Technik und ihre Folgen zu verstehen.

#### Beispiel 1: MES ohne individuelle Kontrolle

So bat der Betriebsrat eines Produktionsbetriebs für Industriearmaturen um die Unterstützung der IG Metall. Das Management hatte Vorarbeiten zur Einführung einer *papierlosen Fertigung* begonnen, ohne Betriebsrat oder betroffene Beschäftigte zu beteiligen. Zusammen mit einer Unternehmensberatung hatten die Manager eine MES-Software ausgewählt, von der sie annahmen, dass sie für die betrieblichen Zwecke passt. Erst anschließend wurde der Betriebsrat überhaupt informiert, dass ein MES eingeführt werde.

Das Management wollte unmittelbar mit einer Erprobung beginnen: An einer Fertigungsinsel mit Dreh- und Fräsmaschinen sowie in einem angeschlossenen Montagebereich sollte das MES pilotiert werden. Vom Betriebsrat erwarteten die Projektleiter Zustimmung. Mit Unterstützung der IG Metall erarbeiteten sich die Betriebsratsmitglieder im ersten Schritt einen Überblick, was *papierlose Fertigung* bedeutet. Daraufhin wurde ihnen klar, dass es hier um die vollständige Vernetzung der Prozesse geht.

Der Betriebsrat erkannte einige Vorteile von MES und stellte sich deshalb nicht gegen das Projekt. Für eine Pilotphase schloss er eine Vereinbarung ab, um ein MES im Betrieb einzuführen. Eine Fertigungsinsel mit acht Dreh- und Fräsmaschinen wurde für den Piloten ausgewählt. Die MES-Planner sahen vor, dass sich die Bediener:innen mit ihrem digitalen Betriebsausweis an den Maschinen anmelden. Die Eingabe der Betriebsdaten wäre dann personenbezogen erfasst worden, beispielsweise Uhrzeit des Auftragsbeginns, Unterbrechung der Arbeit und Gründe dafür. Der Betriebsrat griff hier ein, um diese umfassende Leistungs- und Verhaltenskontrolle zu unterbinden. Der Arbeitgeber wollte jedoch nicht auf die Ermittlung dieser Leistungsdaten verzichten. In Verhandlungen arbeiteten Betriebsrat und Projektleitung heraus, dass bei der Ermittlung der Kennzahlen ein Personenbezug nicht notwendig ist. Der Betriebsrat schlug vor, dass es spezielle Gruppenausweise geben soll. Mit diesen können sich die Kolleg:innen anmelden und sind dann nur als Mitglied einer Gruppe zu identifizieren. Diese Lösung floss in eine Betriebsvereinbarung ein, die parallel zur Pilotphase verhandelt wurde.

Um mehr Leistungsdruck durch strengere Vorgabezeiten zu verhindern, konnte der Betriebsrat erstreiten, dass nicht einzelne Schichten ausgewertet werden dürfen, sondern immer nur ganze 24 Stunden. Ebenso ist in dieser Betriebsvereinbarung für den Regelbetrieb festgeschrieben, dass aus dem MES gewonnene Informationen nicht zum Nachteil von Beschäftigten genutzt werden dürfen. Bei Missbrauch von MES-Daten sieht die Vereinbarung arbeitsrechtliche Konsequenzen vor. Und falls durch den MES-Einsatz die Wertigkeit einer Tätigkeit abnimmt,

hat das Unternehmen unverzüglich Maßnahmen zu ergreifen, um eine Beschäftigung sicherzustellen, die der bisherigen Entgeltgruppe entspricht.

Das Beispiel zeigt die Notwendigkeit der sozio-technischen Gestaltung (z. B. mittels des *Kompass Digitalisierung: IG Metall*, 2019b). Das heißt, dass die drei Dimensionen Mensch, Technik und Organisation gleichrangig betrachtet werden (vgl. Hirsch-Kreinsen, 2018, S. 22–25). In diesem Fall bringt die Technik das Problem der Leistungs- und Verhaltenskontrolle mit sich. Es kann verhindert werden durch eine technische (Gruppenausweise, gröbere Auswertung) und eine organisatorische Maßnahme (Verbot der Datennutzung unter Androhung von Strafe).

### Beispiel 2: Begrenzter Zugriff

Im zweiten Fall handelt es sich um ein arbeitspolitisch sehr erfahrenes Betriebsratsgremium am Standort eines Großkonzerns. Die Werksleitung kündigte die Einführung eines MES an. Das für MES zuständige Betriebsratsmitglied ist neben seinem Ehrenamt in der IT-Abteilung tätig. Das Gremium hatte durch diese Expertise wenig Mühe, die Folgen der vollständigen Vernetzung von Prozessen abzuschätzen. Beim Abschluss einer Betriebsvereinbarung hatte der Betriebsrat allerdings eine Klippe zu umschiffen, um die Leistungs- und Verhaltenskontrolle einzuschränken.

Denn die hergestellten Produkte sind sicherheitsrelevante Geräte, von deren einwandfreier Funktion Menschenleben abhängen. Bundesbehörden oder die Staatsanwaltschaft dürfen aufgrund gesetzlicher Regelungen weitgehende Informationen zum Herstellungsprozess und insbesondere zu den Qualitätsprüfungen einfordern. Die Geschäftsleitung muss in einem solchen Fall offenlegen, welche Person die Qualitätsprüfung bei mangelhaften Geräten durchgeführt hat.

Vor der MES-Einführung wurde in einem solchen Fall in Papierdokumenten recherchiert, die verschlossen waren. Der Aufwand hierzu war hoch, doch nun erfolgt die Recherche im leichter zugänglichen MES. Dies vereinfacht eine Suche und individuelle Nachvollziehbarkeit erheblich. Es gibt Unternehmen, die solche Situationen ausnutzen, um die vollständige Transparenz des MES als scheinbare Notwendigkeit darzustellen. Dann werden entweder rechtliche Vorschriften oder Anforderungen wichtiger Kunden wie Automobil-Endhersteller (OEM) ins Feld geführt.

In diesem Fall war für den Betriebsrat und auch die Geschäftsleitung klar: Eine so weitgehende Recherche personenbezogener Daten mittels MES darf im normalen Betrieb nicht möglich sein. Der Betriebsrat forderte daher Beschränkungen und erreichte in den Verhandlungen für eine Betriebsvereinbarung folgendes: Der Betriebsrat muss informiert werden, falls es eine Anfrage von Staatsanwaltschaft oder Bundesbehörden gibt. Erst anschließend darf die Recherche im MES starten. Es gibt zudem einen klar definierten und sehr kleinen Personenkreis, der so weitgehende Zugangsberechtigungen im MES erhält, dass er diese Recherchen durchführen kann. Auch hier zeigt sich das sozio-technische Zusammenspiel von Technik und organisatorischen Maßnahmen, um das Risiko der individuellen Kontrolle zu minimieren.

## MES strategisch gestalten

Es wäre notwendig, die MES vor ihrer Einführung zu gestalten, also die Programme zu modifizieren. Aber praktisch eröffnet sich den betroffenen Beschäftigten und Betriebsräten dazu nur selten die Möglichkeit (Carls et al., 2023). Vielmehr wählt das Unternehmen meist ohne deren Beteiligung einen MES-Anbieter aus, womit Funktionen und Software-Design bereits weitgehend feststehen. Im Nachhinein kann die grundlegende Struktur der Programme nicht mehr geändert werden. Deshalb sind Steuerungskreise, paritätisch besetzt von Betriebsrat und Management, besonders wichtig. Solche zusätzlichen Gremien, die allerdings nicht gesetzlich vorgesehen sind, können vom Projektstart an wichtige Entscheidungen treffen. Hierzu gehört beispielsweise die Auswahl eines MES-Anbieters und von Software-Modulen, was über traditionelle Mitbestimmungsrechte hinausgeht.

Sollte sich die Unternehmensseite nicht auf einen Steuerungskreis oder ähnliche Strukturen einlassen, ist es umso wichtiger, dass die Beschäftigten und ihre Interessenvertretung rote Linien definiert haben: Was sind Mindestanforderungen, die bei und nach einer MES-Einführung gewahrt sein müssen? Hierzu eignet sich ein *Leitbild für Gute Arbeit* (mit MES), das unter breiter Beteiligung vieler Beschäftigter erarbeitet werden sollte (IG Metall, 2019a; für das Konzept der Guten Arbeit siehe Schmucker, 2020).

Offt ist das Management bei MES als vermeintlichem IT-Projekt sehr auf die Technik fokussiert, die aber für eine gelungene Einführung nur einen Teilaspekt darstellt. Vielmehr handelt es sich bei einem Betrieb um ein sozio-technisches System. Demnach sind Mensch, Technik und Organisation gleich bedeutsam und stehen in wechselseitigen Abhängigkeiten (siehe z. B. Hirsch-Kreinsen, 2018, S. 25).

### Mitbestimmungsrechte und Beteiligung

In Betrieben mit Betriebsrat hat dieser eine Reihe von Mitbestimmungs- und Informationsrechten bei der Einführung eines MES, auf die hier nicht im Detail eingegangen werden kann. Für einen Überblick bietet sich die unten angegebene Broschüre (IG Metall, 2020) an und zur Vertiefung der ausführlichere Ratgeber der TBS NRW (2018). Nur kurz genannt seien § 87 (1) 6 des Betriebsverfassungsgesetzes, der ein Mitbestimmungsrecht bei möglicher Leistungs- und Verhaltenskontrolle vorsieht. Neben diesem sehr offensichtlichen Feld gibt es im Arbeits- und Gesundheitsschutz weitere starke Einflussmöglichkeiten des Betriebsrats auf MES, v. a. über § 87 (1) 7 BetrVG in Verbindung mit § 3 (1) des Arbeitsschutzgesetzes und der Betriebsicherheitsverordnung. Was trocken bis staubig klingen mag, hat in der praktischen Bedeutung große Wirkung: Vor Inbetriebnahme eines MES muss das Unternehmen eine Gefährdungsbeurteilung durchführen, die auch psychische Belastungen berücksichtigt.

Ohne einen Betriebsrat haben die Beschäftigten allerdings keinerlei Mitbestimmungsrechte bei der Einführung von MES oder anderen IT-Systemen. Es ist daher dringend zu empfehlen, einen Betriebsrat zu gründen, falls noch keiner vorhanden ist. Bei einem solchen Vorhaben unterstützt die jeweilige Gewerkschaft vor Ort, die vorab kontaktiert werden sollte.

Aber natürlich ist es mit der Existenz eines Betriebsrats nicht getan, sondern es kommt darauf an, ob und wie dieses Gremium (rechtliche) Möglichkeiten wahrnimmt. Die Beteiligung der Beschäftigten spielt hierbei eine zentrale Rolle für die erfolgreiche Gestaltung von MES (vgl. Kämpf & Langes, 2021). Wie auch bei anderen Techniken und Themenfeldern sollte der Betriebsrat möglichst viele (betroffene) Beschäftigte in die Gestaltung der Systeme einbeziehen, deren Expertise nutzen und Anliegen aufgreifen. Das Betriebsverfassungsgesetz bietet hierzu mit Absatz 2 des § 80 vielfältige Möglichkeiten zur Beteiligung (IG Metall, 2022), damit es gar nicht erst zu einer Stellvertreterpolitik kommt.

### Was bietet die IG Metall an?

Um Betriebsräte und Beschäftigte beim Thema MES zu unterstützen, hält die IG Metall mittlerweile ein breites Angebot bereit. Dabei kommt der gewerkschaftlichen Bildungsarbeit eine zentrale Rolle zu. In Seminaren können sich Betriebsratsmitglieder und hauptamtliche Gewerkschaftssekretär:innen zu MES informieren und ihre eigenen Fälle diskutieren. Hierzu bietet die IG Metall seit 2020 Praxisworkshops in Kooperation mit verschiedenen wissenschaftlichen Instituten an: Unter dem Titel *Digitalisierung der Produktionssteuerung – MES nutzen und gestalten* können Teilnehmende, beispielsweise in der Lernfabrik der Ruhr-Universität Bochum, die Anwendung von MES selbst ausprobieren. Der Fokus solcher Workshops liegt auf den Handlungsmöglichkeiten des Betriebsrats, um Chancen und Risiken von MES einschätzen zu können mit dem Ziel, bessere Arbeitsbedingungen für die Beschäftigten durchzusetzen.

Unsere Erfahrungen hierbei sind positiv. Die Möglichkeit, ein MES in der Lernfabrik selber zu erfahren, fördert das konkrete *Begreifen* der Möglichkeiten von MES. Die Teilnehmenden an den Seminaren sagen, dass ihnen in diesen drei Tagen vieles klargeworden ist, was sie zuvor in abstrakten Abhandlungen über MES bereits gelesen hatten.

Damit solches Wissen langfristig bewahrt und weitergegeben wird, hat die IG Metall 2022 das *Betriebsrätenetzwerk MES* gegründet. Hier kommen – virtuell und in Präsenz – Betriebsratsmitglieder zusammen, die sich mit MES befassen. Neben der Qualifizierung ist ein zentrales Ziel, MES in den Betrieben so zu gestalten, dass die Grundsätze von Guter Arbeit verwirklicht sind (vgl. IG Metall, 2019a, 2020). Beispielhafte Themen sind die Gestaltung attraktiver Arbeitsaufgaben in Produktionsabläufen, die von MES gesteuert sind; oder wie Betriebsräte Entgeltverluste verhindern können, wenn sich Arbeitsinhalte durch eine MES-Einführung verändert haben.

Als niedrigschwelliges Angebot und zur Vermittlung von Grundlagenwissen eignet sich die Broschüre *Gute Arbeit mit MES* (IG Metall, 2020). Die Autoren aus dem Ressort *Zukunft der Arbeit* der IG-Metall-Vorstandsverwaltung bieten außerdem Beratung an für Betriebsräte und ihre hauptamtlichen Betreuer:innen in den über 140 IG-Metall-Geschäftsstellen.

### Fazit

Die *Industrie 4.0* ist in Form der digitalen Fertigungssteuerung mit MES in den Betrieben angekommen. Die geschilderten Praxisbeispiele zeigen, dass es eines engagierten Betriebsrats bedarf; sonst droht eine lückenlose Überwachung der Beschäftigten mittels personenbezogener Daten (exemplarisch in Schaupp, 2022). Eine gut begleitete Einführung von MES bietet hingegen auch die Chance für Arbeitserleichterungen. So kann Transparenz in Logistik und Werkzeugverwaltung die Beschäftigten beispielsweise von der unbeliebten Suche nach Material und Arbeitsmitteln entlasten. Bei dieser möglichen Verbesserung muss aber darauf geachtet werden, dass solche gewonnenen Freiräume nicht zu einer Belastung der Beschäftigten werden, indem ihre Tätigkeit verdichtet und Leistung intensiviert wird (Carls et al., 2023, S. 86).



### Moritz Niehaus und Jürgen Klippert

**Moritz Niehaus** ist politischer Sekretär in der IG-Metall-Vorstandsverwaltung, Frankfurt am Main. In der Zentrale von Europas größter Gewerkschaft arbeitet er als Referent im Ressort *Zukunft der Arbeit*. Inhaltliche Schwerpunkte sind die Digitalisierung von (Industrie-)Arbeit und die sozial-ökologische Transformation. Ziel seiner Vorträge, Veröffentlichungen und Seminare ist es, zu besseren Arbeits- und Lebensbedingungen beizutragen.



**Jürgen Klippert**, Dr. habil., politischer Sekretär im Ressort *Zukunft der Arbeit* in der IG-Metall-Vorstandsverwaltung, Frankfurt am Main. Schwerpunktthema: Entwicklung der Arbeitsqualität im Rahmen der digitalen Transformation. Arbeitsschwerpunkte: humanzentrierte und partizipative Gestaltung von sozio-technischen Systemen. Dr. Klippert ist zudem Lehrbeauftragter für Arbeitssystemgestaltung an der Universität Kassel.

Die Gestaltung der Systeme ist also ein umkämpftes Feld. Hierbei gibt es für den Betriebsrat eine besondere Schwierigkeit: Er sollte schon in den Gestaltungsprozess eingreifen, bevor traditionelle Mitbestimmungsrechte greifen, auch wenn die Tragweite der MES-Einführung noch nicht absehbar ist. Dies wird aktuell immer wichtiger, weil MES zunehmend in Kombination mit *Künstlicher Intelligenz* eingesetzt werden, um die immensen Datensammlungen auszuwerten.

Letztlich wird Industrieunternehmen im marktwirtschaftlichen Wettbewerb kaum eine Alternative zur Einführung von MES bleiben, wenn sie produktiv und damit wettbewerbsfähig bleiben wollen. Solange es bei der aktuellen Wirtschaftsordnung und Eigentumsverhältnissen an Produktionsmitteln bleibt, wäre die reine Abwehr solcher Systeme ein Risiko für den Fortbestand von Betrieben und Arbeitsplätzen. Deshalb wäre es kurzfristig, Betriebsräten zur vollständigen Ablehnung von MES zu raten. Vielmehr gilt es, die Systeme – möglichst frühzeitig – so zu gestalten, dass sie Gute Arbeit ermöglichen und potenzielle Nachteile für Beschäftigte verhindert werden. Betriebsräte sollten zumindest Leistungs- und Verhaltenskontrolle unterbinden, um darüber hinaus bessere Arbeitsbedingungen für die Beschäftigten zu erstreiten.

## Referenzen

acatech. (2013). Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: Abschlussbericht des Arbeitskreises Industrie 4.0. Promotorengruppe Kommunikation der Forschungsunion Wirtschaft – Wissenschaft.

Carls, K., Gehrken, H., Kuhlmann, M., Splett, B. & Thamm, L. (2023). Digitalisierung aus Beschäftigtensicht: Fehlende Mitgestaltung, belastender Zusatzaufwand, mangelnde Wertschätzung. *WSI*, 76(2), 83–92. <https://doi.org/10.5771/0342-300X-2023-2-83>

Hirsch-Kreinsen, H. (2018). Einleitung: Digitalisierung industrieller Arbeit. In H. Hirsch-Kreinsen, P. Ittermann & J. Niehaus (Hrsg.), *Digitalisierung industrieller Arbeit: Die Vision 4.0 und ihre sozialen Herausforderungen* (2. Aufl., S. 13–32). Nomos.

IG Metall. (2019a). Der Weg zu einem Leitbild. Gute Arbeit in der Digitalisierung. Zukunft der Arbeit konkret. [https://www.igmetall.de/download/20230621\\_Der\\_Weg\\_zu\\_einem\\_Leitbild\\_Gute\\_Arbeit\\_in\\_der\\_Digitalisierung\\_b0117d7122a9f9539d4e05df6b2f4229fc2e185e.pdf](https://www.igmetall.de/download/20230621_Der_Weg_zu_einem_Leitbild_Gute_Arbeit_in_der_Digitalisierung_b0117d7122a9f9539d4e05df6b2f4229fc2e185e.pdf)

IG Metall. (2019b). Kompass Digitalisierung: Eine Gestaltungshilfe für gute digitale Arbeit. [http://www.steps-projekt.de/fileadmin/Templates/Steps\\_Projekt/Media/20190524\\_Kompass\\_Digitalisierung\\_final.pdf](http://www.steps-projekt.de/fileadmin/Templates/Steps_Projekt/Media/20190524_Kompass_Digitalisierung_final.pdf)

IG Metall. (2020). Gute Arbeit mit MES: Mensch-Organisation-Technik bei Manufacturing Execution Systems. [https://www.igmetall.de/download/20201005\\_ZdA\\_MES\\_Online\\_cf327f89eda86af0d5bc1bda6b1737ce21a4b7f5.pdf](https://www.igmetall.de/download/20201005_ZdA_MES_Online_cf327f89eda86af0d5bc1bda6b1737ce21a4b7f5.pdf)

IG Metall. (2022). Auskunftspersonen § 80 Abs. 2 Satz 4 BetrVG: Chance für die Vertrauensleutearbeit und die Einbeziehung von Beschäftigten. [https://www.igmetall.de/download/20230310\\_Handlungshilfe\\_Auskunftspersonen\\_def1f87e65b91431b5a40d3a48aa250915f9856c.pdf](https://www.igmetall.de/download/20230310_Handlungshilfe_Auskunftspersonen_def1f87e65b91431b5a40d3a48aa250915f9856c.pdf)

Kämpf, T. & Langes, B. (2021). Künstliche Intelligenz in der digitalen Arbeitswelt: Erste Befunde einer empirischen Bestandsaufnahme. *spw Zeitschrift für sozialistische Politik und Wirtschaft*, 44(6), 26–31. [http://spw.de/data/247\\_kaempfung\\_langes.pdf](http://spw.de/data/247_kaempfung_langes.pdf)

Schaupp, S. (2022). COVID-19, economic crises and digitalisation: How algorithmic management became an alternative to automation. *New technology, work and employment*. Vorab-Onlinepublikation. <https://doi.org/10.1111/ntwe.12246>

Schmucker, R. (2020). Von Decent Work zu Guter Arbeit. Zur Bedeutung arbeitspolitischer Leitbilder. *sozialpolitik.ch*. Vorab-Onlinepublikation. <https://doi.org/10.18753/2297-8224-157>

TBS NRW. (2018). Mit MES-Systemen und Kennziffern auf dem Weg zu Industrie 4.0: Aufbau und Funktionsumfang, Gestaltungs- und Regelungsansätze (1. Aufl.). Reihe Arbeit, Gesundheit, Umwelt, Technik: Heft 86. Technologieberatungsstelle beim DGB NRW e. V.



Markus Hoppe

## Künstliche Intelligenz in der betrieblichen Praxis

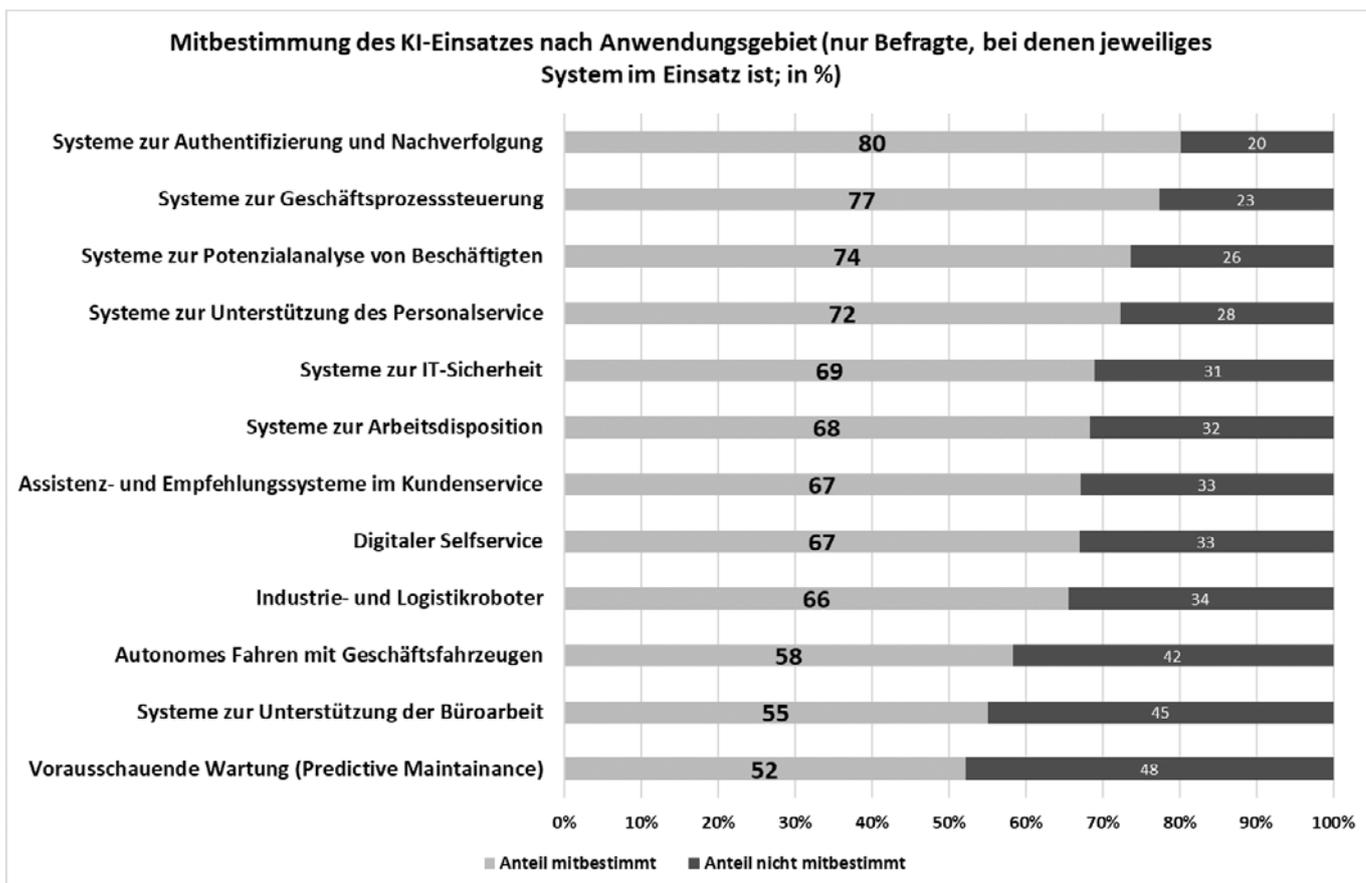
### Durch Mitbestimmung und Beteiligung die Transformation meistern.

*Künstliche Intelligenz (KI) wird zunehmend integraler Teil der Unternehmenspraxis. Obwohl Betriebs- und Personalräte durch ihre Mitbestimmungs- und Beteiligungsrechte Einfluss auf das betriebliche KI-Geschehen haben, sind viele von ihnen skeptisch. Das liegt u. a. auch daran, dass sie sich unzureichend beteiligt sehen. In diesem Beitrag wird gezeigt, dass betriebliche Interessenvertretungen aus der Defensive gelangen können, wenn sie sich aktiv in die KI-Transformation einmischen und auch die Partizipation der Beschäftigten einfordern. Denn dadurch können sie maßgeblich zur Akzeptanz von KI und einer gelingenden Praxis beitragen.*

### Welche Rolle spielt KI inzwischen in der betrieblichen Praxis?

Mit KI sind im Grundsatz IT-Systeme gemeint, die in der Lage sind, selbständig Aufgaben zu lösen, für die bisher menschliche Intelligenz nötig war, und die aus Daten eigenständig Schlussfolgerungen ziehen und sich im Echtzeitbetrieb durch Lerneffekte aus bisherigen Datenverarbeitungen oder zusätzlichem Dateninput selbst verbessern können. Verschiedene Studien zeigen (beispielsweise Sevindik 2022), dass die KI-Verbreitung

steigt und inzwischen etwa jedes zehnte Unternehmen aktiv KI-Technologien einsetzt. Während sich die öffentliche Debatte um KI und die aus ihr resultierenden Risiken für Arbeitsplätze und Geschäftsmodelle gerade an der Sprachassistenten-Software ChatGPT erneut entzündet, geht es konkret in Unternehmen und Verwaltungen auch um Beiträge von KI zur Verbesserung betrieblicher Prozesse und zur Gestaltung von Arbeitsbedingungen unter Berücksichtigung von Job- und Datenschutzrisiken (Giering 2022). Hinsichtlich der zu erwartenden Auswirkungen von KI sind nicht wenige zurückhaltend. Oft werden



*Mitbestimmung des KI-Einsatzes nach Anwendungsgebieten (eigene Darstellung)  
Quelle: INPUT Consulting 2023*

nachteilige Effekte befürchtet, wenn es um die Beschäftigungssicherheit, die Kontrolle am Arbeitsplatz, die Algorithmisierung von Managementfunktionen oder auch die Arbeitsbelastungen geht. Die Skepsis ist insbesondere bei betrieblichen Interessenvertretungen stark ausgeprägt. Als Ursachen lassen sich u.a. die als mangelhaft erlebten Möglichkeiten der Einflussnahme und Mitgestaltung anführen. Diese sind aber notwendig für den Aufbau von Vertrauen und Routine, weil betriebliche Interessenvertretungen nur damit eine Akteursrolle in der digitalen Transformation erhalten. Bislang herrscht bei vielen die Wahrnehmung vor, dass sie Getriebene sind vom Arbeitgeber und dem technologischen Fortschritt. Dieser Eindruck ist nicht unbegründet, wie Zahlen<sup>1</sup> belegen, nach denen sich ein relevanter Teil der KI-Nutzung offenbar mitbestimmungsfrei abspielt (Hoppe et al. 2023).

Wenn die notwendigen Beteiligungsmöglichkeiten gegeben sind, erleichtert dies eine menschenzentrierte, d.h. an den Bedürfnissen der Mitarbeitenden orientierte Gestaltung des betrieblichen KI-Einsatzes. Dieser Beitrag zeigt, wie die Beteiligung von Beschäftigten und ihren Interessenvertretungen an der digitalen Transformation zu einem wichtigen Erfolgs- und Akzeptanzfaktor für KI-Anwendungen werden kann. Partizipation ist dabei ebenso Ausdruck einer Demokratisierung im Betrieb wie Beitrag zur Humanisierung von Arbeit (Becker und Brinkmann 2023). Daran anknüpfend geht es um die Voraussetzungen, unter denen Betriebs- und Personalratsgremien in die Offensive gelangen und Beschäftigten-orientierte Strategien der betrieblichen KI-Nutzung entwickeln können.

### **Warum verspricht Beteiligung eine verbesserte betriebliche KI-Praxis, und wie lässt sie sich erreichen?**

Von der Beteiligung von Beschäftigten und ihren Interessenvertretungen wird zunächst erwartet, dass sie zu einer breiten Akzeptanz von KI führt, weil die Mitwirkung von operativ tätigen Beschäftigten zur anforderungs- und bedarfsgerechten KI-Gestaltung beiträgt (Hoppe und Hermes 2021). Dass Beteiligung eine notwendige Voraussetzung für den menschenzentrierten KI-Einsatz ist, stellt niemand in Abrede. Mit Blick auf die betriebliche Praxis drängt sich dennoch die Frage auf, ob dabei häufig „bloße ‚Mensch-im-Mittelpunkt‘-Rhetorik“ (Bahnmüller et al. 2023, S. 28) strapaziert oder ob Beschäftigtenbeteiligung tatsächlich ernst genommen wird. Es ist nötig, diese Frage aufzuwerfen, weil Beschäftigtenbeteiligung kein Selbstläufer ist, sondern es geeigneter betrieblicher Strukturen, passgenauer Methoden sowie der erforderlichen zeitlichen und personellen Kapazitäten bedarf. Beschäftigte müssen die Gelegenheit, sich am betrieblichen Transformationsgeschehen zu beteiligen, auch tatsächlich haben. Nicht zuletzt ist die Beschäftigtenbeteiligung auch eine Frage der Unternehmenskultur, weil sie einen geeigneten Rahmen der Einbettung von individuellen Kompetenzen und den Möglichkeiten zur Partizipation bildet (Schröder 2020). Dass das Beteiligungsprinzip umsetzbar ist, zeigt das Praxisbeispiel der Deutsche Telekom Service GmbH (im Folgenden DTS). Dort unterstützt etwa der digitale Assistent PIA<sup>2</sup> die Beschäftigten des Kundenservice beim Ausfüllen von Formularen, bei dem Zugriff auf unterschiedliche Datenbestände oder bei der

Informationssuche (Bayer und Bohn 2020). Beteiligung fördert bei PIA die interne Kommunikation über die KI-Technologie und verbessert auch den Wissenstransfer zwischen der für die IT-Entwicklung zuständigen Abteilung und den operativ tätigen Beschäftigten in den dezentralen Callcentern (Hoppe et al. 2022b).

Zur Beschäftigtenbeteiligung führen viele Wege. In der Praxis können dies klassische Beteiligungsinstrumente sein, die Unternehmen seit geraumer Zeit unabhängig von der digitalen Transformation nutzen. Dazu zählen etwa Methoden des Ideenmanagements, das betriebliche Vorschlagswesen, Bedarfsanalysen mittels Beschäftigtenbefragungen oder profan der Dialog zwischen Beschäftigten und ihren unmittelbaren Führungskräften. In der fortschreitenden, vom zunehmenden KI-Einsatz geprägten Digitalisierung werden auch neue oder an die Transformation angepasste Teilnehmungsformate wichtiger. Hierzu gehören u. a. die temporäre Einbindung operativ tätiger Beschäftigter in *agile* Entwicklungsprozesse oder die Beteiligung an *Design-Thinking-Workshops* zur Ideengenerierung für neue KI-Einsatzmöglichkeiten (Hoppe et al. 2022a). Beteiligung kann direkt durch die Mitwirkung der Mitarbeitenden an Innovationen erfolgen oder indirekt, vermittelt über betrieblich etablierte Teilnehmungsstrukturen. Das können betriebsinterne Informations- und Kommunikationskanäle sein, Methoden des Transfers von Wissen und Feedback (wie *KI-Postfächer*, *Multiplikator:innen-Netzwerk*) oder strukturierte Personalentwicklungsprogramme, die das Prozess- und Erfahrungswissen operativ tätiger Beschäftigter aufgreifen (wie *Fachkarrieren* für die IT-Entwicklung).

In der betrieblichen Praxis entsteht häufig der Eindruck, dass Beschäftigtenbeteiligung noch nicht den notwendigen Status hat. Die zu erreichenden Wirtschaftlichkeits- und Effizienzziele haben Priorität. Die positiven Effekte von Beteiligung etwa im Hinblick auf die Technologieakzeptanz oder die Qualität von KI-Anwendungen sind häufig nicht unmittelbar messbar und spielen deshalb betriebswirtschaftlich kaum eine Rolle. Nur wenn es gelingt, den Beitrag Teilnehmungs-orientierten Vorgehens bei der KI-Gestaltung deutlich zu machen, kann auch das von politischer Seite vielfach bemühte Leitbild der *menschenzentrierten KI-Gestaltung* mit Leben gefüllt werden. Dabei geht es um ethische Werte wie die Selbstbestimmtheit des Menschen, Würde und Fairness oder um technologische Anforderungen wie die Robustheit und Verlässlichkeit von KI-Systemen ebenso wie um die sozialpartnerschaftliche Interessenausgewogenheit im Betrieb. Diesem Ansatz folgend sind auch Betriebs- und Personalräte gut beraten, am Aufbau und an der nachhaltigen Nutzung von Methoden der Beschäftigtenbeteiligung mitzuarbeiten.

### Wie können Betriebs- und Personalräte zur partizipativen KI-Praxis beitragen, und welche Herausforderungen müssen sie dabei meistern?

Betriebliche Interessenvertretungen tun sich häufig noch schwer damit, in KI auch positive Ansätze für die Arbeit von Beschäftigten zu sehen und eigene Umsetzungsideen einer gelungenen KI-Praxis zu entwickeln. Betriebs- und Personal-

räte äußern einen hohen Unterstützungsbedarf hinsichtlich der Grundlagen der Einflussnahme und Mitgestaltung, problematisieren aber auch unzureichende rechtliche Möglichkeiten (Hoppe 2022). Für viele ist es schwierig, frühzeitig an Informationen zum Transformationsgeschehen zu kommen und anstehende Veränderungen *auf Augenhöhe* mitzugestalten. Bislang sind für viele Interessenvertretungen die Fragen unbeantwortet, mit welchen Instrumenten und Verfahren Beteiligung und Zugriff möglich ist, welche Auswirkungen KI auf Beschäftigte hat (beispielsweise KI-Folgenabschätzung), wie sie zum Schutz der Persönlichkeitsrechte der Mitarbeitenden beitragen können und welche Ansatzpunkte zur KI-Regulierung geeignet erscheinen (Hoppe et al. 2023).

Der derzeitige Rechtsrahmen gibt Betriebs- und Personalräten aber durchaus Möglichkeiten, das KI-Geschehen im Sinne einer Teilnehmungs-orientierten Praxis mitzugestalten. KI-Systeme haben häufig die gleichen betriebspolitischen Auswirkungen wie der Einsatz konventioneller (IT-) Technologie, die betrieblichen Interessenvertretungen müssen sich also keineswegs auf eine defensiv-abwartende Position zurückziehen. Handlungsfelder, die zum Kerngeschäft der Beschäftigtenvertretung gehören, sind etwa Entgelt-, Arbeitszeit-, Qualifizierungs- und Personalentwicklungs- oder auch Arbeitsorganisationsfragen, alles hinreichende Gründe, das Thema KI offensiv und selbstbewusst anzugehen. Bezüge bestehen etwa durch die Mitbestimmung des Betriebsrats bei der Einführung technischer Systeme zur Verhaltens- und Leistungskontrolle (§ 87 Abs. 1 Nr. 6 BetrVG), die Beteiligung bei Maßnahmen der Personalplanung (§ 92 BetrVG), in Fragen der Beschäftigungssicherung (§ 92a BetrVG) oder bei betrieblichen Bildungsmaßnahmen (§§ 96 und 98 BetrVG). Zudem hat der Gesetzgeber 2021 für Betriebsräte den Zugang zu Expert:innenwissen bei der Anwendung von KI im Betrieb durch das Betriebsrätemodernisierungs-Gesetz erleichtert: Muss der Betriebsrat die Einführung oder Anwendung von Künstlicher Intelligenz beurteilen, so hat er Anspruch auf die Hinzuziehung eines Sachverständigen (§ 80 Abs. 3 BetrVG).

Gleichwohl gibt es bei KI eine Reihe von Herausforderungen: KI ist – anders als herkömmliche IT – dadurch gekennzeichnet, dass aus Daten eigenständig Schlussfolgerungen gezogen und somit Ergebnisse hervorgebracht werden können, die nicht vorherbestimmt sind. Diese „Zone des Ungewissen“ in der *Blackbox KI* (Martini 2019) löst mitunter ein gewisses Unbehagen aus und prägt auch die Einstellung von Betriebs- und Personalräten. KI-Mitbestimmung ist fast immer auch Handeln unter Unsicherheit. Bei konventionellen IT-Systemen lässt sich Mitbestimmung vor dem Technologieeinsatz ausüben, etwa indem geprüft wird, ob es Handlungsbedarf gibt, weil sie sich zur Leistungs- und Verhaltenskontrolle *eignen*. Bei KI ist die Situation schwieriger, weil die Anwendungen definitionsgemäß lern- und selbstoptimierungsfähig sind. Die Auswirkungen von KI auf Beschäftigte können sich deshalb im laufenden Einsatz verändern. Für die betriebliche Mitbestimmung bedeutet dies, dass bei KI-Systemen ein einmaliger Mitbestimmungsakt vor dem Praxisbetrieb zu kurz greift und es notwendig wird, die Folgen für Beschäftigte fortlaufend im Prozess zu überprüfen. Dieser Ansatz wird unter dem Stichwort der prozessualen oder prozessorientierten Mitbestimmung diskutiert (Gerst 2020). Betriebs- und Personalratsgre-

mien zeigen sich damit jedoch häufig überfordert. Die Gründe liegen neben dem benötigten Sachverstand, der vielfach mangelnden unzureichenden Informationslage und der zunehmenden Komplexität der betrieblichen IT- und KI-Landschaften insgesamt auch in den zeitlichen und personellen Ressourcen für die Bearbeitung von KI-Themen.

## Wo ist der Ausgang aus dem Mitbestimmungslabyrinth bei KI-Systemen?

Auch wenn das Betriebsrätemodernisierungs-Gesetz die Situation etwas verbessert hat, sind noch weitere Schritte notwendig, die Betriebs- und Personalräte im Interesse einer menschenzentrierten KI-Nutzung gehen sollten. In unserer Arbeit im Projekt humAI work lab haben wir fünf Anforderungen identifiziert, die eine gelingende KI-Mitbestimmungspraxis fördern (Hoppe et al. 2022a).

Die *erste* Anforderung besteht darin, *innenpolitisch* aktiv zu werden und die betriebliche Sozialpartnerschaft neu auszurichten mit dem Ziel, den Mitbestimmungsprozess zu verschlanken. In unserem Praxisbeispiel DTS haben sich die betrieblichen Sozialparteien zunächst auf gemeinsame Mitbestimmungsleitlinien verständigt, die eine frühe Einbindung der Interessenvertretung ebenso enthalten wie deren Bereitschaft zur proaktiven Beteiligung an der KI-Transformation. Das bei der DTS gewählte Vorgehen entspricht dem Leitbild einer Agilisierung der sozialpartnerschaftlichen Zusammenarbeit im Betrieb, wie sie verschiedentlich für notwendig gehalten und gefordert wird (Holzapfel 2021).

Die *zweite* Anforderung zielt darauf ab, Mitbestimmung entlang des Prozesses der KI-Nutzung zu organisieren. Das ist deshalb erforderlich, weil KI-Systeme eine stetige Veränderung erwarten lassen. Dem muss die betriebliche Mitbestimmung Rechnung tragen. Die prozessorientierte Mitbestimmung gilt bei der DTS als *Win-Win*-Situation. Das Unternehmen profitiert in Form einer beschleunigten und vereinfachten Mitbestimmung, der Betriebsrat durch frühzeitige Information und Beteiligung.

Als *dritte* Anforderung sind geeignete Regulierungsformen zu schaffen. Bei der DTS wird die prozessorientierte Mitbestimmung über eine konzernweite IT-Rahmenvereinbarung um-

gesetzt, konkretisiert durch eine Gesamtbetriebsvereinbarung Digitalisierung. In dieser ist ein *Digiboard* als gemeinsames Gremium zwischen Arbeitgeber und Gesamtbetriebsrat vereinbart, das den Austausch und die Einbindung der Interessenvertretung sichert. Zudem gibt es eine halbjährlich zu erstellende *Roadmap*, auf der das Unternehmen Auskunft über anstehende Entwicklungen gibt. Zusätzlich muss der Arbeitgeber bei der Einführung neuer IT und KI in einem *Steckbrief* über die Funktionsweise, die Nutzungsmöglichkeiten und möglichen Folgen für Beschäftigte informieren. Ähnlich wird auch in anderen Vorreiterunternehmen verfahren. So verfügt Siemens über das Instrument der *AI Card* zur Einschätzung des Risikopotenzials von KI (Grasy und Seibold 2023). Als Orientierung für betriebliche Praktiker:innen wäre ein Gesamtüberblick über die angewandten Regelungen hilfreich, der jedoch bislang nicht existiert.

Die *vierte* Anforderung fokussiert darauf, die angesprochenen Wissensdefizite auf Betriebsratsseite zu beheben und KI-Expertise aufzubauen. Bei der DTS gelingt das durch Spezialisierung, etwa durch die Mitarbeit im IT-Ausschuss des Gesamtbetriebsrats. Neben der Teilnahme an einschlägigen Qualifizierungsmodulen wird IT- und KI-Wissen auch durch die Mitwirkung an Erprobungsverfahren (*friendly user tests*) vermittelt. Zudem sollen Personen mit einem beruflichen IT-Hintergrund für die Mitarbeit im Gremium gewonnen werden. Im Allgemeinen hilft beim Kompetenzerwerb auch die Orientierung an den vielfältig verfügbaren Handlungshilfen und Ratgebern (beispielsweise ver.di Bildungswerk Hessen 2022).

Die *fünfte* Anforderung verankert Beteiligung als KI-Gestaltungsprinzip. Bei der Telekom ist die Beschäftigten- und Betriebsratsbeteiligung an der digitalen Transformation auch Ausdruck der digitalen Ethik, die nicht zuletzt seit Ende 2022 in einem sozialpartnerschaftlichen „KI-Manifest“ (Deutsche Telekom AG 2022) verbrieft ist. Kombiniert mit der prozessorientierten KI-Mitbestimmungspraxis und ergänzenden Vorschriften zur Gewährleistung von IT-Sicherheit bieten die darin formulierten Leitlinien die Grundlage dafür, dass sich die Entwicklung und Einführung neuer KI-Anwendungen beschleunigen lässt und gleichzeitig die Bedürfnisse von Beschäftigten im Hinblick auf notwendige Qualifizierungen, Entlastung, Datenschutz und zukünftige Beschäftigungsfähigkeit berücksichtigt werden.



**Markus Hoppe**

**Markus Hoppe** ist seit 2017 als wissenschaftlicher Mitarbeiter bei der in Stuttgart ansässigen INPUT Consulting gGmbH angestellt, einer Forschungs- und Beratungsgesellschaft, die sich mit der Entwicklung der Arbeitswelt und des Dienstleistungssektors beschäftigt. In seiner Arbeit befasst er sich schwerpunktmäßig mit Fragen der digitalen und zunehmend auch durch KI geprägten Transformation von Unternehmen und Beschäftigung. Im Rahmen mehrerer Forschungsprojekte hat er analysiert, wie sich der Einsatz von KI auf die Beschäftigungssituation der Arbeitnehmer:innen auswirkt und welche Herausforderungen betriebliche Interessenvertretungen dabei meistern müssen.

## Fazit

Beschäftigtenbeteiligung an der betrieblichen KI-Praxis ist kein Selbstzweck. Dem menschenzentrierten Leitbild folgend sind Verfahrensweisen sinnvoll, die systematisch auf den Anforderungen von Beschäftigten aufbauen. Dazu bedarf es geeigneter Rahmenbedingungen, die es Beschäftigten ermöglichen, sich in technologische Innovationsprozesse einzubringen. Das erfordert Zeit, die an anderer Stelle möglicherweise fehlt und vom Unternehmen eingeräumt werden muss, etwa für die Beteiligung an agilen Entwicklungsprojekten oder für Qualifizierungszwecke. Solche Investitionen lohnen sich aber: Sie begünstigen nicht nur einen nachhaltigen KI-Einsatz, sie können sich vielmehr auch in unternehmerischer Perspektive auszahlen. Die Stichworte lauten: technologische Passgenauigkeit, Fehlerfreiheit und nicht zuletzt geringere Beschaffungs- und Unterhaltungskosten. Hinzu kommt, dass KI-Technologien, die beteiligungsorientiert entwickelt und erprobt werden, breiter akzeptiert sind, weil Beschäftigte sie zu *ihrer Sache machen* können. Dies fördert auch die Motivation, die Technologie sinnvoll zu nutzen, weil es Anerkennung und Wertschätzung zeigt. Beteiligung kann angesichts des in vielen Branchen bestehenden Personal- und Fachkräftemangels außerdem ein Element der Arbeitgeber-Attraktivität bei der Gewinnung und Bindung motivierter, qualifizierter und zufriedener Beschäftigter sein und so – im Interesse von Unternehmen und Beschäftigten gleichermaßen – bei der Bewältigung des digitalen Wandels helfen.

## Referenzen

- Bahn Müller, Reinhard; Kutlu, Yalçın; Mugler, Walter; Salm, Rainer; Seibold, Bettina; Kirner, Eva; Klatt, Sandra (2023): Mitsprache bei der Digitalisierung? Beteiligung von Betriebsrat und Beschäftigten in digitalisierungsaktiven Betrieben. Düsseldorf.
- Bayer, Ferdinand; Bohn, Leonie (2020): Fallstudie Telekom Service GmbH. Was ist und kann PIA? In: IBM Deutschland GmbH und ver.di (Hg.): Künstliche Intelligenz. Ein sozialpartnerschaftliches Forschungsprojekt untersucht die neue Arbeitswelt. Ehningen, Berlin, S. 62–64.
- Becker, Karina; Brinkmann, Ulrich (2023): Partizipation. In: Rainer Bohn, Hartmut Hirsch-Kreinsen, Sabine Pfeiffer und Mascha Will-Zocholl (Hg.): Lexikon der Arbeits- und Industriezoologie. 3. Auflage. Baden-Baden: Nomos, S. 296–300.
- Deutsche Telekom AG (2022): Manifest zwischen dem Konzern Deutsche Telekom und dem Konzernbetriebstat über die Einführung und Nutzung lernender informationstechnischer Systeme (sog. Künstliche Intelligenz). (KI Manifest). Bonn.
- Gerst, Detlef (2020): Mitbestimmung in digitalen und agilen Betrieben – das Modell einer prozessualen partnerschaftlichen Konfliktkultur. In: Verena Bader und Stephan Kaiser (Hg.): Arbeit in der Data Society. Wiesbaden: Springer, S. 35–56.
- Giering, Oliver (2022): Künstliche Intelligenz und Arbeit: Betrachtungen zwischen Prognose und betrieblicher Realität. In: Z. Arb. Wiss. 76 (1), S. 50–64.
- Grasy, Jonas; Seibold, Bettina (2023): Die Komplexität von KI-Systemen durch Steckbriefe bewältigen – Portrait über den Einsatz sogenannter AI-Cards bei der Siemens AG. Hg. v. I.M.U. Institut für Mitbestimmung und Unternehmensführung. Hans Böckler Stiftung. Düsseldorf.
- Holzappel, Marco (2021): Agilität der betrieblichen Zusammenarbeit – Vom klassischen Verhandlungsprozess zum agilen Mitbestimmungs-Scrum. Köln. Online verfügbar unter <https://marco-holzappel.de/2021/03/06/agilitaet-der-betrieblichen-zusammenarbeit-vom-klassischen-verhandlungsprozess-zum-agilen-mitbestimmungs-scrum/> (zuletzt geprüft am 13.7.2023)
- Hoppe, Markus (2022): Über die Einstellung von Mitbestimmungsakteuren zu KI-Systemen: Ergebnisse einer Onlinebefragung von Betriebs- und Personalräten zu Vertrauensfaktoren und Regulierungserfordernissen. In: Lothar Schröder und Petra Höfers (Hg.): Praxishandbuch künstliche Intelligenz. Frankfurt am Main: Bund-Verlag, S. 351–377.
- Hoppe, Markus; Hermes, Adrian (2021): Beschäftigteninteressen und Regulierungserfordernisse bei KI-Anwendungen. Stuttgart: Fraunhofer Verlag.
- Hoppe, Markus; Klueß, Daniel; Roth, Ines; Suriano, Giovanni (2022a): Praxis: Künstliche Intelligenz im Kundenservice. In: Gute Arbeit (4), S. 14–18.
- Hoppe, Markus; Roth, Ines; Suriano, Giovanni; Klueß, Daniel (2022b): Künstliche Intelligenz menschenzentriert gestalten. Ein Praxisbericht aus dem Kundenservice (humAI in work lab KI-Lernraum). Online verfügbar unter <https://www.humain-worklab.de/wp-content/uploads/2022/04/KI-menschenzentriert-gestalten.-Ein-Praxisbericht-aus-dem-Kundenservice.pdf> (zuletzt geprüft am 13.7.2023)
- Hoppe, Markus; Roth, Ines; Wedde, Peter (2023): Online-Umfrage: KI-Einsatz und Mitbestimmung. In: Computer und Arbeit (CuA) (5), S. 36–38.
- Hoppe, Markus; Roth, Ines; Wedde, Peter (2022c): Praxis: Online-Umfrage zu KI in der Mitbestimmung. In: Computer und Arbeit (CuA) (12), S. 26–37.
- Martini, Mario (2019): Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Schröder, Laura (2020): Digitalisierung gelingt nur mit den Beschäftigten. In: Heilberufe 72 (6), S. 56–58.
- Sevindik, Ugur (2022): Verbreitung und Einsatz von Künstlicher Intelligenz in Deutschland – Auswirkungen auf berufliche Anforderungen und Strukturen. Hg. v. Bundesinstitut für Berufsbildung (BiBB). Bonn. Online verfügbar unter [https://datapool-bibb.bibb.de/pdfs/Sevindik\\_Kuenstliche\\_Intelligenz.pdf](https://datapool-bibb.bibb.de/pdfs/Sevindik_Kuenstliche_Intelligenz.pdf) (zuletzt geprüft am 13.7.2023)
- ver.di Bildungswerk Hessen (2022): Kompetenzen über KI aufbauen. Handlungsleitfaden für Betriebsräte. Hg. v. Bildungswerk ver.di im Lande Hessen e.V. Frankfurt am Main. Online verfügbar unter [https://www.btq-kassel.de/wp-content/uploads/2023/06/2022-06-12-Handlungsleitfaden-KI\\_2022.pdf](https://www.btq-kassel.de/wp-content/uploads/2023/06/2022-06-12-Handlungsleitfaden-KI_2022.pdf) (zuletzt geprüft am 13.7.2023)

## Anmerkungen

- 1 Die Daten entstammen einer Onlinebefragung von Betriebs- und Personalräten zu den (mitbestimmungs-)rechtlichen Herausforderungen des betrieblichen KI-Einsatzes, die Ende 2022 von der INPUT Consulting gGmbH durchgeführt und zu deren Teilnahme in der Zeitschrift „Computer und Arbeit“ geworben wurde (Hoppe et al. 2022c). An der Befragung beteiligten sich insgesamt 260 Personen.
- 2 Die Abkürzung PIA steht für Persönlicher Interaktiver Assistent. PIA ist eine Robotic Desktop Automation (RDA), die Kundenberater:innen im Callcenter unterstützt, indem sie Routineaufgaben übernimmt.

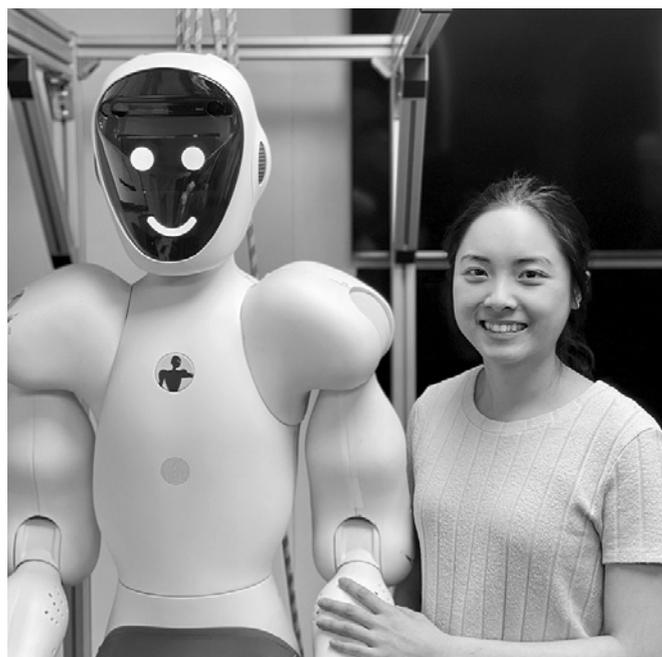


## Gute Digitalisierung von Interaktionsarbeit?

Chatbots im Unternehmenskontext betreffen Kund:innen wie Mitarbeiter:innen. Letztere sind von der Frage nach der human friendly automation besonders betroffen<sup>1</sup>: Kommen die Kund:innen mit der automatisierten Kund:innenschnittstelle klar? Falls nicht, kommt es zu eher unerfreulichen intensiven Kund:innenkontakten und Mehrarbeit. Ist die Technik nach den eigenen Organisations- und Tätigkeitsanforderungen gestaltet? Davon hängt schließlich ab, inwieweit aktuelle Digitalisierungsprozesse gute Arbeit unterstützen oder das Gegenteil bewirken. Beim Stichwort gute Arbeit richtet sich der Blick insbesondere auf Interaktionsarbeit als Arbeit an und mit Menschen. Sie definiert sich nach einem in der Arbeitsforschung bewährten Konzept durch die voraussetzungsvolle Herstellung einer Kooperationsbeziehung zwischen Kund:innen und Dienstleister:innen, was nicht selbstverständlich ist, denn in der Regel haben sie unterschiedliche Interessen. Zudem kann in der Kundenbeziehung die Arbeit an den eigenen Emotionen notwendig werden, genauso wie die Arbeit an den Gefühlen des Gegenübers (zum Beispiel, um beruhigend einzuwirken). Interaktionssituationen sind kaum planbar, ihre Bewältigung zeichnet sich durch ein schrittweises Vorgehen, eine Wahrnehmung mit allen Sinnen, ganzheitliches und assoziatives Denken und eine empathische Beziehung zum Arbeitsgegenstand aus (vgl. Böhle/Wehrich 2020; Böhle 2017). Gründe und Herausforderung genug, um sich genauer anzusehen, wie gelungene Technikgestaltung aus Sicht der Mitarbeiter:innen zur Unterstützung ihrer Interaktionsarbeit aussehen sollte. Dazu wird in dem Artikel beispielhaft ein Partizipations- und Reflexionsprozess im Rahmen eines Chatbot-Entwicklungsprozesses vorgestellt, bei dem sich insbesondere zeigt<sup>2</sup>: Wie gut die Beschäftigten am Arbeitsplatz entlastet werden, hängt auch bei der Entwicklung eines inter(aktions)arbeitsensiblen Chatbots davon ab, ob die Perspektive der Kund:innen antizipiert wird. Qualitätsmerkmale und Gestaltungsempfehlungen für KI gilt es daraufhin zu vertiefen.

### KI-gestützte Servicearbeit

Chatbots erfreuen sich auch in Unternehmen mit Kontakt zu Kund:innen immer größerer Beliebtheit (EOS-Chatbot-Studie 2021). Dementsprechend werden Mitarbeiter:innen an Serviceschnittstellen und Kund:innen immer häufiger mit Chatbots konfrontiert. Vollautomatisierte Varianten finden sich hier bislang noch selten (in der oben genannten Studie bei jedem zehnten befragten Unternehmen). Weitaus verbreiteter ist, dass Chatbots den Erstkontakt sowie die thematische Vorauswahl bei Kund:innenanfragen übernehmen<sup>3</sup>. Das gilt umso mehr im Mittelstand, dem zur Entwicklung KI-gestützter Servicearbeit in der Regel nur begrenzte Mittel zur Verfügung stehen. Gleichzeitig sind Automatisierungsansprüche an der Kund:innenschnittstelle hier besonders empfindlich, zählt doch der persönliche Kontakt hier ganz besonders.



Halodi Robotics' perception engineer with a humanoid collaborative robot, Foto: Nicholas-halodi, CC BY-SA 4.0

Generell wird der Mehrwert von Chatbots gerne mit der Entlastung von Mitarbeiter:innen erklärt, die sich dann intensiver qualitativ anspruchsvolleren Interaktions(arbeits)themen widmen können sollten (Böhmman et al. 2023; Gentsch 2019)<sup>4</sup>. Zur Substitution von Interaktionsarbeit durch Technik gibt es jedoch auch zahlreiche kritische Gegenstimmen (vgl. verschiedene Beiträge in Ciesinger et al. 2022). Untersuchungen zeigen, dass auch ein interaktionssensibler Chatbot schwerlich die gleichen Fähigkeiten wie Servicemitarbeiter:innen ausweist – zumindest was seine Fähigkeiten zur Interaktionsarbeit betrifft (ausführlich dargestellt in Porschen-Hueck/Zylowski 2021). Umso wichtiger ist ein gemeinsamer Partizipations- und Reflexionsprozess für die *menschengerechte Gestaltung* der Mensch-Maschine-Interaktion.

### Fallbeispiel zur „menschengerechten“ Gestaltung eines Chatbots

Im Folgenden wird ausschnitthaft ein konkreter Entwicklungsprozess betrachtet, der den Anspruch und die Praxis *guter Digitalisierung* verfolgt hat. Mit Hilfe des Reflexionsinstruments *Humanizing Men-Machine-Interaction with AI (HAI-MMI)* (Huchler 2020) wurde für den Arbeitskontext diskutiert, welchen Ansprüchen ein Chatbot im Unternehmen mindestens genügen sollte<sup>5</sup>. Zugleich wurde das Reflexionsinstrument für den Einsatz von KI auf seine Tauglichkeit für den Einsatz in der Chatbot-Entwicklung getestet.

In unserem Fallbeispiel eines mittelständischen IT-Unternehmens für die Herstellung und den Vertrieb von Business-Software besteht die Herausforderung darin, im Rahmen einer Geschäftsmodell-Transformation hin zu vermehrten B2C-Schnittstellen die Dienstleistungsqualität gegenüber Kund:innen bei knapper Personalkapazität hochzuhalten. Wenn das Unternehmen sein Geschäftsmodell so umstellt, dass Endkunden nicht mehr über ein Partnernetzwerk, sondern auf einer Plattform betreut werden, wird es sich vermehrt um die Endkunden kümmern müssen. Das Unternehmen legt großen Wert auf ein hohes Service-Niveau,

das bislang durch personenbezogene Dienstleistungen erreicht wurde, für die Interaktionsarbeit der Mitarbeiter:innen einen besonderen Stellenwert hat (siehe oben). Nun soll ein Chatbot an der Kund:innenschnittstelle die Beschäftigten entlasten und bei der Servicetätigkeit unterstützen, sei es mit Blick auf Produktinformationen, Konnektivitätsthemen oder Fehleranfragen etc. Das ist allerdings *nicht ohne*, denn Servicemitarbeiter:innen bringen für Interaktionsarbeit zahlreiche Fähigkeiten mit, die die Technik erst noch lernen muss, was gar nicht 1:1 möglich ist. Deshalb gehört auch die Kenntnis der Grenzen von Technik mit Blick auf Interaktionsarbeit dazu. Um Kund:innen des IT-Unternehmens weiterhin zu binden, muss der Chatbot hohen Ansprüchen genügen<sup>6</sup>. So liegt es nahe, mit Kund:innenanliegen befasste Mitarbeiter:innen in den Entwicklungsprozess einzubeziehen und Chancen und Fallstricke im Entwicklungsverlauf zu reflektieren. In unserem Fallbeispiel wurde dazu ein partizipativer Entwicklungsprozess angestoßen, der die Anliegen der betroffenen Mitarbeiter:innen ins Zentrum gestellt hat. Da der Chatbot an der Kund:innenschnittstelle eingesetzt wird, wurden bei der Entwicklung außerdem Kriterien für die menschengerechte Gestaltung auch mit Blick auf Kund:innen diskutiert. Das adressiert die in der Software-Entwicklung vertraute Frage – wer sind die Kunden oder Nutzer, auf die sich die Technikentwicklung letztendlich auswirkt? Bekanntlich gibt es darauf unterschiedlich ernsthafte und passende Antworten. Die Diskussion der Kriterien fand begleitend zum Entwicklungsprozess statt und nicht erst im Nachhinein, womit sie in der Umsetzung berücksichtigt werden konnten.

### Partizipativer Entwicklungsprozess

Wie wurde in diesem partizipativen Entwicklungsprozess vorgegangen? In unserem Fallbeispiel wurden dazu Mitarbeiter:innen an Kund:innenschnittstellen zu ihrem Arbeitshandeln und den Quellen von Unterbrechungen interviewt. Eine Schlüsselfrage dabei lautete, inwieweit und inwiefern technisch-organisatorische Lösungen wie die Einführung eines Chatbot-Prozesses dabei entlasten können. Anschließend wurden auf Basis der Befragungsergebnisse in Workshops mit Domänenexpert:innen die gewünschten Fähigkeiten des Chatbots herausgearbeitet:

In den Interviews und den Feedback-Workshops war für die Mitarbeiter:innen das Stichwort *Customer Experience* zentral. Die beginnt mit einem guten Einstieg der Kund:innen in den Chatbot-Prozess. Als wichtig wurde erachtet, ob der Chatbot *die Sprache* der Kund:innen *verstehen* und dafür im besten Fall eine *Kompetenzeinschätzung/-einordnung* der Kund:innen bewerkstelligen kann. Es geht um die Einschätzung, ob man es mit Expert:innen oder Laien in Bezug auf das Produkt zu tun hat, womit unterschiedlich voraussetzungsreiche Antworten möglich sind. Die zentrale Anforderung an den Chatbot liegt des Weiteren in der guten Auffindbarkeit von Informationen für die Kund:innen. Das gilt besonders für Standardfragen, die Kund:innen zunehmend am liebsten 24/7 stellen können möchten. Eine gute *Customer Experience* geht aber auch damit einher, dass die Prozesskette transparent visualisiert wird. Je besser nachvollziehbar die Möglichkeiten für Feedback sind, je vertrauenswürdiger die Datensouveränität der Anfragenden gewahrt und vermittelt wird sowie tatsächlich befriedigende Lösungen zu finden sind, desto mehr werden Mitarbeiter:innen

auch tatsächlich von Basisanfragen entlastet und gewinnen Zeit für individuelle Kund:innenberatungen. Die individuellen Kund:innenberatungen werden weiterhin insbesondere unter dem Stichwort *Second Level Support* als zentral und obligatorisch bewertet. Im *First Level Support* soll der Chatbot dabei helfen zu bewerten, wie dringend die Anfragen sind, sie den fachlich passenden Ansprechpartner:innen zuweisen und sie bei der dazu notwendigen Terminkoordination und Vermittlung unterstützen. In abstraktere Fähigkeitskategorien übersetzt heißt das: (1) FAQ & Weiterleitung, (2) direkte Antworten, (3) Rücksprache mit einem Mitarbeiter/einer Mitarbeiterin, (4) Terminvereinbarung, (5) Produktberatung und (6) Smalltalk. Diese Kategorien dienen als Grundlage für die Gestaltung von Anwendungsfällen und deren technische Umsetzung. Aus Sicht des Unternehmens können dazu bereits viele Daten aus den bestehenden und ausgewerteten FAQs gesammelt und exemplarische Kundenanfragen und Antworten zusammengefasst werden. Eingebunden werden sollen auch Whitepaper, Schulungs- und Informationsvideos etc. Durch die Verknüpfungslogik des Systems in dem Unternehmen wird die Informationsabfrage, Einordnung der Kund:innenanfrage und darüber hinaus die Zuordnung zu fachlich kompetenten Mitarbeiter:innen (Übergabe) technisch abbildbar.

### Reflexionsprozess zur menschengerechten KI-Gestaltung

Der oben beschriebene Partizipationsprozess wurde durch einen Reflexionsprozess im Entwickler:innenteam ergänzt, der auf die Gestaltungskriterien der *Humanizing AI Men-Machine Interaction* (HAI-MMI) (Huchler 2020) zurückgriff<sup>7</sup>. Das Konzept ist für unseren Fall deshalb interessant, weil es auf die Verschiebung im Zusammenspiel zwischen technischen Prozessen und menschlichem Arbeitshandeln durch den Einsatz von KI eingeht (ebd.) und berücksichtigt, dass KI-Technologie systematisch auf menschliche Arbeit (Interaktionsarbeit) angewiesen bleibt. Es geht um eine sozial nachhaltige *Arbeitsteilung* zwischen Mensch und Maschine für die nicht nur ethische Gesichtspunkte, sondern insbesondere auch die systematisch erschlossenen jeweiligen Kompetenzen berücksichtigt werden, die es weiterzuentwickeln gilt. Im Mittelpunkt soll eine *ausgewogene, qualitativ hochwertige Interaktion* zwischen Mensch und KI stehen und damit weder eine alleinige Dominanz menschlichen Arbeitshandeln noch eine alleinige Dominanz einer vollautomatisierten KI-Steuerung.

Um dem näher zu kommen, wurden in unserem Entwicklungs- und Gestaltungsprozess die im HAI-MMI-Konzept herausgestellten Qualitätsebenen der Mensch-KI-Interaktion in der Diskussion vertieft. Dazu zählen *Koordination, Arbeitsteilung, Lernen, Adaptivität* und *Empowerment* (detailliert beschrieben siehe Huchler 2020 a. a. O.). Werfen wir exemplarisch einen Blick auf die Qualitätsebene *Koordination*<sup>8</sup>, die auf unterschiedliche Koordinationsniveaus der Mensch-Maschinen-Interaktion eingeht: In den Reflexionsrunden mit dem Entwicklerteam und betroffenen Mitarbeiter:innen sowie Testnutzer:innen wurde eine vorgesehene Funktionalität des Chatbots mit Blick auf die Koordinationsqualität kritisch diskutiert. Zwar wurde zugestimmt, dass die angebotene Terminkoordination des Chatbots mit dem Kunden für ein persönliches Servicegespräch auf Ba-

sis der Mitarbeiter:innen-Kalender eine Entlastung darstellt, da diese Koordinationsarbeit nicht mehr selbst durchgeführt werden muss. Zugleich wurde aber mit Blick darauf, dass die Terminkoordination fremdgesteuert ohne große eigene Eingriffsmöglichkeiten erfolgt, kritisch angemerkt, dass damit lediglich eine kompensierende Koordinationsmöglichkeit beim Menschen verbleibt. Besser wäre eine interaktive Gestaltung der Mensch-Maschinen-Interaktion, die eine proaktive Koordinationsbeziehung befördern kann. Hier wird die Frage, wer steuert und entscheidet, transparent und interaktiv im Prozess abgestimmt. Das System bietet aktiv Aufgaben an, die es übernehmen kann (Huchler 2020).

Positiv wurde bewertet, dass der Chatbot den Chatverlauf als Kontextinformation zur Unterstützung der Berater:innen übermittelt. Nach Feedback-Runden mit Mitarbeiter:innen stellte das Entwicklerteam hier aber ebenfalls Ausbaupotenzial fest: So sollte der Chatbot auf den Chatverläufen basierende Beratungsvorschläge einbringen, um mit seinem Dienst über eine einseitige Koordination hinauszugehen. Von Seiten der Mitarbeiter:innen könnten solche Vorschläge aufgegriffen und auf Basis ihrer Expertise weiterentwickelt bzw. abgewogen in das Beratungsgespräch eingebracht werden. Die Verschränkung der Ergebnisse der Datenanalyse mit Expert:innenwissen der Beschäftigten ist ein gutes Beispiel für die angestrebte ausgewogene und qualitativ hochwertige Mensch-Maschinen-Interaktion. Zu der gehört über die Mitarbeiter:innensicht hinaus auch die Berücksichtigung der Kund:innenperspektive, die sich im Entwicklungsprozess immer weiter in den Vordergrund geschoben hat. Im Zuge dessen wurden verschiedene Interaktionsstufen bis hin zu Grenzen der Mensch-Maschine-Interaktion und Übergabe-Szenarien mit ihrer Bedeutung für die Kund:innen herausgearbeitet. In der ersten Stufe bietet der Chatbot eine einseitige Informationsabfrage im Sinne von FAQs an (einseitige Koordination). In der nächsten Stufe gibt der Chatbot Antworten aufgrund von Schlüsselfragen/-wörtern (einfache Koordination). Eine weitere Stufe besteht im Erkennen eines umfassenderen Wissensbedarfs und der Generierung von Angeboten für Schulungen, Webinare, Schulungsvideos etc. aus dem Fundus, der zu der Problemanfrage abrufbar ist (interaktive Koordination). Nach dem HAI-MMI-Konzept ist mit der Beförderung einer proaktiven Koordinationsbeziehung ein qualitativ hochwertiges Niveau erreicht (siehe Endnote 8).

An der Kund:innenschnittstelle ist zudem wesentlich, die Grenzen der automatisierten Antwortmöglichkeiten im Blick zu ha-

ben und Kund:innen in diesen Fällen direkt an kompetente Ansprechpartner:innen zu vermitteln. Technisch wird hier zwischen einem *sprachlichen Nicht-Verstehen (Natural Language Understanding)* und dem Nicht-Wissen, wie der Dialog weitergeführt werden soll, unterschieden. Auch in dem Fall, wenn Kund:innen an ungeeigneter Schnittstelle ihre Anliegen äußern, ist eine aktive Vermittlung an zuständige Einheiten im Unternehmen (*Routing via Intent Classification*) durch den Chatbot erforderlich. Darüber hinaus ist die Herstellung einer Mensch-Mensch-Kommunikation an definierten Stellen (Übergabe) im Fehlerfall angeraten, also schlicht, wenn der Chatbot nicht helfen kann. Damit Kund:innen dabei nicht aus dem Dialog herausfallen und erneut Kontakt herstellen müssen, empfiehlt sich die oben angesprochene automatische Terminkoordination, beispielsweise per E-Mail an Mitarbeiter:innen mit Link zu dem entsprechenden Chat oder ein Termineintrag. Das Beispiel der automatisierten Terminkoordination zeigt, dass auch für *gute Digitalisierung* von Interaktionsarbeit Abwägungen getroffen werden müssen. In dem Fall ist zwischen der Entlastung von Koordinationsarbeit und dem fremdbestimmten Terminkalender abzuwägen.

### Die Perspektive wechseln

Über die Reflexion der Qualitätsebenen hinaus wurde auch bei der Diskussion der Gestaltungskriterien für eine menschengerechte Mensch-Maschinen-Interaktion (aus Huchler et al. 2020) *Schutz des Individuums, Vertrauenswürdigkeit, förderliche Arbeitsbedingungen und sinnvolle Arbeitsteilung* (Huchler 2020) offensichtlich, dass es nicht ausreicht, nur auf den eigenen Arbeitskontext zu schauen. Am Beispiel *digitales Vertrauen* wird der notwendige Perspektivenwechsel besonders deutlich: Ist transparent, dass ein Chatbot eingesetzt wird? Kund:innen müssen sich bewusst sein können, dass sie mit einem Chatbot kommunizieren (siehe auch den EU-Gesetzesentwurf zur KI), wenngleich es Hinweise auf einen Vertrauensbonus für vermenschlichte Chatbots gibt (siehe hierzu kritisch Hardi 2019). Vertrauen in einen Chatbot-Prozess geht weiter mit der Frage, wann ein Chatbot erscheinen sollte, um nicht aufdringlich zu wirken. Das *Onboarding* der Kunden bedarf einer zielgruppenspezifischen Aufbereitung, um einzuladen und nicht abzuschrecken. Die Gestaltungsfrage betrifft auch das Marken-Image bzw. die Vermittlung der Unternehmenskultur. Stellen die Funktionalitäten keine angemessene Alternative für persönlichen Austausch sicher, besteht die Gefahr, dass die Nutzer:innen den Chatbot schnell verlassen.



### Stephanie Porschen-Hueck

**Stephanie Porschen-Hueck** ist langjährige Wissenschaftlerin am Institut für Sozialwissenschaftliche Forschung ISF München e.V. Sie forscht zu Veränderungen und zukunftsfähiger Gestaltung von Arbeit im Zusammenspiel von Mensch, Organisation und Technik.

Präventiv kann hier ein kontinuierlicher Verbesserungsprozess – z. B. mit Blick auf die verlässliche Sachauskunft und hohe Datenqualität (aktuelle Produktdaten und Angebote/Empfehlungen), gute Übersichtlichkeit, schnelle Reaktion und 24/7-Erreichbarkeit angestoßen werden. Zum Lernen des Chatbots gehört auch, zu erkennen, wie gut die Antworten passen (etwa, wenn sie vermehrt Rückfragen generiert haben) und ein Korrekturverfahren. Bei alledem ist nicht zu vergessen, dass *Machine-Learning* sehr große Datenmengen benötigt. Mittelständische Unternehmen, wie in unserem Fallbeispiel, stoßen hier an ihre Grenzen.

### Resümee zum Praxislabor

Aus Medien und im eigenen Gebrauch sind hohe technische Standards von Chatbots wohlbekannt. Diese funktionieren allerdings noch nicht zuverlässig im Business-Umfeld. In der Praxis wurde mit Blick auf die Gestaltung deshalb nach dem KISS-Prinzip vorgegangen (keep it simple). Chatbots funktionieren demnach schnell nicht mehr, wenn sie zu ambitioniert gestaltet werden. Hier kommt Erwartungsmanagement ins Spiel. In dem begleiteten Chatbot-Entwicklungsprozess des mittelständischen Unternehmens wurden Basis- und Ausbaustufen des Chatbots auch auf ihre Realisierbarkeit hin diskutiert.

Ein Reflexionsprozess zu einer kompetenzbasierten menschengerechten KI-Entwicklung, wie er im Artikel ausschnitthaft aufgezeigt wurde, hat das Potenzial zu einer guten mitarbeiterorientierten Technikgestaltung, und das in mehrerlei Hinsicht: 1) Die an der Arbeitspraxis ansetzende Anforderungsanalyse geht von den Herausforderungen der Mitarbeiter:innen aus. 2) Die Mitarbeiter:innen wägen technische Optionen mit Blick auf ihre Arbeitsprozesse ab und verweisen auf Grenzen der Digitalisierung/Automatisierung der Interaktionsarbeit mit den Kund:innen. 3) Der Stellenwert von Interaktionsarbeit als Arbeit mit Menschen wird gerade mit Blick auf die Grenzen der Automatisierung deutlich. Hieran knüpft sich die Frage nach deren guter Rahmung (organisatorisch, technisch, kompetenzseitig). Der Chatbot als Unterstützung hochwertiger Servicearbeit ist dabei nur ein Aspekt. 4) Zudem bedient die Reflexion aus der Kund:innenperspektive die in der Software-Entwicklung hinlänglich bekannte Frage, wer die Kunden/(End-)Nutzer sind. Lassen sich bei den Kund:innen Irritationen, Unterbrechungen, Umwege vermeiden? Unzufriedenheit schlägt letztendlich immer auf die Beratungs-/Servicemitarbeiter:innen zurück und belastet anstatt zu entlasten. Kunden-Orientierung wirkt sich auf die Arbeitsentlastung aus. Bezüglich Fruchtbarkeit des HAI-MMI-Konzepts für die Chatbot-Entwicklung kam das IT-Entwickler:innenteam zu dem Resümee, dass die herausgearbeiteten Qualitätsebenen und konkretisierten Gestaltungskriterien potenziell als Bewertungskriterien für eine Beta-Version dienen können. Dieser Test wie auch das tatsächliche Kund:innen-Feedback sind den nächsten Schritten nach Beendigung des Forschungs- und Entwicklungsprojektes vorbehalten.

### Referenzen

André, Elisabeth; Aurich, Jan; Bauer, Wilhelm; Bullinger, Angelika; Heister, Michael; Huchler, Norbert; Neuburger, Rahild; Peissner, Matthias; Stich, Andrea; Suchy, Oliver; Ramin, Philipp; Wächter, Michael (Hrsg.) (2021):

Kompetenzentwicklung für Künstliche Intelligenz – Veränderungen, Bedarfe und Handlungsoptionen. Whitepaper aus der Plattform Lernende Systeme. München.

- Böhle, Fritz (2017): Arbeit als Subjektivierendes Handeln. Handlungsfähigkeit bei Unwägbarkeiten und Ungewissheit. Springer VS, Wiesbaden.
- Böhle, Fritz; Wehrich, Margit (2020): Das Konzept der Interaktionsarbeit. In: Zeitschrift für Arbeitswissenschaft, 74, pp. 9–22.
- Böhmman, Tilo; Bittner, Eva; Biemann, Chris (2023): Das 3X3 der Chatbots. Erfolgsfaktoren für die Einführung und Verstetigung sprachbasierter KI in Organisationen. Universität Hamburg.
- Ciesinger, Kurt-Georg; Fink-Cvetnic, Tanja; Jungtäubl, Marc; Wehrmann, Jonas; Wehrich, Margit (2022): Bots & Co. – Die Zukunft der Interaktionsarbeit? In: Praeview. Zeitschrift für innovative Arbeitsgestaltung und Innovation, Nr. 1/2022, Essen: GMF / Gathmann Michaelis und Freunde Kommunikationsdesign.
- Gentsch, Peter (2019): Künstliche Intelligenz für Sales, Marketing und Service: Mit AI und Bots zu einem Algorithmic Business – Konzepte, Technologien und Best Practices. Wiesbaden: Springer Gabler.
- Hadi, Rhonda (2019): Chatbots als Persönlichkeiten. In: NIM Marketing Intelligence Review, Vol. 11, No. 2.
- Huchler, Norbert (2020): Die Mensch-Maschine-Interaktion bei Künstlicher Intelligenz im Sinne der Beschäftigten gestalten – Das HAI-MMI-Konzept und die Idee der Komplementarität. In: Digitale Welt vom 31.7.2020. Online unter: <https://digitaleweltmagazin.de/2020/07/31/die-menschmaschine-interaktion-bei-kuenstlicher-intelligenz-im-sinne-der-beschaeftigten-gestalten-das-hai-mmi-konzept-und-die-idee-der-komplementaritaet/> (letzter Zugriff: 6.6.2023).
- Huchler, Norbert; Adolph, Lars; André, Elisabeth; Bender, Nadine; Müller, Nadine; Neuburger, Rahild; Peissner, Matthias; Steil, Jochen, Jochen; Stowasser, Sascha; Suchy, Oliver (2020): Kriterien für die menschengerechte Gestaltung der Mensch-Maschine-Interaktion bei Lernenden Systemen, Whitepaper aus der Plattform Lernende Systeme, München.
- Porschen-Hueck, Stephanie; Zylowski, Thorsten (2021): Interaction-sensitive chatbot. Updating customer support in the platform-based ecosystem for business software. In: Bernd Bienzeisler, Katrin Peters, Alexander Schletz (Hrsg.): The disruptive Role of Data, AI and Ecosystems in Services – Proceedings of the 31. RESER Conference, S. 195-211, <http://publica.fraunhofer.de/dokumente/N-642928.html> (letzter Zugriff 12.6.2023).

### Anmerkungen

- 1 <https://de.humanfriendlyautomation.com/>, letzter Zugriff 16.07.2023
- 2 *Der Chatbot wurde im Rahmen des Forschungs- und Entwicklungsprojektes UMDIA, gefördert vom Bundesministerium für Bildung und Forschung und dem Europäischen Sozialfonds im Rahmen des Forschungsschwerpunktes „Zukunft der Arbeit: Arbeiten an und mit Menschen“ im Rahmen des FuE-Programms „Zukunft der Arbeit“ als Teil des Dachprogramms „Innovationen für die Produktion, Dienstleistung und Arbeit von morgen“ (FKZ 02L18A123) durchgeführt. Neben dem Dank an die Förderer gilt dieser insbesondere auch den Unternehmenspartner:innen mit ihrem Engagement in dem Teilprojekt zu unternehmensbezogener Interaktionsarbeit. Hintergründe zum Projekt siehe <https://unterbrechungen-bei-interaktionsarbeit.de/>, letzter Zugriff 13.7.2023.*
- 3 <https://de.eos-solutions.com/chatbot-survey-2021>, letzter Zugriff 28.6.2023
- 4 <http://projekt-hiss.>, letzter Zugriff 28.6.2023
- 5 <https://labora.digital/2020/markt-der-ideen/humanizing-machine-interaction-with-ai-hai-mmi/>, letzter Zugriff 16.7.2023
- 6 Vgl. die Diskussion zum interaktions(arbeits)sensiblen Chatbot in (Porschen-Hueck/Zylowski 2021).

- 7 Für das HAI-MMI-Instrument wurde unter anderem auf den Katalog der vom BMBF initiierten Plattform Lernende Systeme zurückgegriffen (Huchler et al. 2020). Siehe auch Andre et al. (2021).
- 8 Sie kann einseitig sein, d. h. dass die KI nicht als Interaktionspartner agiert. Das System arbeitet nur Hintergrundprozesse und Aufträge ab. Sie kann einfach sein, dann kann der Mensch verschiedene Funktionen des Systems ansteuern, bietet aber keine situationsadäquaten Lösungen an. In beiden Fällen werden die Potenziale der Mensch-Maschinen-Interaktion nicht voll ausgeschöpft, die einseitige Koordination steht sogar für eine schlechte Qualität der Mensch-Maschinen-Interaktion. Einem hohen Niveau der Mensch-Maschinen-Interaktion entspricht dagegen eine interaktive Gestaltung der Mensch-Maschi-

nen-Interaktion, die eine proaktive Koordinationsbeziehung befördern kann! Hier wird die Frage, wer steuert und entscheidet, transparent und interaktiv im Prozess abgestimmt. Das System bietet aktiv Aufgaben an, die es übernehmen kann. Sind die Koordinationsrollen zwischen Mensch und Maschine determiniert verteilt, sinkt diese Qualität wieder. Hier wäre beim Technikdesign der Einbezug des Menschen an konkreten Stellen festgelegt. Wenn der Mensch ferner nur noch für eine kompensierende und präventive Koordination in der Technikgestaltung vorgesehen ist, beispielsweise zur Fehlerbehebung, -vermeidung oder Verantwortungsübernahme, sinkt die Qualität der MMI wieder, sie wird dem HAI-MMI Konzept nach sogar als qualitativ schlecht bewertet.



Christoph Schmitz

## IT-Gestaltung für Gute Arbeit

### Pro Partizipative Technik- und Arbeitsgestaltung

*Gestaltung von Informationstechnik (IT) für Gute Arbeit heißt in erster Linie: mit der Technik – wenn sie im Arbeitsprozess eingesetzt werden soll – auch direkt die Arbeitsbedingungen gut zu gestalten, diese von vornherein mitzudenken, beides zusammenzudenken. Voraussetzung für eine gute Gestaltung von IT und Arbeit ist die Beteiligung der Beschäftigten und ihrer Interessenvertretungen – also derjenigen, die mit der Technik arbeiten und umgehen sollen.*

*ver.di hat bereits vor Jahren Leitlinien für gute digitale Arbeit aufgestellt: zuerst in der Enquête-Kommission Internet und digitale Gesellschaft (2010 – 2013), in der erstmals solche Leitlinien fraktionsübergreifend verabschiedet wurden; dann auch im gewerkschaftlichen Kontext, wo ein Leitbild auf dem 20. DGB-Kongress 2014 (ver.di 2014, S. 22 ff.) und auf dem ver.di-Kongress 2015 beschlossen wurde.*

### Gute digitale Arbeit als Leitbild

Die Leitlinien fordern insbesondere einen dringenden Ausbau der Mitbestimmung, der der Digitalisierung Rechnung trägt (vgl. ver.di 2016). Das wurde zum Teil in einen Vorschlag der Gewerkschaften für eine Reform des Betriebsverfassungs-Gesetzes aufgenommen (im *Gesetzentwurf für ein modernes Betriebsverfassungsgesetz*, Sonderausgabe der Zeitschrift *Arbeit und Recht* vom April 2022), denn das inzwischen verabschiedete Betriebsrätemodernisierungs-Gesetz reicht bei Weitem nicht aus. Das bedeutet, dem § 80 Absatz 3 folgenden Satz hinzuzufügen: „Muss der Betriebsrat zur Durchführung seiner Aufgaben die Einführung oder Anwendung von Künstlicher Intelligenz beurteilen, gilt insoweit die Hinzuziehung eines Sachverständigen als erforderlich.“ (vgl. Müller 2023, S. 84)

ver.di fordert bereits seit Langem mit den anderen Gewerkschaften u. a. ein Recht auf Weiterbildung und eine staatlich geförderte Weiterbildungsteilzeit. Neben der Förderung von Beschäftigung in Form von gut gestalteten Arbeitsplätzen mit fairer Entlohnung ist das Teil des Leitbildes gute digitale Arbeit. Zudem setzt sich ver.di seit vielen Jahren dafür ein, dass die psychischen Belastungen durch eine zu hohe Arbeitsmenge und überzogene Erreichbarkeits-Erwartungen abgebaut werden müssen. Dazu soll der Arbeits- und Gesundheitsschutz verbindlich umgesetzt werden. Deshalb fordert ver.di Sanktionen, wenn beispielsweise in den Betrieben und Verwaltungen die gesetzlich vor-

geschriebene Gefährdungsbeurteilung nicht durchgeführt wird. Auch muss dringend das Aufsichtspersonal aufgestockt werden, das die Umsetzung in den Betrieben und Verwaltungen kontrolliert. Die Beschäftigten sollen an der Gefährdungsbeurteilung mitwirken und sich einbringen können. Nur mit ihrer Beteiligung können bei der Beurteilung aussagekräftige und praxisrelevante Ergebnisse herauskommen; ihre Erfahrungen münden in der Regel auch in wirksame und vernünftige Abhilfe- und Schutzmaßnahmen.

Ein weiterer wichtiger Aspekt für gute digitale Arbeit ist der Schutz der Persönlichkeitsrechte, wozu es in der Arbeitswelt ein Beschäftigtendatenschutz-Gesetz braucht – denn abhängig Beschäftigte benötigen besondere Schutzrechte. ver.di fordert nicht nur solch ein Gesetz, sondern unterstützt auch gesetzliche Interessenvertretungen mit Schulungen zum Thema wie auch einem Werkzeug zur Bewertung des Beschäftigtendatenschutzes im Betrieb bzw. in der Verwaltung (vgl. Brandl 2023).

### Die Realität der digitalen Arbeitswelt

Gute Arbeitsbedingungen sind keine Selbstverständlichkeit, sie müssen erkämpft und gestaltet werden – auch im Zuge der Digitalisierung, wenn IT und vor allem Software zunehmend eine zentrale Rolle in der Arbeitswelt spielen. Eine erste Befragung mit dem DGB-Index *Gute Arbeit* schon im Jahr 2016 hat für

den Dienstleistungssektor ergeben, dass vor allem der Stress bei digitaler Arbeit größer ist als bei nicht oder kaum digitalisierter Arbeit (vgl. Roth 2017). In der IKT-Branche<sup>1</sup> zeigt sich das besonders deutlich: Fast die Hälfte der Befragten gibt an, dass die Arbeitsbelastung mit der Digitalisierung gewachsen ist. Zugewonnen hat auch die Arbeitsmenge. Die Belastungen wiederum beeinflussen maßgeblich die Arbeitsfähigkeit der Beschäftigten.

Die Ergebnisse belegen: Um Belastungen abzubauen und Arbeitshetze zu verringern, müssen die Beschäftigten an der Gestaltung ihrer Arbeit beteiligt werden. Wer seine Arbeit selbständig planen kann und Einfluss auf die Arbeitsmenge hat, gibt seltener an, sich (sehr) häufig in der Arbeit gehetzt zu fühlen. Auch schon damals konnte festgestellt werden: 75 % der Beschäftigten haben keinen oder nur geringen Einfluss auf die Art und Weise des Technikeinsatzes, was dazu beiträgt, dass sich 45 % der Befragten der digitalen Technik sehr häufig oder oft ausgeliefert fühlen (ebd., S. 32 ff.).

Seit einigen Jahren liegen also die gewerkschaftlichen Forderungen, Zahlen zur Qualität der Arbeitsbedingungen und Vorschläge zu ihrer Verbesserung im Kontext der Digitalisierung vor. Jedoch muss anhand aktueller Zahlen festgestellt werden, dass weiterhin großer Handlungsbedarf besteht: Die Corona-Krise hat den Trend zur Digitalisierung und Überarbeitung verstärkt. In der Pandemie habe die Zahl der geleisteten Arbeitsstunden um 10 % zugelegt.<sup>2</sup> Vor allem durch den steigenden Anteil von Homeoffice sei in vielen Branchen die Grenze zwischen Freizeit und Arbeit verschwommen, viele Beschäftigte würden und müssten länger arbeiten.

Die Digitalisierung der Arbeitswelt hat durch Corona einen Schub erfahren. Das zeigten schon die Daten der Repräsentativbefragung mit dem DGB-Index *Gute Arbeit 2021*. 46 % der Befragten berichten von neuer Software und neu eingesetzten Apps. Nach dem DGB-Index *Gute Arbeit* bekommen jedoch ein Viertel der Beschäftigten nicht die Unterstützung, die sie für den Umgang mit der neuen Technik brauchen würden. Ebenfalls ein Viertel erhält nicht die notwendigen Schulungen, um mit den neu eingeführten digitalen Arbeitsmitteln gut umgehen zu können (vgl. Institut DGB 2021, S. 10 f.). Das führt zu mehr Belastungen (vgl. Müller 2023, S. 77).

Im Jahr 2022 hat der DGB eine erneute Beschäftigtenbefragung mit dem Schwerpunkt Digitalisierung durchgeführt. ver.di hat dazu wieder eine Sonderauswertung vorgelegt. Es zeigt sich: Noch immer liegt die durchschnittliche Arbeitsintensität in einem kritischen und gesundheitsgefährdenden Bereich. 54 % der Beschäftigten im Dienstleistungssektor fühlen sich in der Arbeit gehetzt oder stehen unter Zeitdruck. Auf den Einsatz digitaler Technik am Arbeitsplatz können lediglich knapp ein Viertel (24 %) der befragten Beschäftigten mit einem hohen Digitalisierungsgrad der Arbeit in (sehr) hohem Maß Einfluss nehmen. Das war bereits 2016 so. Auch hier gaben lediglich 25 % der Befragten an, in (sehr) hohem Maß Einfluss auf die Technikgestaltung am Arbeitsplatz nehmen zu können. Unter den Frauen ist der Anteil derjenigen, die Einfluss nehmen können, sogar um 11 Prozentpunkte geringer als unter den Männern. Je geringer der Einfluss auf die Technikgestaltung am Arbeitsplatz ist, desto eher dominiert das Gefühl der digitalen Technik ausge-

liefert zu sein: 45 % der Beschäftigten, die keinen Einfluss haben, fühlen sich in (sehr) hohem Maß der Technik ausgeliefert. Demgegenüber sind es von denjenigen mit einem sehr hohen Einfluss auf die Technikgestaltung an ihrem Arbeitsplatz 19 % (vgl. Roth 2023).

### Künstliche Intelligenz: Für *Gute Arbeit by Design*

Mit dem Einsatz von Künstlicher Intelligenz (KI) wird es für die Beschäftigten und ihre Interessenvertretungen noch dringlicher, Einfluss auf die Technikgestaltung zu nehmen. Nicht erst seit es ChatGPT gibt, beschäftigt sich ver.di mit KI. Dass KI nach ein paar so genannten *KI-Wintern* (an ihrer Entwicklung wird ja bereits seit 1955 gearbeitet) wieder einen größeren Stellenwert in der Digitalisierungsdiskussion einnimmt, hat im Wesentlichen zwei Ursachen:

*Erstens* sind im Bereich der *Künstlichen Intelligenz* vor allem auf dem Gebiet des *maschinellen Lernens* und des *tiefen Lernens* (*deep learning*) Fortschritte aufgrund der zunehmenden Daten und Rechen- sowie Speicherleistungen gemacht worden. Im Unterschied zu herkömmlicher Software wird bei einer KI nicht jeder Schritt programmiert, sondern die Systeme *lernen* bzw. verbessern sich bis zu einem gewissen Grad *selbstständig* und in unvorhergesehener bzw. unvorhersehbarer Weise. Haben sich die Algorithmen *eingelernt*, sind die getroffenen Entscheidungen bzw. Ergebnisse des Systems schwer nachvollziehbar. Das birgt Risiken, und es gibt Bemühungen, dieses Problem auch technisch zu lösen.

*Zweitens* hatten vor allem die USA und China ihre Investitionen in diese Technik stark erhöht. Die Unternehmensberatung McKinsey & Company schätzt, dass 2016 weltweit etwa 39 Mrd. US-Dollar für die Forschung und Entwicklung von KI-Technologien bereitgestellt wurden, dreimal so viel wie noch 2013 – Tendenz stark steigend. Damit ist ein Wettbewerb eingeläutet worden, bei dem es wohl vor allem darum geht, Geschäftsmodelle mit KI zu entwickeln und Gewinne zu machen: Gewinne mit einer Technik, die einerseits Fortschritte mit sich bringt, was sich etwa bei der enormen Verbesserung in der Diagnostik seltener Erkrankungen und Tumore zeigt. Andererseits birgt KI eben Risiken. Und so sind auch in Europa und der BRD in den letzten Jahren die Investitionen in die Erforschung von KI gestiegen. „In den Jahren 2013 bis 2022 beliefen sich die gesamten privaten Investitionen im Bereich Künstliche Intelligenz in Deutschland auf insgesamt rund sieben Milliarden US-Dollar. Im gleichen Zeitraum wurden in den USA private Investitionen in Höhe von rund 249 Milliarden US-Dollar getätigt.“ (Statista Research Department, 20. April 2023)

Inzwischen haben sich verschiedene Akteure in die Debatte eingeschaltet. Nach Frankreich hatte die Bundesregierung ebenfalls eine KI-Strategie vorgelegt, in der auch „Ziele für Gute Arbeit“ vorkommen: „Wir setzen uns für eine im gesellschaftlichen Diskurs erarbeitete gemeinsame Verständigung auf eine Definition von Zielen für ‚Gute Arbeit by Design‘ ein, die in einer ‚Digital Bill of Rights‘ mit Blick auf die Bedürfnisse und Herausforderungen der Informationsgesellschaft präzisiert werden“ (Bundesregierung 2018, S. 40). An einer EU-KI-Verordnung wird derzeit gearbeitet (vgl. Müller 2023, S. 79).

## Leitlinien KI

In diesen Prozess haben sich auch die Gewerkschaften eingebracht, nicht nur in den Arbeitsgruppen der *Plattform Lernende Systeme*, die 2017 vom Bundesministerium für Bildung und Forschung (BMBF) auf Anregung des Fachforums *Autonome Systeme* des Hightech-Forums und acatech gegründet wurde, sondern auch in weiteren politischen Gremien wie der Enquête-Kommission zu KI. ver.di hat ihre Positionen für eine progressive KI zum Wohl der Allgemeinheit, der Demokratie, guter Dienstleistungen und Guter Arbeit zeitgleich mit der KI-Strategie der Bundesregierung im November 2018 vorgestellt (vgl. ver.di 2018).

### KI: Erste ver.di-Positionierung – ein Auszug

- Dialogische Erarbeitung einer gesellschaftlichen Vision als Leitgedanke für den KI-Einsatz in Deutschland
- Rückbindung der KI-Entwicklung an demokratische Verfahren und Einbindung in gesellschaftliche Diskussionen
- klare Ethikregeln und Kriterien für Gute Arbeit, u. a. Verbot autonomer Waffensysteme
- Schaffung einer verantwortlichen Stelle in den zu überwachenden Regeln: stringente Zweckbindung, Folgenabschätzung entsprechend der EU-DSGVO, Eingriffsmechanismen
- Rechenschaftspflichten: Transparenz der Funktionsmechanismen und Entscheidungsparametern (white box), Whistleblower-Schutz
- eine Ausweitung der KI-Sicherheits- und Arbeitsforschung sowie Technikfolgenabschätzung
- eine gründliche Erprobung und gründliches Testen von KI-Anwendungen (konzeptionelle Entschleunigung)
- Kennzeichnungspflicht für Chatbots (Erklärung s. u.) bei deren Einsatz und von KI-generierten Medieninhalten
- Offenlegung und Umverteilung von Effizienzgewinnen durch KI in gesellschaftliche Bedarfssfelder wie Gesundheit, Pflege, Bildung und Mobilität

- frühestmögliche Beteiligung von Mitbestimmungsträgern und Beschäftigten: Stärkung und Ausbau der Mitbestimmung (u. a. Vereinfachung der Gründung von Betriebsräten) sowie KI-Beratung für Mitbestimmungsakteure
- Wahrung von Persönlichkeitsrechten und der Menschenwürde

Diese Positionen sind im Jahr darauf in einer überarbeiteten Form auf dem jährlichen ver.di-Digitalisierungskongress zum Schwerpunkt KI nochmals vorgestellt und diskutiert worden (vgl. ver.di 2019). Zudem hat ver.di auf dem Bundeskongress im September 2019 im Beschluss *Künstliche Intelligenz und neue Arbeitsformen gemeinwohldienlich und menschengerecht gestalten* formuliert, dass sie „ihre Leitlinien zu guter digitaler Arbeit insbesondere mit Blick auf aktuelle Entwicklungen in den Bereichen Künstliche Intelligenz und neuer Arbeitsformen weiterentwickeln“ wird. Es gab den Bedarf in unserer Organisation, sich auch in die Debatte um ethische Leitlinien für den KI-Einsatz einzubringen. ver.di entwickelte solche Leitlinien und stellte sie zu Beginn des Jahres 2020 zur Diskussion. Die Leitlinien sollen als gewerkschaftlicher Beitrag in der KI-Debatte insbesondere Orientierung und Hilfestellung für diejenigen sein, die KI-Anwendungen entwickeln, einführen und nutzen.

Neben der Beschäftigungssicherung und besseren Qualifikationsmöglichkeiten geht es in den ethischen Leitlinien aus gewerkschaftlicher Sicht darum, KI-Systeme so zu gestalten, dass sie die Handlungs- und Gestaltungsspielräume der Erwerbstätigen erweitern. Statt Tätigkeiten zu entwerten, sind sie durch gezielte Qualifizierung aufzuwerten. Das lässt sich durch Tarifverträge flankieren. KI-Systeme, die auf der Verarbeitung großer Datenmengen basieren, dürften nicht zu einer Gefährdung der Persönlichkeitsrechte führen, und die Zweckbindung der Daten ist sicherzustellen. Zudem dürfen Verantwortlichkeiten und Haftung nicht auf die Technik übertragen werden. Der Mensch bleibt in der Verantwortung.

Im Vordergrund steht der Ansatz *Gute Arbeit by Design*: So wie beim Prinzip *Privacy by Design* die Persönlichkeitsrechte von Beginn an zu beachten sind, muss bei der Konzeption und Gestaltung Künstlicher Intelligenz von vornherein die Arbeitsgestaltung im Fokus sein. Das kann am besten dadurch gewährleistet werden, dass Beschäftigte und ihre Interessenvertretungen rechtzeitig im Gestaltungs- und Einführungsprozess beteiligt werden. Künstliche Intelligenz ist Mittel zum Zweck, für

## Christoph Schmitz



**Christoph Schmitz** ist seit 2009 bei der Gewerkschaft ver.di beschäftigt und seit 2019 Mitglied im ver.di-Bundesvorstand als Leiter des Fachbereichs Finanzdienste, Kommunikation und Technologie, Kultur, Ver- und Entsorgung und auch zuständig für den Bereich Innovation und Gute Arbeit. Digitalisierung ist für Christoph Schmitz ein Gestaltungsthema, das nur dann gut funktioniert, wenn die Beschäftigten und ihre Interessenvertretungen mitgenommen werden. ver.di steht deshalb auch für einen Demokratisierungsprozess in Wirtschaft und Arbeitswelt und hat Gute digitale Arbeit als Leitbild entwickelt.

Foto: Kay Herschelmann

Beschäftigte ist es vor allem ein Arbeitsmittel. Sie sollten bei der Zweckbestimmung mitreden. Gerade für KI, die je nach Kritikalitätsstufe Risiken birgt, sind rote Haltelinien einzuziehen: Es darf keine KI-Anwendungen geben, die Menschen schaden, die gegen Menschen- und Grundrechte verstoßen. Bestehendes Recht und geltende Gesetze; insbesondere die Tarifautonomie, das Arbeitsschutz-Gesetz, das Betriebsverfassungs-Gesetz, das Urheberrecht und die europäische Datenschutzgrundverordnung sind einzuhalten.

Was den Einsatz von KI im Dienstleistungssektor betrifft, gibt das ver.di-Innovationsbarometer erste Hinweise zu deren Auswirkung: So rechnen 66 % der befragten Betriebs-, Personal- und Aufsichtsräte in betroffenen Unternehmen mit einer sinkenden Zahl von Arbeitsplätzen durch den KI-Einsatz. 42 % der Befragten verzeichnen häufigere Störungen der Arbeitsabläufe. Dass die Kompetenz- und Qualifikationsanforderungen zunehmen, davon gehen 51 % aus. Und 52 % der mit dem ver.di-Innovationsbarometer befragten Betriebs-, Personal- und Aufsichtsräte in ver.di meinen zudem, dass sich (auch) mit dem KI-Einsatz die Arbeitsintensität erhöht hat. 50 % der Befragten erkennen eine Zunahme der Transparenz des Arbeits- und Leistungsverhaltens der Beschäftigten. 60 % berichten von einer Verringerung der Handlungs- und Entscheidungsspielräume durch KI und 88 % meinen, dass sie starke Mitbestimmungsrechte schon bei der Planung des KI-Einsatzes benötigen (vgl. Zanker et al 2019; Müller 2023, S. 83 f.).

### Gestalten mit ver.di

Um bessere Arbeitsbedingungen wie auch mehr Mitbestimmung zu erreichen, hat ver.di sich nicht nur in politischen Gremien für diese Forderungen eingesetzt, sondern in den letzten Jahren vermehrt Kollektivvereinbarungen abgeschlossen. Dazu hat ver.di Empfehlungen vorgelegt (vgl. ver.di 2020, 2021). Sie beziehen sich im Wesentlichen auf Prozessvereinbarungen, die die Beteiligungs- und Mitbestimmungsmöglichkeiten für die Beschäftigten und v. a. ihre gewerkschaftlichen sowie betrieblichen Vertreter:innen erweitern. Folgende Themen sollten geregelt werden – wenn das nicht durch andere Kollektivvereinbarungen bereits geschehen ist:

- Qualifizierung und Personalentwicklung
- Führungskultur
- Arbeitszeit
- Gesundheitsschutz
- Beschäftigungssicherung bzw. sozialverträglicher Ausstieg
- Beschäftigtendatenschutz

Beispiele sind der IBM-Tarifvertrag zum Gesundheitsschutz; der Belastungsschutz-Tarifvertrag bei der Telekom, die Entlastungs-Tarifverträge bei den Kliniken wie der Tarifvertrag zur Personalbemessung bei der Charité und auch der Eurogate-Tarifvertrag Zukunft enthalten Elemente zum Arbeits- und Gesundheitsschutz wie vor allem eine Stärkung der Mitbestimmung. Solch eine Prozessvereinbarung in Form einer „Vorgehensvereinbarung für den Umgang mit digitalen Projekten ...“, die einen intensiven Beratungs- und Beteiligungsprozess vorsieht und auch beschreibt“ (ver.di 2020a), hat zudem der Betriebsrat bei der ERGO-Versicherungsgruppe abgeschlossen. Bei Sportscheck

(ehemals Karstadt Sport) ist im Tarifvertrag *Gute und gesunde Arbeit* eine regelmäßige Befragung nach DGB-Index Gute Arbeit verankert worden, mit den Index- und bestimmten Zusatzfragen zur Digitalisierung. Die Befragung soll alle zwei Jahre durchgeführt werden. Die erste Befragung ist abgeschlossen, und es hat sich knapp die Hälfte der Beschäftigten beteiligt. Ziel ist vor allem, gute und gesunde Arbeitsbedingungen beteiligungsorientiert zu gestalten.

Auch zu anderen Themen im Kontext der Digitalisierung hat ver.di in den letzten Jahren vermehrt Vereinbarungen geschlossen, um Verbesserungen für die Erwerbstätigen zu erzielen. Dazu gehören beispielsweise der *Tarifvertrag Digitalisierung Bund* sowie *zu den mobilen Arbeitsformen*, weiter ein Tarifabschluss bei den Versicherungen zu Beschäftigungssicherung und Qualifizierung und der IBM-Tarifvertrag *Qualifizierung und Nachwuchssicherung* (vgl. Müller 2023, S. 82 ff.).

Für die Empfehlungen zu tariflicher und betrieblicher Gestaltung hat ver.di die Publikationsreihe *Praxis gestalten* aufgelegt. Die Broschüren bieten Tipps für ein effektives, strategisches Vorgehen und bieten jeweils Bausteine für die Ausgestaltung von Betriebs- und Dienstvereinbarungen sowie Tarifverträgen. Der ver.di-Bereich *Innovation und Gute Arbeit* sowie die Grundsatzabteilung Tarifpolitik haben dazu insgesamt vier Broschüren der Reihe publiziert, zuletzt die beiden Bände *Gesunde Arbeit* und *Digitale Arbeit*, um Veränderungsprozesse in der Digitalisierung zu gestalten (zuvor *Mobile Arbeit* und *Agiles Arbeiten*). Die Bände der Reihe *Praxis gestalten* vermitteln kurz und bündig Hintergrund, Anregungen für betriebliche Lösungen und einen Fahrplan für das Vorgehen. Die Unterpunkte enthalten jeweils Kriterien und Regelungspunkte für Gute Arbeit sowie Auszüge aus bereits umgesetzten Vereinbarungen und Tarifverträgen. Die Gliederung der Publikationen orientiert sich am Aufbau von Vereinbarungen.

### Ausblick

Auch auf dem ver.di-Bundeskongress im September 2023 sind die Digitalisierung und damit die Gestaltung von IT, KI und ihres Einsatzes ein Schwerpunkt, bei dem die Arbeitsbedingungen eine zentrale Rolle spielen. ver.di ist und bleibt aktiv in diesem Bereich. Beispielsweise bietet ver.di seit dem letzten Sommer eine Online-Veranstaltungsreihe zum Thema KI an. In Zukunft wird ver.di sich noch intensiver in den Branchen mit der betrieblichen Realität des Einsatzes von KI beschäftigen und ggf. Forderungen und Aktivitäten schärfen. ver.di bietet bereits Seminare für gesetzliche Interessenvertretungen nicht nur zum Thema KI an. Es gibt dazu auch erste Handlungshilfen für gewerkschaftliche Akteure (vgl. Schröder, Höfers 2022, S. 196; TBS Berlin, ver.di 2020). Nach und nach können erste Ergebnisse aus verschiedenen von den Ministerien geförderten KI-Projekten bei der Gestaltung von Technik und Arbeit berücksichtigt werden.<sup>3</sup> Dreh- und Angelpunkt bleibt die Beteiligung der Beschäftigten und ihrer Interessenvertretungen: Deshalb braucht es einen Demokratisierungsprozess in Wirtschaft und Arbeitswelt (vgl. Schmitz/Urban 2021). Die Erwerbstätigen müssen stärker an der nachhaltigen und guten Gestaltung ihrer Arbeitsprozesse und beim Technikeinsatz beteiligt werden. Auch deshalb hat ver.di auf ihrem letzten Bundeskongress eine *Offensive für Gute Arbeit* beschlossen.

## Referenzen

Brandl, Ch. (2023): Projekt BeDaX – ein Prüftool für den Beschäftigtendenschutz, in: Roth, I. (2023)

Bundesregierung (2018): Strategie Künstliche Intelligenz, 2020

Institut DGB-Index Gute Arbeit (2021): Jahresbericht 2021. Ergebnisse der Beschäftigtenbefragung zum DGB-Index Gute Arbeit, Berlin

Müller, Nadine (2023): Digitalisierung nach vorn gedacht, in: Jahrbuch Gute Arbeit 2023 hrsg. von Schmitz, Ch./Urban, H.-J., S. 77-88

Roth, I., u. Mitarb. v. Müller, N. (2017): Digitalisierung und Arbeitsqualität. Eine Sonderauswertung des DGB-Index Gute Arbeit 2016 für den Dienstleistungssektor, [www.innovation-gute-arbeit.verdi.de/themen/digitale-arbeit](http://www.innovation-gute-arbeit.verdi.de/themen/digitale-arbeit)

Roth, I., u. Mitarb. v. Müller, N. (2023, im Erscheinen): Arbeitsbedingungen in der Digitalisierung. Eine Sonderauswertung des DGB-Index Gute Arbeit 2022 für den Dienstleistungssektor, [www.innovation-gute-arbeit.verdi.de/themen/digitale-arbeit](http://www.innovation-gute-arbeit.verdi.de/themen/digitale-arbeit)

Schmitz, Christoph/Urban, Hans-Jürgen (Hrsg., 2021): Demokratie in der Arbeit. Eine vergessene Dimension der Arbeitspolitik? Frankfurt am Main

Schröder, L.; Höfers, P. (2022): KI Lagom. Die Balance zwischen künstlicher Intelligenz und menschlichen Werten, Frankfurt/M. <https://innovation-gute-arbeit.verdi.de/themen/digitale-arbeit/publikationen/++co++85856aec-2035-11ec-9925-001a4a160100>

Statista Research Department (2023): Private Investitionen in KI in ausgewählten Ländern 2013-2022; <https://de.statista.com/statistik/daten/studie/1321707/umfrage/private-investitionen-in-ki-in-ausgewaehlten-laendern/#:~:text=In%20den%20Jahren%202013%20bis%202022%20beliefen%20sich,in%20H%C3%B6he%20von%20rund%20249%20Milliarden%20US-Dollar%20get%C3%A4tigt>.

TBS Berlin GmbH, ver.di (2020): Handreichung zur Gestaltung und Anwendung Künstlicher Intelligenz (KI), Berlin

ver.di (2014): Digitalisierung und Dienstleistungen – Perspektiven Guter Arbeit, Berlin

ver.di (2016): Arbeiten 4.0 braucht gleichberechtigte Teilhabe! Mehr Mitbestimmung und Demokratie in der digitalen Arbeitswelt, Diskussionspapier,

Berlin; <https://innovation-gute-arbeit.verdi.de/themen/digitale-arbeit/beschlusse-und-positionen>

ver.di (2018): WOZU, WAS und WIE sollen Algorithmen lernen? Gute Arbeit und Künstliche Intelligenz, Berlin, <https://innovation-gute-arbeit.verdi.de/themen/digitale-arbeit/beschlusse-und-positionen>

ver.di (2019): Künstliche Intelligenz und Gute Arbeit gestalten. Wir mischen uns ein, gestalten, regeln – mit Euch gemeinsam! Berlin. [www.verdi.de/themen/digitalisierungskongresse/kongress-2019](http://www.verdi.de/themen/digitalisierungskongresse/kongress-2019)

ver.di (2020): Gesunde Arbeit. Empfehlungen für die tarif- und betriebspolitische Gestaltung, Berlin. [www.innovation-gute-arbeit.verdi.de/themen/digitale-arbeit](http://www.innovation-gute-arbeit.verdi.de/themen/digitale-arbeit)

ver.di (2020a): Die ERGO Versicherungsgruppe „Wie ein Mensch denken, das kann der Bot nicht“. Praxisreport zur Digitalisierung in der Versicherungswirtschaft, Berlin

ver.di (2020b): Ethische Leitlinien für die Entwicklung und den Einsatz von Künstlicher Intelligenz (KI): Gemeinwohl und Gute Arbeit by Design, Diskussionspapier, Berlin, <https://innovation-gute-arbeit.verdi.de/themen/digitale-arbeit/beschlusse-und-positionen>

ver.di (2021): Digitale Arbeit. Veränderungsprozesse tarif- und betriebspolitisch gestalten, Berlin, [www.innovation-gute-arbeit.verdi.de/themen/digitale-arbeit](http://www.innovation-gute-arbeit.verdi.de/themen/digitale-arbeit)

Zanker, C., Roth, I.; Hoppe, M. (2019): ver.di-Innovationsbarometer Künstliche Intelligenz, hrsg. vom ver.di-Bereich Innovation und Gute Arbeit, Berlin, [www.innovation-gute-arbeit.verdi.de/innovation/innovationsbarometer](http://www.innovation-gute-arbeit.verdi.de/innovation/innovationsbarometer)

## Anmerkungen

- 1 *Informations- und Kommunikationstechnik*
- 2 *WHO-Experte Pega mit Verweis auf eine Studie des National Bureau of Economic Research in 15 Ländern.*
- 3 *Beispiele sind das Projekt KIDD – Künstliche Intelligenz im Dienste der Diversität (<https://kidd-prozess.de/>); das Projekt Humain WorkLab (<https://www.humain-worklab.de/>) oder auch KomKI (<https://projekt-komki.de/>).*



Sabine Pfeiffer

## KI als Kollegin?

### Wie wünschen sich Beschäftigte Künstliche Intelligenz am Arbeitsplatz?

*Über die Folgen von Künstlicher Intelligenz (KI) für Arbeit und Gesellschaft gibt es einen breiten Diskurs. Was aber Beschäftigte über KI in ihrem direkten Arbeitsumfeld denken, ist bislang kaum erforscht. Der Beitrag stellt Ergebnisse einer repräsentativen Erhebung vor und konzentriert sich vor allem auf Beschäftigte im Gesundheits-/Sozialwesen und der Logistik.*

### KI und Arbeit – zur Einleitung

Der Einsatz von Künstlicher Intelligenz oder Maschinellem Lernen (nachfolgend KI/ML) am Arbeitsplatz steht hier im Mittelpunkt, wird ihr doch vor allem das Potenzial zu weitgehender Automatisierung – vor allem nicht-stofflicher – Arbeitsprozesse zugesprochen. Hier aber soll es weder um Prognosen zu Automatisierungspotenzialen noch um die Thematik des Bias gehen – etwa bei KI-gestützten Personalauswahlverfahren. Im Zentrum steht hier, wie Beschäftigte die Thematik sehen und ob und wie sie sich KI am Arbeitsplatz vorstellen können.

Bisher sind Studien zur Einstellung von Beschäftigten dünn gesät, beziehen sich auf nur einen Anwendungskontext (z. B. Medizin) oder wurden in angelsächsischen oder asiatischen Arbeitswelten erhoben – wie etwa die Studie *Fear of autonomous robots and artificial intelligence (FARAI)* (Liang und Lee 2017) – und erlauben wegen unterschiedlicher Arbeitsmarktregulierung und anderen Mitbestimmungskulturen keine Übertragung auf Deutschland. Einzelne deutsche Studien wie der *Meinungsmonitor Künstliche Intelligenz* (Dosenovic et al. 2020) oder eine Umfrage der *D21-Initiative* (2020) befragen nur zum Teil Beschäftigte und beziehen sich in ihren Fragen kaum auf die Arbeitswelt.

Unsere Befragung *KI als Kollegin* (KIK) wurde als repräsentative<sup>1</sup> Online-Befragung mit N=2.018 im Februar 2022 durchgeführt. Über ein aktives Quoten-Management und eine iterative Gewichtung konnte ein nach Geschlecht, Alter, Qualifikationshöhe und Branche den aktuellen Arbeitsmarktdaten entsprechendes Sample gebildet werden. Die Rekrutierung erfolgte über ein professionelles Befragungsinstitut, für die Erhebung wurde das Online-Werkzeug *SoSci-Survey* verwendet (Leiner 2019), das DSGVO-konform auf einem lehrstuhleigenen Server des Rechenzentrums der FAU Erlangen-Nürnberg installiert ist (Pfeiffer 2021).

Für diesen Artikel werden auf Wunsch der Herausgeber:innen Beschäftigte mit Berufen im Bereich Logistik (6 %; nachfolgend LOG) und Gesundheitswesen (10 %; nachfolgend GES) gesondert ausgewertet und den Daten aller anderen Beschäftigten (84 %; nachfolgend DER) gegenübergestellt. *Abbildung 1* zeigt, dass die drei Stichproben nach Geschlecht, Alter und Qualifikationshöhe differieren. Grob gesagt und erwartbar zeigt sich LOG männlicher und geringer qualifiziert und GES weiblicher und höher qualifiziert, die Altersverteilungen sind weitgehend vergleichbar.

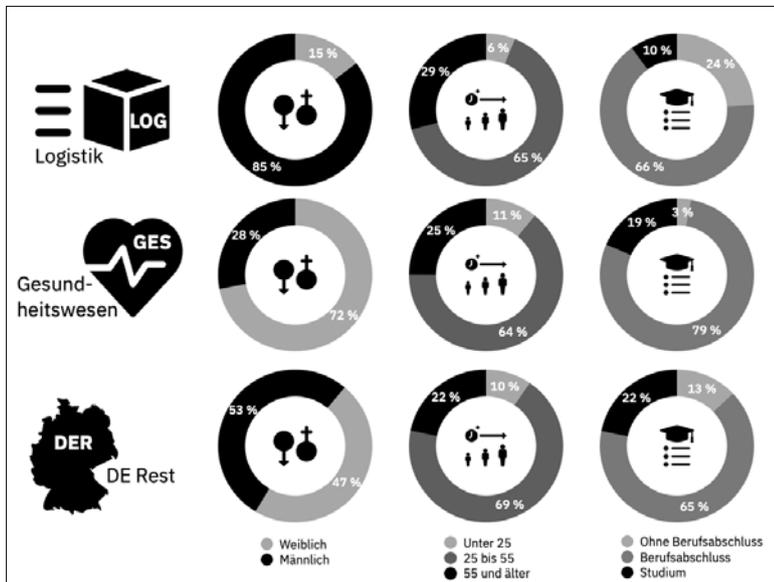


Abbildung 1: Zentrale Stichprobenmerkmale im Vergleich

Nachfolgend werden einige der Ergebnisse der Studie für die drei Stichproben vorgestellt:

### KI am Arbeitsplatz, grundsätzliche Einstellung, Wissensstand

*Abbildung 2* (obere Balkengrafik) zeigt, dass im Vergleich zu allen anderen Beschäftigten in Deutschland in der Logistik 7 % (LOG) und 9 % (GES) KI aktuell weniger verbreitet (DER 11 %). In den Unternehmen je weiterer 11 % (GES und DER) bzw. 14 % (LOG) ist eine konkrete KI-Umsetzung bereits geplant. Trotzdem und wenig überraschend erlebt die eindeutige Mehrheit (LOG: 68 %, GES 67 % und DER 61 %) KI überhaupt nicht als Thema am Arbeitsplatz.

Möglicherweise, weil die Auswirkungen algorithmischer Steuerung dort schon in der Breite erlebt werden, findet sich in LOG

mit 24 % die höchste grundsätzliche Ablehnung von KI am eigenen Arbeitsplatz (*Balkengrafik Mitte*). Mit 17 % eindeutiger Ablehnung liegt auch GES über dem Wert für DER von 15 %. Interessanter ist vielleicht, dass die Mehrheit sehr wohl offen ist: In allen drei Stichproben würde rund die Hälfte der Beschäftigten KI am Arbeitsplatz unter bestimmten Bedingungen akzeptieren und zwischen einem Viertel (LOG 26 %) und rund einem Drittel (32 % GES und 36 % DER) stehen KI in der Arbeit grundsätzlich offen. Insgesamt also keine Spur von den medial gerne beschworenen ängstlichen Beschäftigten, allerdings zeigen sich LOG und GES skeptischer als der Rest der Beschäftigten.

Die untere Balkengrafik in *Abbildung 2* verdeutlicht den selbst eingeschätzten Wissensstand zum Thema KI. Obwohl sich in allen drei Stichproben grafisch kaum mehr darstellbar kleine Anteile selbsterklärter KI-Expert:innen finden, gibt mit jeweils über 70 % die deutliche Mehrheit an, mindestens in etwa zu wissen oder sogar erklären zu können, was man unter KI versteht. Zum Vergleich: in einer ähnlichen Abfrage der Initiative D21 in 2020 waren dies 61 Prozent. Der in unserer Abfrage höhere Wert spiegelt zum einen wider, dass hier alle Befragten im Arbeits-

leben stehen und zudem das Thema eine deutliche Dynamik aufweist: auch in der D21 Studie stieg der Wert dazu zwischen der Befragung 2018 und 2020 um 9 Prozent (Initiative D21 2020, S. 30). Mit 8 % hat GES den geringsten Anteil von Beschäftigten ohne Kenntnis von KI, in LOG ist dieser Anteil mit 11 % am größten und noch leicht vor DER (10 %).

### Sorgen und Befürchtungen

Auch die Einschätzung zur Abwägung von Nutzen und Risiko beim Einsatz von KI fällt eher positiv aus. Die beiden oberen *Balkengrafiken* in *Abbildung 3* zeigen die Antwort einmal für eine Abwägung in Bezug auf KI allgemein (links) und im Hinblick auf Beschäftigung (rechts). Es fällt dabei für alle Stichproben ins Auge: a) dass die Mehrheit jeweils Nutzen und Risiko als gleich einschätzt oder sogar einen höheren Nutzen erwartet und im Vergleich jeweils geringere Anteile ein höheres Risiko annimmt; b) ein höheres Risiko im Vergleich zum Nutzen wird jeweils für KI im Hinblick

auf Beschäftigung von mehr Befragten gesehen als für den allgemeinen Einsatz; c) jeweils die deutliche Mehrheit ist der Meinung, dass Nutzen und Risiken ungefähr gleich groß seien, was sowohl Ausdruck von Unsicherheit beim Thema wie Ausdruck der Hoffnung in gelingende und partizipativer Gestaltungsprozesse sein kann. Im Vergleich der drei Stichproben zeigt sich auch, dass die GES-Beschäftigten mit 23 % für generellen KI-Einsatz und 33 % für KI und Beschäftigung deutlich höhere Anteile einen Risikoüberhang erwarten als dies bei LOG und DER der Fall ist.

Auch die weiteren *sechs Balkengrafiken* der *Abbildung 3* zeigen, dass Ängstlichkeit nicht das vorherrschende Gefühl von Beschäftigten in Deutschland zu sein scheint, wenn es um KI am Arbeitsplatz geht. Bei den drei Balkengrafiken auf der *linken Seite*, geht es um die eigenen Fähigkeiten im Verhältnis zur KI (4-er-Skala zusammengefasst), *rechts* werden verschiedene Ängste abgefragt. Auch hier sprechen die Detailergebnisse für sich, es

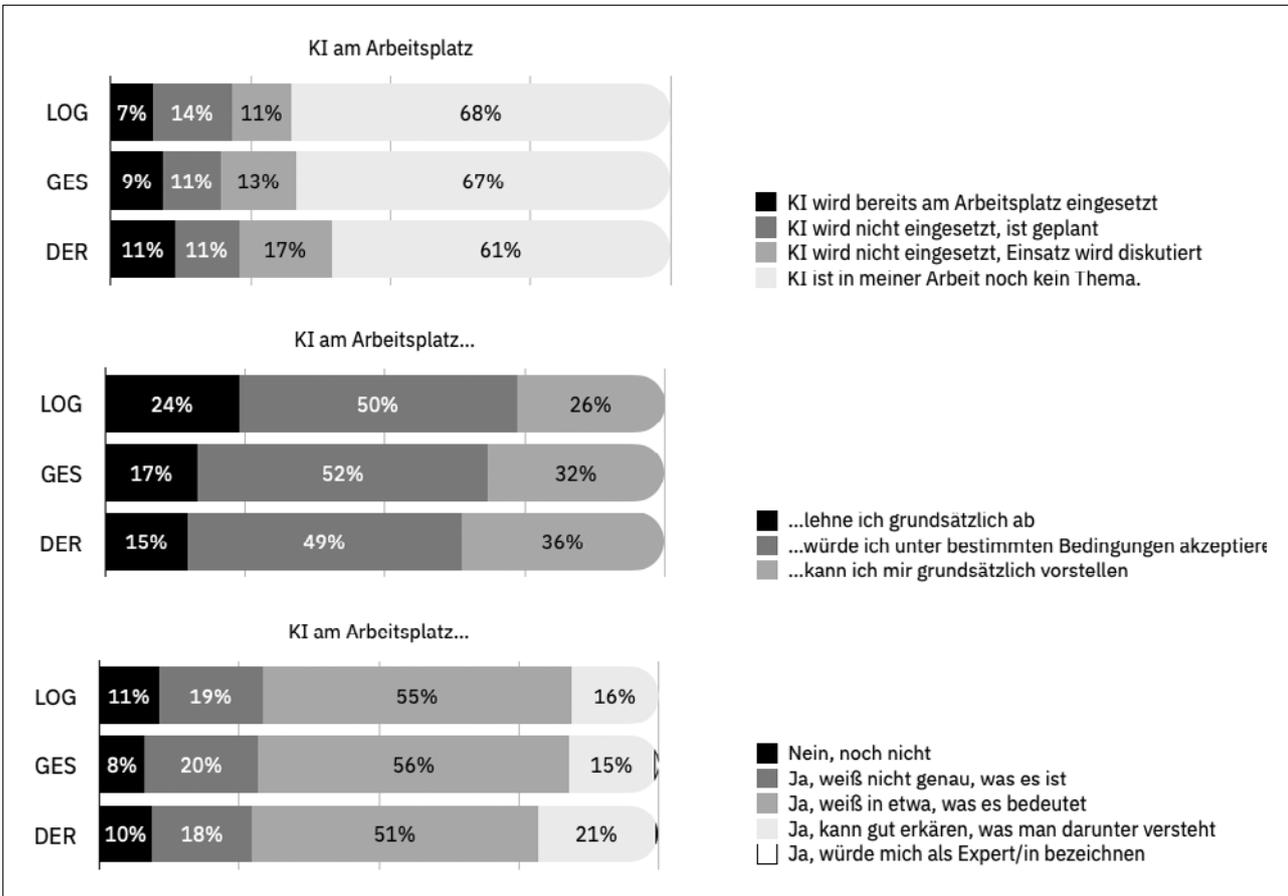


Abbildung 2: KI am Arbeitsplatz, grundsätzliche Einstellung, Wissensstand

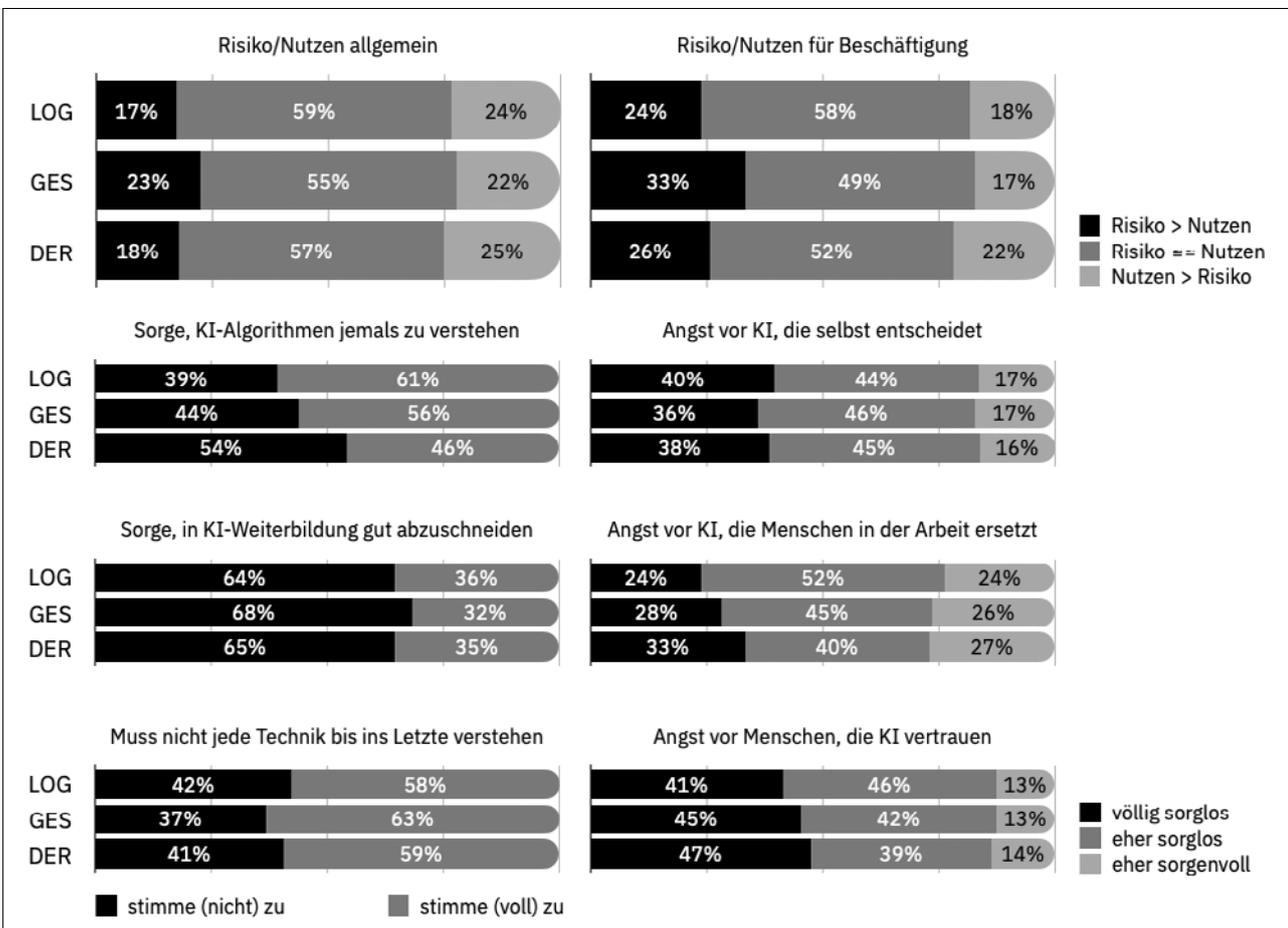


Abbildung 3: Sorgen und Befürchtungen

zeigt sich aber wiederum: Ängste und Sorgen finden sich weniger als der mediale Diskurs um KI nahelegen würde. Am wenigsten Sorgen machen sich die Befragten im Hinblick auf ihren Erfolg bei KI-bezogener Weiterbildung.

### Wie KI sein soll: proaktive Kollegin oder stille Assistentin?

Abbildung 4 zeigt einige Fragen der Erhebung, die sich mit der von Beschäftigten gewünschten Ausgestaltung der KI beschäftigen. Die vier oberen Balkengrafiken zeigen, in welcher Rolle sich Beschäftigte KI wünschen. Dabei ist die Einschätzung am deutlichsten, dass KI keine Chefin sein soll, dies lehnt LOG am eindeutigsten ab (89%), dicht gefolgt von GES mit 88%. Auch als Kollegin ist die KI kaum gewünscht, aber für Anteile zwischen 40% GES, 46% LOG und 50% DER vorstellbar. Jeweils zwischen „wäre ok“ und „geht gar nicht“ hält sich die Einschätzung jedoch fast die Waage. Die Zustimmungswerte wachsen eindeutig bei den eher untergeordneten, zuarbeitenden Rollen der KI als Beraterin oder Assistentin, wobei letzteres in allen drei Stichproben die am stärksten vorstellbare Rolle darstellt.

Abgefragt wurde auch, wie sichtbar bzw. aktiv die KI in ihrer Rolle am Arbeitsplatz werden soll (vier untere Balkengrafiken in Abbildung 4). Es dominiert auch hier jeweils das etwas uneindeutige und verhaltene „wäre ok“. Am eindeutigsten ist die Ablehnung von KI in Form eines sichtbaren Roboters oder Avatars: hier finden sich für alle drei Stichproben die größten „geht gar nicht“-Anteile. Die GES-Beschäftigten wünschen sich mit 23% am deutlichsten eine „stille“, im Hintergrund agierende KI. Bei

den drei Aktivitäts- und Sichtbarkeitsvarianten kristallisieren sich keine eindeutige Präferenzen heraus. Insgesamt zeigen diese Werte vor allem eines: KI sollte so gestaltet werden, dass die mit ihr umgehenden Beschäftigten die Rollenverteilung und Sichtbarkeit der KI individuell bestimmen und beeinflussen können – am besten nicht nur während der Implementierungs- und Lernphase, sondern auch im späteren Prozess.

### Was KI dürfen soll: Einstellungen zu Zugriffsdimensionen

Die aktuelle Digitalisierung eröffnet an vielen Stellen neue Formen der Steuerung und Kontrolle von Arbeit – bis auf die Ebene von Mikrohandeln und Vitalzeichen. So zeichnen manche kollaborative Roboter – aus ihrer Funktion heraus notwendig – Metadaten zu Körperbewegungen von Beschäftigten auf. Das gilt auch für das Aufzeichnen der Stimme im Call-Center oder die laufende Messung von Vitalzeichen durch einen smarten Arbeitshandschuh. Viele dieser Datenzugriffe sind an Arbeitsplätzen in Deutschland (noch?) verboten oder rechtlich stark eingeschränkt. All diese neuen, beschäftigtennahen Datenquellen aber werden und sollen auch für KI-gestützte Analysen geöffnet werden – so zumindest die Strategie und das Geschäftsmodell vieler aktuell aus dem Boden schießender Start-ups. Wie sehen das die Beschäftigten? Auch hierzu ein Auszug vieler darauf zielender Abfragen in unserer Studie, siehe Abbildung 5. Alle acht Balkengrafiken stellen Varianten von Zugriffsdimensionen vor.

Mit dieser Abfrage – das muss mitgedacht werden – ist aber nicht adressiert, was die aufgezeichneten Daten über den eigentlichen Zweck der Aufzeichnung hinaus potenziell ermög-

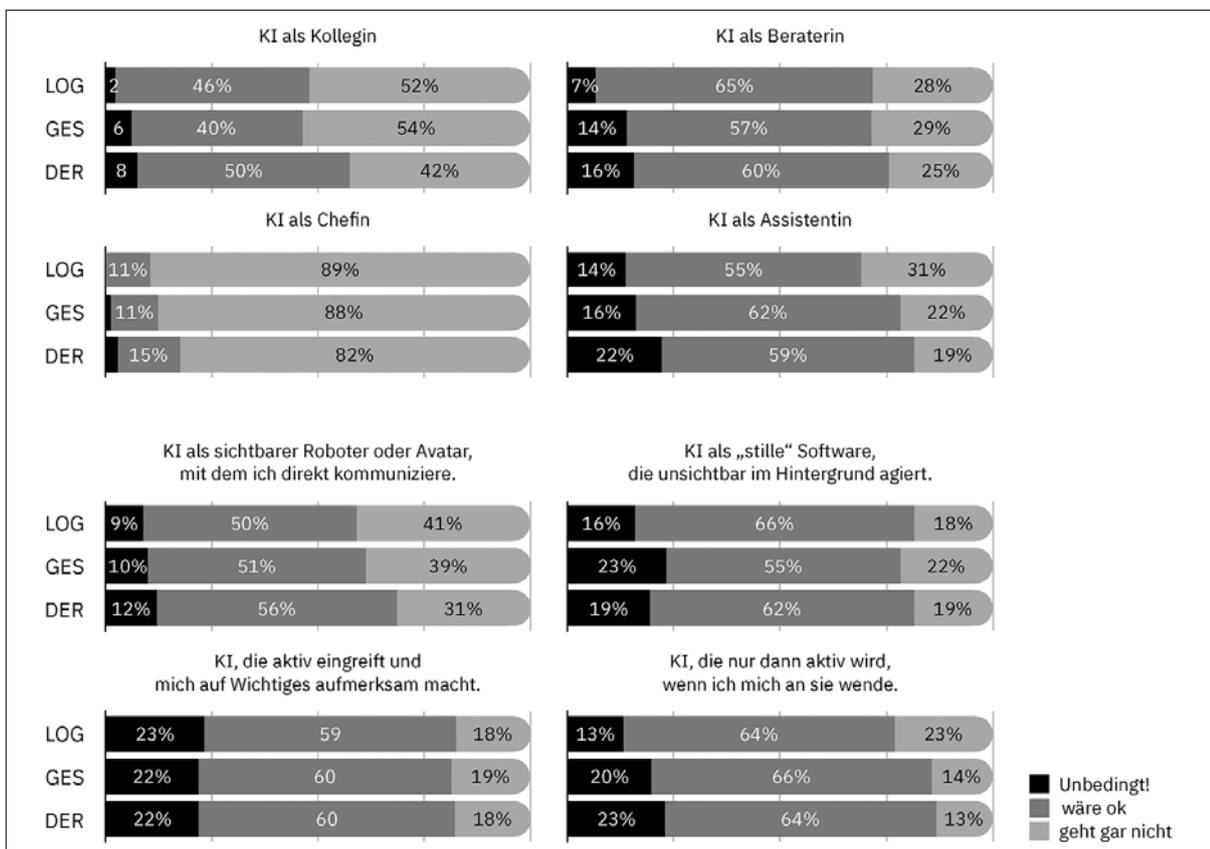


Abbildung 4: KI als Kollegin? Rollen- und Funktionserwartungen

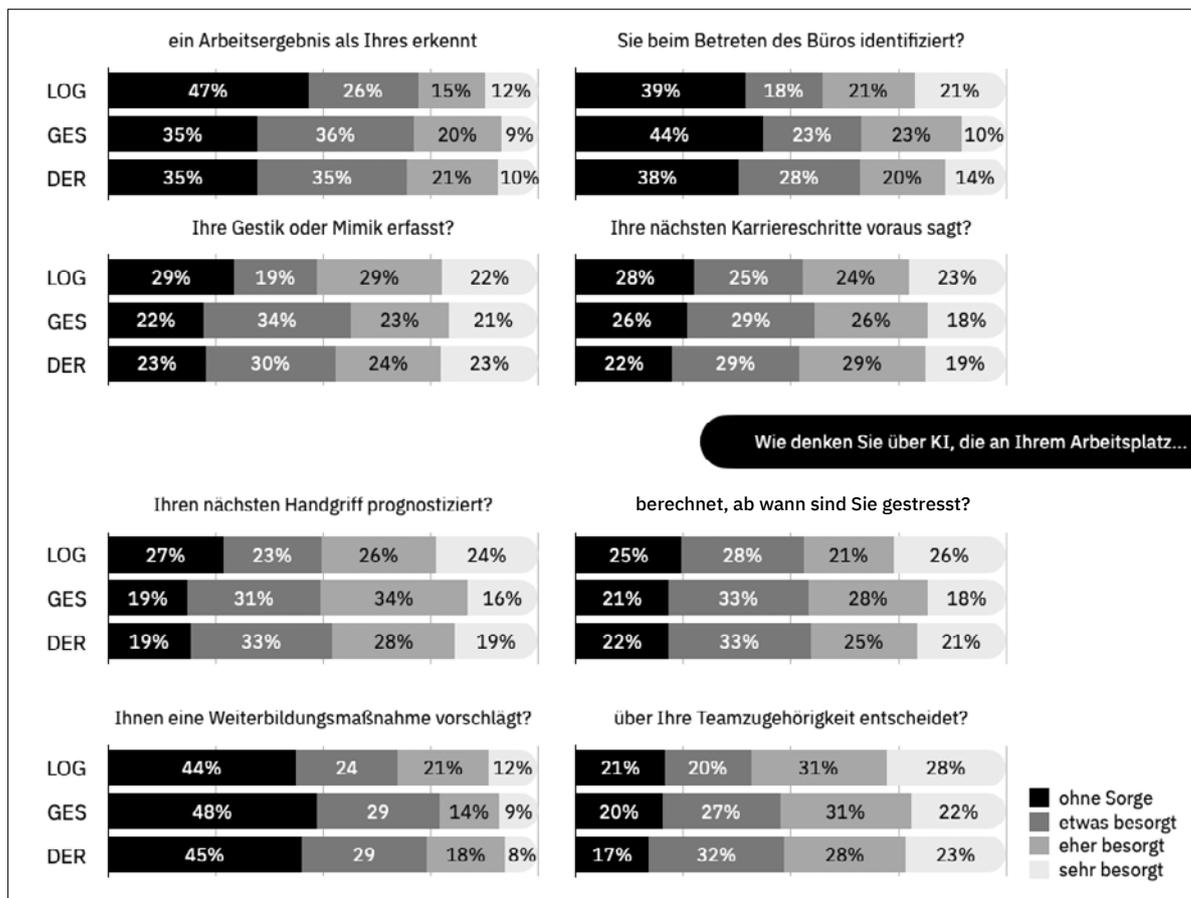


Abbildung 5: Einstellung zu Zugriffsoptionen der KI am eigenen Arbeitsplatz

lichen. Am entspanntesten, d. h. mit den höchsten Werten bei „ohne Sorge“, sind die Beschäftigten beim KI-gestützten Vorschlagen von Weiterbildungsmaßnahmen; bei der KI-Nutzung, die einen beim Betreten des Büros erkennt und beim Erkennen des eigenen Arbeitsergebnisses (mit hier den höchsten Anteilen von LOG mit 47 %). Alle anderen Items erhalten deutlich negativere Einschätzungen. Deutlich skeptischer zeigen sich die Befragten bei Einblicken in Mikrohandeln und Körper, bei der Teamzugehörigkeit und bei der eigenen Karriere-Vorausschau. Insgesamt überwiegen bei allen Stichproben und all diesen Themen jedoch deutlich die Sorgen. Weiter oben wurde bereits deutlich, dass die Beschäftigten keine ausgeprägten Ängste der KI selbst gegenüber haben. Wenn sich hier mehr Sorgenvolles findet, ist davon auszugehen, dass die bisherige Arbeitserfahrung dabei eine Rolle spielt und das damit verbundene Wissen darum, dass Technik wo immer möglich von Arbeitgeberseite gerne für die Kontrolle und Steuerung der Beschäftigten eingesetzt wird.

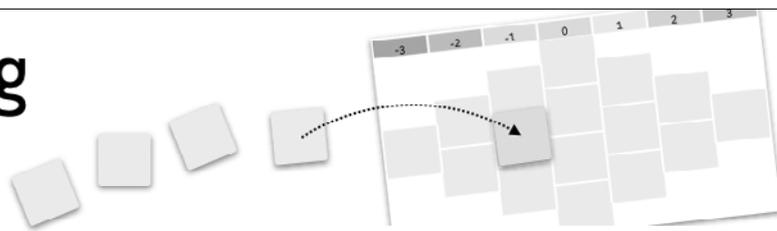
### Denkzeug: Komplexe, arbeitsbezogene Meinungsbilder

Einstellungen zu KI sind komplexer und widersprüchlicher als es klassische Erhebungsmethoden abbilden können. Daher wurde in den Fragebogen unserer Studie *KI als Kollegin* als besondere Erhebungsmethode das am Lehrstuhl entwickelte „Denkzeug“ integriert, eine Abwandlung der sogenannten Q-Methode (Brown 1980). Dieses Erhebungs- und Analyseverfahren eignet sich für eine fast qualitative Erfassung menschlicher Subjektivität innerhalb einer quantitativen Erhebung. In Zusammenarbeit mit

SoSci eine ältere HTMLQ-Applikation (Banasick 2022) technisch und methodisch aktualisiert und responsiv auf der Basis von Angular neu gestaltet. In einer quasi-haptischen Weise können die User:innen die von uns inhaltlich entwickelten 16 Karten zu acht Aussagedimensionen (Q-set; Müller und Kals 2004) in ein vorgegebenes Grid sortieren. Damit konnten sie ihre Vorstellung zu KI am eigenen Arbeitsplatz in ein komplexes Meinungsbild legen und zu den beiden Extremkarten am Ende noch Kommentare in ein freies Textfeld abgeben (was über 300 der Befragten überwiegend mit sehr ausführlichen Texten auch getan haben). *Abbildung 6* zeigt eine stilisierte Form des Erhebungsinstrumentes und die 16 Aussagekarten. Die Ausgangsfrage lautete (verkürzt): „Stellen Sie sich bitte vor, KI kommt an Ihrem Arbeitsplatz an und spielt bei Ihrer alltäglichen Arbeit eine Rolle. (...) Wie wünschen Sie sich eine solche Zusammenarbeit? Wie sollte es sein? Und was sollte nicht oder auf keinen Fall passieren?“.

Auf Basis der gelegten Meinungsbilder konnten (mit *PCA* und *Varimax Rotation*) vier Einstellungstypen identifiziert werden: Die Störungsjonglierenden, die Arbeitsfokussierten, die Verantwortungstragenden und die Entlastungsuchenden. Für diese vier Typen zeigt sich trotz aller Unterschiede: Die Ansichten zu KI als Kollegin sind stärker vom Arbeitskontext geprägt als von allgemeinen Vorstellungen zu Technik und KI. Selbst andere Merkmale wie Geschlecht, Alter oder Qualifikation spielen demgegenüber eine untergeordnete Rolle. *Abbildung 7* im unteren Bereich skizziert die vier Typen inhaltlich und weist Anteile und varianz-erklärende Anteile für die Gesamtstichprobe aus. Die *Balkengrafik oben* zeigt die Anteile für die hier verwendeten Stichproben: Demnach finden sich in LOG mit 35 % überdurchschnittlich viele

# DENKzeug



|                               |                         |   |        |   |                               |
|-------------------------------|-------------------------|---|--------|---|-------------------------------|
| <b>Entscheidungshoheit</b>    | beim Mensch             | Die letzte Entscheidung liegt immer bei mir, nie bei der KI. Das wäre ja noch schöner!  | ←<br>→ | Die meisten Entscheidungen nimmt mir die KI ab. Sie fragt mich nur dann, wenn die Situation unklar oder kritisch ist.                     | bei der KI                    |
| <b>Entscheidungskriterien</b> | nachvollziehbar         | Warum die KI genau zu einer Entscheidung kommt, das muss nachvollziehbar bleiben. Das ist bei meiner Arbeit ja auch nicht anders.             | ←<br>→ | Wie und warum die KI so oder so entscheidet, das muss ich gar nicht wissen. Dann bin ich auch nicht dafür verantwortlich.                 | Blackbox                      |
| <b>Lernfähigkeit</b>          | des Menschen            | Wenn doch mal was schief läuft, muss ja doch ich ran. Deshalb bleibt wichtig, dass ich immer wieder Neues lernen kann – trotz KI.             | ←<br>→ | In der Arbeit ändert sich ja ständig etwas, deswegen bleibt die KI in der Lage, immer weiter zu lernen.                                   | des Algorithmus               |
| <b>Entscheidungsprozess</b>   | beeinflussbar           | Bei der Arbeit geht es schließlich um etwas. Daher kann ich die KI jederzeit stoppen oder ihre Entscheidungen wieder zurücknehmen.            | ←<br>→ | Damit sich das Ganze auch lohnt, arbeitet die KI möglichst viel bis zum Ende völlig automatisiert ab, ohne dass ich eingreifen muss.      | automatisiert                 |
| <b>Aufgabenlast</b>           | beim Mensch             | Weil die KI mir die Routinesachen abnimmt, kann ich mich darauf konzentrieren, wo es hakt, ganz dringend ist oder „brennt“.                   | ←<br>→ | Meine Arbeit ist viel weniger anstrengend, weil mir die KI sehr viel abnimmt und so den Rücken freihält.                                  | bei der KI                    |
| <b>Arbeitsanforderung</b>     | wird anspruchsvoller    | Die langweiligen und lästigen Aufgaben übernimmt die KI. So habe ich Zeit für meine anspruchsvollen und wichtigen Aufgaben.                   | ←<br>→ | Die KI nimmt mir die komplexen und unübersichtlichen Aufgaben ab. Dafür ist sie da. Damit bei mir nicht nur das Schwierige übrig bleibt.  | wird einfacher                |
| <b>Kontextwissen</b>          | relevant für Einführung | So eine KI ist auch nur Software. Damit da was Ordentliches bei rauskommt, bin ich bei der Einführung von Anfang an dabei.                    | ←<br>→ | KI ist was für Experten. Da halte ich mich raus. Die Technik soll sauber funktionieren, wenn sie bei mir am Arbeitsplatz eingesetzt wird. | nicht relevant für Einführung |
| <b>Verantwortung</b>          | bei Mensch/Anwender*in  | Bei KI ist das auch nicht anders als wenn ich andere Technik einsetze: am Ende stehe nur ich gerade für meine Arbeit und das ist auch gut so. | ←<br>→ | Wenn meine KI Quatsch macht, dann bin nicht ich schuld. Da muss dann der Programmierer ran oder die Softwarefirma.                        | bei KI/Hersteller             |

Abbildung 6: Q-set – Aussagen zu KI des Denkzeugs



Abbildung 7: KI-Einstellungstypen – sozio-demografische Merkmale

Störungsjonglierende, in GES sind es mit 26 % weniger (auch als in DER). LOG und GES haben mit je 33 % mehr Arbeitsfokussierte als der Rest der Beschäftigten in Deutschland. Am wenigsten variieren über die Stichproben die Anteile der Verantwortungstragenden; auf den ersten Blick erstaunlich finden sich in LOG deutlich weniger Entlastungsuchende. Da gerade bei diesem Typ der Anteil der Akademiker:innen höher ist, bilden sich hier eventuell auch stark die für diese Beschäftigtengruppe typischeren Belastungsformen ab, die sich im von anderen Stressfaktoren geprägten Logistik-Alltag weniger finden. Oder aber: Logistikbeschäftigte sind stressresistenter als Akademiker:innen in allen Branchen. Das lässt sich mit den hier vorgestellten Daten und Analysen nicht sagen.

## KI am Arbeitsplatz: Eine Frage der Gestaltung – und gestaltungswillige Beschäftigte

Diese komprimierte Darstellung einiger zentraler Ergebnisse unserer repräsentativen Beschäftigungsbefragung zu *KI als Kollegin* und mit dem gewünschten Fokus auf Logistik und Gesundheitswesen zeigt: KI hat erst begonnen, in der Arbeitswelt anzukommen. Beschäftigte zeigen sich im Hinblick auf KI nicht so ängstlich und sorgenvoll, wie üblicherweise unterstellt wird, sie wünschen sich als Rolle der KI am Arbeitsplatz nur bedingt KI als Kollegin, sondern eher als dienstbare und im Hintergrund still wirkende Assistentin. Ein zentraler und besonders interaktiver Baustein der Erhebung – das so genannte „Denkzeug“ – zeigt auf der Basis multivarianter Analysen, dass Einstellungsmuster zu KI am Arbeitsplatz nicht nur komplex sind, sondern dabei die Perspektive aus der eigenen Arbeit maßgeblicher ist als allgemeine Einstellungen zu Technik und KI. Es finden sich im Detail durchaus Unterschiede zwischen den drei präsentierten Stichproben Logistik (LOG), Gesundheitswesen (GES) und dem Rest der Beschäftigten in Deutschland (DER) – allerdings nicht durchgängig und selten dramatisch.

Bei vielen der dargestellten Ergebnisse dominieren die uneindeutigen „es kommt darauf an“ Antworten. Das deutet auch aus Beschäftigtenperspektive auf ein hohes Interesse an Mitbestimmung hin. Auch das haben wir abgefragt: die überwiegende Mehrheit ist der Meinung, dass Mitbestimmung bei KI am Arbeitsplatz wichtiger wird als bisher – unsere beiden im Zentrum stehenden Stichproben sehen das ganz besonders so (LOG 77 %, GES 76 %, DER 71 %; ohne Abbildung).

Die für diesen Text ausgewählten Ergebnisse sind vor dem Hintergrund zu sehen: Die Mehrheit der Beschäftigten erlebt KI noch nicht an ihrem Arbeitsplatz und viele Beschäftigte fühlen

sich zwar informiert zum Thema KI, haben evtl. aber trotzdem noch keine ausreichenden Vorstellungen über die Arbeitsweise von KI. Daher sind viele der erfragten Einstellungen und Erfahrungen zu KI als Kollegin heute noch nicht erfahrungsgesättigt. Je nachdem, welche Erfahrungen in der Nutzung von KI in den kommenden Jahren gemacht werden, werden sich Einstellungen dazu verändern, ausdifferenzieren oder auch verfestigen. Deswegen kann diese Befragung nur eine Momentaufnahme sein.

## Referenzen

- Banasick, Shawn. 2022. EQ Web Sort. <https://github.com/shawnbanasick/easy-htmlq>. Zugegriffen: 23. Juli 2022.
- Brown, Steven R. 1980. *Political Subjectivity – Applications of Q Methodology in Political Science*. New Haven, CT: Yale University Press.
- Dosenovic, Pero, Kimon Kieslich, und Birte Keller. 2020. *Meinungsmonitor Künstliche Intelligenz – Methodensteckbrief*. Düsseldorf: Heinrich-Heine-Universität. [https://uni-duesseldorf.datalion.net/custom/img/16/200629\\_Methodensteckbrief%20MeMoKI.pdf](https://uni-duesseldorf.datalion.net/custom/img/16/200629_Methodensteckbrief%20MeMoKI.pdf).
- Initiative D21. 2020. *Wie digital ist Deutschland 19/20. Jährliches Lagebild zur Digitalen Gesellschaft*. Kantar.
- Leiner, Dominik Johannes. 2019. SoSci Survey (Version 3.1.06) [Computer software]. <https://www.socisurvey.de>.
- Liang, Yuhua, und Seungcheol Austin Lee. 2017. Fear of Autonomous Robots and Artificial Intelligence: Evidence from National Representative Data with Probability Sampling. *International Journal of Social Robotics* 9: 379–384.
- Müller, Florian H., und Elisabeth Kals. 2004. Die Q-Methode. Ein innovatives Verfahren zur Erhebung subjektiver Einstellungen und Meinungen. *Forum Qualitative Sozialforschung* 5: Art. 34.
- Pfeiffer, Sabine. 2021. *Datenmanagementkonzept zur Erhebung und Verarbeitung personenbezogener und/oder sensibler Daten für Forschungszwecke im Rahmen qualitativer und quantitativer Erhebungen. v1.0 vom 6.11.2021*. Nürnberg: Lehrstuhl für Soziologie (Technik – Arbeit – Gesellschaft) am NCT der FAU Erlangen-Nürnberg. <https://www.labouratory.de/files/downloads/WP-04-2021-Datenmanagementkonzept.pdf>.
- Schnell, Rainer. 2018. *Warum ausschließlich 'online' durchgeführte Bevölkerungsumfragen nicht 'repräsentativ' sind*. Duisburg, Essen: Universität Duisburg-Essen. [https://www.uni-due.de/imperia/md/content/soziologie/schnell/repraesentativ\\_online.pdf](https://www.uni-due.de/imperia/md/content/soziologie/schnell/repraesentativ_online.pdf).

## Anmerkung

- 1 *Im mathematischen Sinn ist Repräsentativität bei einem Online-Panel nicht zu erreichen, da dies eine echte Zufallsstichprobe erfordern würde, an der jede Person der Grundgesamtheit die gleiche Chance hat, ausgewählt zu werden – was für die meisten Befragungswege nicht in Gänze garantiert werden kann (vgl. dazu kritisch Schnell 2018).*



**Sabine Pfeiffer**



**Sabine Pfeiffer**, geboren 1966 in Unna, ist Professorin für die Soziologie mit dem Schwerpunkt Arbeit – Technik – Gesellschaft an der Friedrich-Alexander-Universität Erlangen-Nürnberg. 2021 erschien ihr Buch *Digitalisierung als Distributivkraft. Über das Neue am digitalen Kapitalismus*; 2022 in englisch unter dem Titel: *Digital Capitalism and Distributive Forces*; in OpenAccess bei transcript. Foto: [www.thomasriese.com](http://www.thomasriese.com)

## Was bleibt von der Arbeitswelt, wie wir sie kennen?

### Ein Praxisbericht aus Sicht der Beschäftigten und der Interessenvertretungen

*Dieser Artikel betrachtet die durch die Digitalisierung und die damit verbundenen Veränderungen in der Arbeitswelt verursachte aktuelle Situation in den Unternehmen aus der Perspektive von Interessenvertretungen und Beschäftigten. Er behandelt einerseits, wie sich die Arbeit durch Digitalisierung und die damit möglich gewordenen neuen Arbeitsformen verändert, und andererseits, wie sich die damit einhergehende massive Ausweitung der Verarbeitung von Beschäftigtendaten auswirkt. Während die Veröffentlichungen der Wirtschaftsverbände, Wissenschaft, Politik, Medien und Anbieter von Digitalisierungslösungen vor allem den Nutzen der Digitalisierung für Unternehmen wie Beschäftigte in helles Licht rücken, geht es hier um den Schatten.*

#### Digitalisierung und Kollaboration sind super, wenn ...

Die Digitalisierung hat nicht erst mit dem *Internet der Dinge* oder *Industrie 4.0* begonnen. *Digitale* Werkzeuge gibt es in der Arbeitswelt seit etwa Mitte letzten Jahrhunderts. Nach den ersten Automatisierungswellen von Teilen der Büroarbeit mittels Großrechnern ab 1950 oder in den 70ern in der Produktion durch die Einführung von Robotern, beschleunigte sich die Digitalisierung in den letzten Jahren massiv. Vernetzung, Cloud-Technologien und die Möglichkeiten der Künstlichen Intelligenz (KI) sind nur einige Beispiele. Die Corona-Pandemie Anfang 2020 war ein wichtiger Katalysator. Sie hat die Einführung von digitalen Kollaborations-Werkzeugen erzwungen und so eine Unternehmensübergreifende Zusammenarbeit unter Einhaltung der gebotenen Quarantäne-Bedingungen möglich gemacht. Beschäftigte, Betriebsräte und das Management haben gleichermaßen den Nutzen von mobiler Arbeit und Homeoffice erkannt und schätzen gelernt. Der Zugriff auf Daten von jedem Ort, die Möglichkeit mittels Videokonferenzsystemen zu kommunizieren oder gar gemeinsam in Echtzeit an Dokumenten oder virtuellen Flipcharts zu arbeiten, haben überzeugt. Es stellt sich allerdings die Frage, welche Auswirkungen diese Optionen auf die Arbeitsgestaltung, Arbeitsabläufe, Arbeitsergebnisse, Aufgabensteuerung und Ergebniskontrolle haben. Die Digitalisierung erfüllt vor allem Wünsche der Arbeitgeber nach Transparenz der Arbeit und damit die Möglichkeit der Arbeitskontrolle. So manche Führungskraft kompensiert die fehlenden Möglichkeiten zur Führung im Homeoffice und ihre Überwachungsinteressen mittels direktem Zugriff auf abgelegte Arbeitsergebnisse oder gar in Echtzeit in den Arbeitsvorgang hinein, wie das z. B. bei Office-Dokumenten möglich ist. Beschäftigte wie Arbeitgeber freuen sich allerdings über die flexiblen Arbeitszeiten. Zwei Seiten der Medaille: *ständige Erreichbarkeit* versus *flexible Arbeitszeit*, die insbesondere jüngeren Beschäftigten entgegenkommt. Im Arbeitseifer ist damit allerdings oft eine Verletzung des Arbeitszeitgesetzes oder eine Überschreitung der vereinbarten Arbeitszeiten oder -menen verbunden. Die damit verbundenen Rechtsverstöße bleiben ungeahndet oder werden sogar gegen die Beschäftigten verwendet, und langfristig wirkt sich das auch noch negativ auf deren Gesundheit aus.

Flexibilisierung vs. Entgrenzung oder Transparenz vs. Kontrolle – wie sieht der Ausweg aus den Dilemmata der Digitalisierung aus? Wie lässt sich der Wunsch der Beschäftigten nach einer gesundheitsförderlichen Arbeit erfüllen? Wie ist *New Work* zu gestalten? Die Antwort finden Arbeitgeber, Beschäftigte und Inte-

ressenvertretungen vermutlich nur gemeinsam, im besten Fall mit Unterstützung der Politik und der Gesetzgebung. Aktuell verfolgen die verschiedenen Interessengruppen wohl noch unterschiedliche Wege.

Bisher scheinen die Arbeitgeber die Digitalisierung besser für sich und für eine Machtverschiebung zu ihren Gunsten nutzen zu können. In der Praxis möchten sie die Gestaltungshoheit für die neuen Technologien anscheinend nicht teilen. Da die Digitalisierung und KI *Stand der Technik* sind, handeln viele Arbeitgeber nach dem Motto: „Es wird digitalisiert, was geht“. *„Digitalisierung first, Bedenken second“*, 2017 ein Wahlkampflogan der FDP, scheint so manchen Arbeitgeber inspiriert zu haben. Flankierend fordern Arbeitgeberverbände (BDA), die gesetzlichen Regelungen der Mitbestimmung, des Arbeitsschutzes und zur Arbeitszeit deutlich aufzuweichen. Das ist bereits 2015 in den Forderungen des BDA zu *Arbeitswelt 4.0 – Chancen nutzen, Herausforderungen meistern* sowie aktuell in dem Papier des BDA-Digitalrats *Zwei Welten, ein Rechtsrahmen* nachzulesen.

Den Interessenvertretungen ist klar, dass die klassische statische Betriebsvereinbarung mit vollständiger Dokumentation zu einer Datenverarbeitung heute, bei monatlichen Softwareupdates, nicht mehr die Lösung sein kann. Weil Digitalisierung ohne Mitbestimmung auch keine Lösung liefern kann, lassen sich gute Arbeitsbedingungen und Arbeitsergebnisse nur mit dem gemeinsamen Willen für kreative betriebliche Regelungen sichern, die der Digitalisierung und den Schutzbedürfnissen der Beschäftigten gleichermaßen gerecht werden. Daher muss auf der politischen, der tariflichen und der betrieblichen Ebene die Beteiligung der Beschäftigten und die Mitbestimmung durch Interessenvertretungen gestärkt und modernisiert werden. Die Sozialpartner, also Gewerkschaften und Arbeitgeberverbände, haben daher bereits einzelne Tarifverträge zur Digitalisierung abgeschlossen und sind in ständiger Diskussion, um die Machtverhältnisse auszugleichen. Die Gesetzgeber arbeiten auf europäischer wie nationaler Ebene an verschiedenen Gesetzen, Richtlinien und Verordnungen.

*Miteinander reden hilft*, kann der Ansatz sein, die verschiedenen Interessen der Akteure einzubeziehen. Ziele vieler Digitalisierungsprojekte sind eine höhere Effizienz, Transparenz oder Flexibilität. In den Ohren von Beschäftigten klingt das nach Arbeitsverdichtung, Überwachung oder Entgrenzung. Zu Gunsten des Erfolgs sollten die Ziele eines Digitalisierungsprojekts verhandelt und gemeinsam bestimmt werden.

## Transparenz im Betrieb oder Überwachung der Beschäftigten

Die schnelle und erfolgreiche Einführung von Werkzeugen zur digitalen Kollaboration, zur Überwindung der Folgen eines Corona-Lockdowns, haben den Wunsch nach mehr Digitalisierung befördert. Doch *Wir digitalisieren alles, was geht!* ist keine zielgerichtete Strategie. Wird erst digitalisiert und danach überlegt, welcher Nutzen damit verbunden ist, bleiben die echten Probleme unbehandelt. Vielmehr macht sich Unfrieden breit über mangelnden Datenschutz oder vernachlässigte Mitbestimmung. Signifikantes Beispiel dafür ist, dass heute Arbeitsgeräte (Handscanner, Mobiltelefone, Tablets, PCs u. v. m.) nicht genutzt werden können, ohne dass sich die nutzende Person namentlich am Werkzeug anmeldet, natürlich begleitet von einem Protokoll über alle An- und Abmeldeprozesse. Diese Entwicklung kennen wir schon von Softwaresystemen, Maschinen und Leitständen in der Produktion, nun kommen Elektrowerkzeuge, Gabelstapler oder Fahrzeuge noch dazu. Die individuelle Arbeitsleistung und das Arbeitsverhalten mit diesen Arbeitsgeräten wird so zu 100 % transparent und auswertbar. Solche Datensammlungen werden gern oberflächlich begründet: etwa als Diebstahlschutz, zur Optimierung der Abläufe oder aus versicherungstechnischen Gründen, da nur berechnete Personen eben bestimmte technische Geräte nutzen dürfen.

Unter der Überschrift *Optimierung und Qualitätssicherung* werden mittels Sensoren zusätzlich die Produktionsschritte oder Dienstleistungen über den gesamten Arbeitsprozess hinweg immer kleinteiliger protokolliert. Die Kombination aus den Daten der Sensorik mit den Login-Daten der einzelnen Beschäftigten macht aus einer Produktionsüberwachung eine Überwachung der Beschäftigten. Der gläserne Beschäftigte wird noch gläserner.

Die Interessenvertretung muss in derartigen Konstellationen oft viel Energie aufbringen, um die Datenerhebung auf die Erforderlichkeit und die Nutzung auf den formulierten Zweck zu begrenzen (das *Need-to-know*-Prinzip). In Europa haben wir dazu zwei gesetzlich verankerte Hebel, das Prinzip der Datensparsamkeit im Rahmen der Datenschutznormen und die Mitbestimmungsrechte. Bei beiden geht es um den Ausgleich der Machtverhältnisse, in unserem Kontext zwischen Arbeitgebern und Beschäftigten. Das heißt, die Spreu vom Weizen zu trennen: Telematiksysteme auf Lkws liefern oft Verhaltensdaten des Fahrers, Geodaten zur Route, Temperaturverläufe zum Motor und Spritverbrauch sekundengenau. Eine vollständige Transparenz der Tätigkeiten oder des Verhaltens von Beschäftigten im Zugriff für den Arbeitgeber ist mit der Wahrung der Persönlichkeitsrechte und Schutzbedürfnisse nicht vereinbar. Somit ist es

erforderlich mit dem Arbeitgeber, auch gegen dessen Widerstand, die konkrete Zweckbestimmung zur Verarbeitung derartiger Daten zu vereinbaren, etwa „... ausschließlich zur zeitbezogenen Reisekostenabrechnung“, dies rechtfertigt die Erfassung von Start- und Endzeitpunkt einer Fahrt, nicht jedoch die Erhebung oder die Auswertung von Fahrverhalten oder kleinteiligen Geokoordinaten.

## Ist die Künstliche Intelligenz in den Betrieben schon angekommen?

Zur Digitalisierung gehörten bereits seit geraumer Zeit eher abstrakt auch Technologien der Künstlichen Intelligenz, aktuell werden sehr konkrete Anwendungen wie ChatGPT oder Datenanalysen mittels maschinellem Lernen eingebunden. Bisher waren KI-Anwendungen sowohl für die Auftraggeber als auch die Beschäftigten und ihre Interessenvertretungen kaum sichtbar, weil eine KI meist nicht eigenständig als KI eingeführt wurde sondern als punktuell wirkendes Werkzeug oder eine Hilfsfunktion in einem umfassenden Produktions- oder Dienstleistungssystem. So existierten KI-Komponenten als Teile einer Gesamtlösung, z. B. zur Spracherkennung und -übersetzung, steuernde Sensoren oder Hilfsmittel, um Anomalien zu identifizieren. Heute wird KI dagegen gezielt eingesetzt, um etwa bei der Beurteilung von Kandidat:innen in einem Bewerbungsverfahren zu unterstützen, durch Analysen der Stimmmodulation, Körpersprache, Hautausdünstungen und Sprachkompetenz. Die Unterstützung liegt etwa im Vergleich vorliegender Daten zu tausenden vergleichbarer Datensätze, die in neuronalen Netze strukturiert und bewertet sind. Für die in diesem Datenbestand erfassten Personen ist die Zukunft als Beschäftigte bereits bekannt. Fehler, die andere Arbeitgeber gemacht haben, sollen so vermieden werden. Für die Interessenvertretung wird die Mitbestimmung hier erschwert, da eine KI-Lösung oft gar keine Leistungs- und Verhaltenskontrolle ermöglicht oder beabsichtigt, so wie es das BetrVG erwartet. Eher geht es um Themen wie Logistikoptimierung, Prozessvereinfachung, Automatisierung oder das Ersetzen von menschlicher Arbeitsleistung. Darüber hinaus sind diese Datenbanken bereits strukturell anonym. Vor diesem Hintergrund ist das klassische Vorgehen nach § 87 (1) 6<sup>1</sup>, bei IT-Einführungen nicht erschöpfend auf ein KI-System anwendbar.

Anfang 2023 änderte sich die Wahrnehmung in der Gesellschaft und damit auch in den Betrieben deutlich. Durch kostenlose Angebote wie *ChatGPT* (Textgenerator) oder *Craiyon* (Bildgenerator), ist es niederschwellig möglich, KI-Lösungen kennen zu lernen. Jeder kann diese Produkte nutzen und Ideen für ihre Verwendung im Betrieb entwickeln.



### Frank Steinwender

**Frank Steinwender**, Dipl.-Ing. (FH), begann seinen Berufsweg und damit auch seine gewerkschaftlichen Aktivitäten mit der Ausbildung zum Betriebsschlosser. Nach dem Studium arbeitete er unter anderem als selbständiger IT-Dienstleister im Bereich Netzwerktechnik, Logistik und IT. Seit 2011 berät er Interessensvertretungen bei der Technologieberatungsstelle des DGB NRW e. V. in Fragen des Beschäftigtendatenschutzes und im Rahmen der Mitbestimmung bei IT Projekten.

Nun haben die beiden großen Systemanbieter SAP und Microsoft Allianzen mit KI-Anbietern geschlossen und angekündigt, die Integration von KI-Lösungen in ihre klassischen Produkte intensiv voranzutreiben. Bisher haben sich die Interessenvertretungen mangels konkreter KI-Projekte kaum mit konkreten Regelungen und der Gestaltung von KI-Anwendungen befasst. Das wird sich in naher Zukunft deutlich ändern müssen. Angekündigt oder bereits vorhanden sind KI-Komponenten für Personalsysteme, Produktions- und unternehmensweite ERP-Systeme, bei Kollaborationswerkzeugen sowie bei IT-Sicherheitssystemen, zur Erstellung von Texten, auch wissenschaftlichen Abhandlungen, zur Automatisierung der Produktionsplanung oder von Bewerbungs- und Einstellungsverfahren.

### Ohne Login geht es nicht ... oder doch?

Die Zeiten scheinen vorbei zu sein, in denen sich ein Mobilfunkgerät, eine Drehbank, ein Drucker, eine Eingangstür oder sogar ein Computer anonym oder mit generischen Benutzerkonten nutzen ließen. Ein praktisches Beispiel ist die Digitalisierung von Schließsystemen/-anlagen. Das alte Prinzip, mit einem Schlüssel eine Zutrittsberechtigung zu vergeben, ist um den Aspekt der Speicherung und Kontrolle der Schlüsselbenutzung erweitert worden. Aus Mitbestimmungssicht ist damit aus der technischen Funktion, Zutritt zu gewähren, eine Überwachungseinrichtung geworden. Die ursprüngliche Ankündigung des Arbeitgebers, man wolle lediglich ein elektronisches Schließsystem einführen, es würde sich an den Verfahren nichts ändern, war schnell überholt. Bald wurde deutlich, dass nicht nur die Berechtigung zum Zutritt geprüft, sondern bei der Verwendung der Schlüssel ein Datensatz mit Schlüsselnummer, Uhrzeit und Ort gespeichert wird. In Verbindung mit der Liste, wem welcher Schlüssel ausgehändigt wurde, ist die namentliche Zuordnung möglich. Die Information, wer wann welche Tür genutzt hat, bedeutet für die Beschäftigten Überwachungsdruck. Es fragt sich, ob dieses System überhaupt geeignet ist, den Zweck des Zutritts zu kontrollieren oder zu helfen, einen Diebstahl aufzuklären. Wahrscheinlich sind dazu sogar Drehkreuze besser geeignet, da sie erlauben, die Personen zu zählen, die einen Raum betreten. Eine wenigstens datenschutzfreundliche Lösung wäre die Vergabe von elektronischen Schlüsseln ohne Personenbezug.

Die Forderung, sich immer und überall authentifizieren zu müssen, scheint Normalität geworden zu sein. Die Vorstellung, man könne eine Maschine einfach nur nutzen, um die Arbeit durchzuführen, oder eine Schraube an einem Bauteil festziehen, ohne den Namen, Zeitpunkt, Ort und Drehmoment zu hinterlassen, wird teilweise als absurd dargestellt. „*Lasst mal die Kirche im Dorf!*“ Das neue Normal ist, dass sich an allen Systemen alle immer anmelden müssen, im ergonomischen Falle über *Single-sign-on* oder biometrische Merkmale. Eine Schlüsselkarte ohne Namensbezug ist ein Beispiel für eine zwischen den Interessen abgewogene Lösung. Sie entspräche dem Need to know-Prinzip und der Forderung nach Datensparsamkeit.

In den Digitalisierungs-Projekten wird die Vorstellung, sich mit Mitbestimmung und Datenschutz auseinandersetzen zu müssen, allzu oft als überholt angesehen. Die Frage: *Was machen Sie konkret mit den gespeicherten Daten?* wird nicht oder nur

vage mit Begründungen wie der Qualitätssicherung beantwortet. Damit bleibt die Verarbeitung unzureichend begründet. Die Frage nach der Verhältnismäßigkeit und den Persönlichkeitsrechten der Beschäftigten stellt sich gar nicht erst. Viel häufiger sollen alle Tätigkeiten bzw. Leistungen personenbezogen detailliert erfasst und verarbeitet werden. Sicher ist die Erfassung von Teilen, deren Bearbeitungsfortschritt oder Standort sehr sinnvoll für die Optimierung von Durchlaufzeiten und für eine Echtzeitsteuerung oder eine schnelle Fehlererkennung. Ob aber ein einzelner Bearbeitungsschritt mit den personenbezogenen Daten von Beschäftigten verbunden werden muss, ist im Einzelfall zu begründen.

Es gibt auch positive Beispiele. In diesen Betrieben wird zwar die gesamte Produktion mit Sensorik ausgestattet, jeder Messwert genommen und gespeichert, allerdings ohne Verknüpfung zu den Menschen, die am Produktionsprozess beteiligt sind. Diese sehr großen anonymen Datenmengen werden mittels KI analysiert und Optimierungen abgeleitet, etwa für die Produktionssteuerung, Instandhaltung, Ablaufsteuerung, Verarbeitungsverfahren, Qualitätsmerkmale und Energieverbrauch. Trotz eines hohen Grads der Digitalisierung schränkt das die Persönlichkeitsrechte der Beschäftigten nicht ein. Digitalisierung kann also ohne *gläserne Beschäftigte* gut funktionieren.

### Informations-Sicherheit versus Datenschutz und Mitbestimmung

Ein weiterer Grund für die Speicherung der Logins/Logouts ist die IT-Sicherheit. Das ist ein legitimer Zweck. Doch ist damit die lückenlose Erfassung aller Aktivitäten von Beschäftigten mit ihren Arbeitsmitteln gerechtfertigt? Worin liegt die Anforderung etwa für ein Protokoll eines Filmvorführers in einem Kino? Der Arbeitgeber hielt das für dringend erforderlich, um sicher zu stellen, dass keine Horrorfilmwerbung im Kinderprogramm läuft. Allerdings hatte der Vorführer keinerlei Möglichkeit, außer der zentral geplanten überhaupt Werbung einzuspielen. Damit scheint das Login-Protokoll ungeeignet, das Problem zu lösen. Der Arbeitgeber verglich dann alternativ den Sicherheitsbedarf des Kinos mit dem eines Atomkraftwerks (!) und forderte es erneut. Glücklicherweise wurde im Beschluss einer Einigungsstelle das Login-Protokoll abschließend als nicht erforderlich erachtet. So kann die Filmvorführung nun doch ohne Protokollierung der Beschäftigten erfolgen.

Da die Anzahl der Cyberangriffe auf IT-Systeme unbestritten ständig zunimmt und weitgehend professionell stattfindet, wächst die Bedeutung von Sicherheitsfunktionen wie Firewalls, Virenscannern oder anderen Werkzeugen zur Erkennung von Schadsoftware. Um das Ungewöhnliche und damit eine mögliche Bedrohung automatisiert erkennen zu können, muss erkannt werden, was gewöhnlich ist. Hierzu wird das Verhalten von Beschäftigten vollständig dokumentiert und auf Muster hin analysiert. *Security Information and Event Management (SIEM)* Systeme leisten das. Kann sich ein zugelassener Nutzer im Büro in Dortmund einloggen und zwei Stunden später aus New York nochmal? Voraussetzung für das Erkennen einer solchen sicherheitsrelevanten Anomalie ist, dass das Login-Verhalten dieser betroffenen Person protokolliert und in Echtzeit überwacht wird, sowie definiert ist, was ein regelkonformes Verhalten ist.

Ein Verhalten außerhalb des *Normalen*, also die Abweichung, der Sicherheitsvorfall wird so erkennbar und muss dann weiter bewertet werden.



*Shadow Dexterous Robot Hand holding a lightbulb, with permission CC-BY SA 3.0 unported of Richard Greenhill and Hugo Elias of the Shadow Robot Company*

Der Gesetzgeber hat die Notwendigkeit solcher Kontrollverfahren für die Informationssicherheit gesehen und sieht sie daher im IT-Sicherheits-Gesetz (IT SiG 2.0) zwingend vor, als Maßnahmen für Unternehmen der kritischen Infrastruktur (wie Energieversorgern). Das IT SiG 2.0 verlangt eine Angriffserkennung in Echtzeit (§ 8a Absatz 1a). Erst recht in einem solchen Setting, wenn der Gesetzgeber die Überwachung fordert, müssen die Mitbestimmung wahrgenommen und der Datenschutz sichergestellt werden. Zweckbindung, Speicherfristen und Zugriffsberechtigungen, Verfahren zur Bestimmung des Normalen und Analyseprozesse sind in einer Betriebsvereinbarung zu regeln.

### **Datalakes und andere Datenspeicher**

Einen Paradigmenwechsel finden wir bei den Verfahren zum Erstellen von Auswertungen von Geschäftsprozesse vielleicht. Der klassische Ansatz, die für das Unternehmen erforderlichen Auswertungen fest zu definieren, in einem Softwarecode festzu-

schreiben und mit Berechtigungen zu versehen, wird durch die Speicherung in externen, hoch performanten Speichern mit flexiblen Auswertemöglichkeiten ersetzt. Die so genannten *Data-lakes* oder *Analytics Clouds* sind in der Lage, beliebige Datenquellen zu integrieren und extrem schnelle Auswertungen durchzuführen. Sie stellen sehr einfach zu bedienende Werkzeuge zur Verfügung. Mit dieser Technologie lassen sich Daten aus allen möglichen Datenquellen miteinander verknüpfen und anreichern. Gedacht sind diese Werkzeuge für die Analyse von Geschäftsprozess-Bewegungen, Netzwerk- und Systemprotokollen. Zusätzlich kann mit diesen Daten eine KI trainiert werden, die Verhaltensmuster und Benutzerprofile erkennt. Damit liegt eine Leistungs- und Verhaltenskontrolle vor. Die Regelungen entsprechender Systeme müssen der Leistungsfähigkeit der Technologie gerecht werden und etwa Berechtigungen und Zweckbindung auch bei Analytics-Systemen festschreiben. Einfacher Ausweg wäre der Verzicht auf das Verarbeiten von personenbezogenen Daten in diesen Analytic-Cloud-Systemen, zumal die Bezüge von Beschäftigten für Unternehmensanalysen nur selten erforderlich sind. Einige Unternehmen tun genau das. Auch hier ist die Frage nach dem *Need-to-know-Prinzip* zu stellen.

### **Fazit**

Digitalisierung in den Unternehmen führt zu einer neuen Dimension von Daten mit Personenbezug und damit zu einer neuen Dimension der Kontrollmöglichkeiten. „Bedenken second“ bei der Datenverarbeitung widerspricht Gesetzen und es entstehen Risiken für die Beschäftigten. Der schon lange beschriebene gläserne Mitarbeiter wird in einer Form Realität, wie wir uns das bisher nicht vorstellen konnten. Sanktionen aufgrund von erkanntem Fehlverhalten und gesundheitsschädlicher Überwachungsdruck sind mögliche Folgen. Auf der anderen Seite ist sicher ein großer Nutzen in der Digitalisierung zu realisieren. Der Autor hat im Vorfeld zu diesem Text erstmalig einzelne Themen mittels KI recherchiert. Da er keine Texte übernommen hat, gibt es keine Kennzeichnung hinsichtlich KI-Ergebnissen. Eine Kennzeichnungspflicht von KI-basierten Entscheidungen oder Arbeitsergebnissen, einhergehend mit der Beschreibung der zugrundeliegenden Datenbasis, sind allerdings wesentliche Forderungen im Hinblick auf eine Regelung zur Verwendung künstlicher Intelligenz im Betrieb.

Wenn der Gesetzgeber rechtliche Vorgaben macht und die unterschiedlichen Interessengruppen sich auf ihren verschiedenen Ebenen Regeln im Umgang mit der Digitalisierung geben und sie einhalten, besteht die Chance, eine gute Digitalisierungsgeschichte zu schreiben und vielleicht zu exportieren bzw. als neuen Business Case zu vermarkten. Eine entfesselte Digitalisierung ist sicher ein schlechter Weg.

### **Anmerkung**

1 „6. Einführung und Anwendung von technischen Einrichtungen, die dazu bestimmt sind, das Verhalten oder die Leistung der Arbeitnehmer zu überwachen;“ §87 (1) 6 BetrVG



## Mitbestimmung bei modernen ITK-Systemen

### und die Notwendigkeit, das Betriebsverfassungsgesetz (BetrVG) zu ändern

*Ist das Mitbestimmungsrecht noch zeitgemäß? Wie soll man vor Inbetriebnahme eines Systems wissen, welche zukünftigen Änderungen welche Auswirkungen auf die Persönlichkeitsrechte Beschäftigter haben werden? Das Mitbestimmungsrecht des § 87 (1) Nr. 6 BetrVG stammt aus der Frühzeit der betrieblichen Mitbestimmung, ohne dass bisher eine wesentliche gesetzliche Anpassung erfolgt ist.*

#### Die Ausgangslage

Betriebsräte haben bei der Einführung technischer Systeme mitzubestimmen, wenn mithilfe des jeweiligen Systems das Verhalten und die Leistung der von ihnen vertretenen Beschäftigten kontrolliert oder überwacht werden können. Diese Bestimmung findet sich im Betriebsverfassungsgesetz in einer Auflistung verschiedener Tatbestände, die die Mitbestimmung auslösen können.

Da moderne IT-Systeme in aller Regel mindestens über einen Berechtigungsmechanismus und, schon allein aus Sicherheitsgründen, über eine Nutzungsprotokollierung verfügen, sind die Voraussetzungen des § 87 (1) Nr. 6 BetrVG fast immer erfüllt: IT-Systeme sind fast immer mitbestimmungspflichtig. Die langjährige Rechtsprechung des Bundesarbeitsgerichts (BAG) stellt klar, dass die Eignung zur „Leistungs- und Verhaltenskontrolle“ ausreicht, um das Mitbestimmungsrecht auszulösen. Nicht nötig ist die gewollte oder tatsächliche Nutzung der anfallenden Beschäftigtendaten zu einer Kontrolle oder Überwachung.



Fahrradkuriere, Foto: Bilderwelt68, Pixabay

Der Zeitpunkt der Mitbestimmungspflicht – also wann der Betriebsrat mit dem Ansinnen des Arbeitgebers befasst wird – muss vor der Einführung und dem Betrieb eines mitbestimmungspflichtigen Systems liegen. Eine Nachfrage beim Betriebsrat *in letzter Minute* verbietet sich. Eine fachgerecht geplante System-einführung muss die Information, die Verhandlung und die Zustimmungserteilung des Betriebsrats als Meilensteine im Projektplan vorsehen.

Die Zustimmung kann sehr unterschiedliche Formen haben. Theoretisch kann der Betriebsrat dem Arbeitgeber nach entsprechendem Beschluss des Gremiums eine Mail des Inhalts „Mach

was Du willst“ schreiben. Das dürfte allerdings eher die Ausnahme sein. Häufiger finden sich schriftlich fixierte Absprachen darüber, welche Grenzen für den Systembetrieb gelten sollen. Das können schlichte *Regelungsabreden* sein oder aber, der Standardfall, *Betriebsvereinbarungen*, in denen ausführliche Rahmenbedingungen für den Systembetrieb und die Nutzung der anfallenden Daten vereinbart werden.

Das Betriebsratsgremium muss einen offiziellen Beschluss über die Tatsache und die Rahmenbedingungen der Zustimmung schließen. Die Entscheidung nur eines einzelnen Betriebsratsmitglieds oder der bzw. des Vorsitzenden genügt nicht.

Der Arbeitgeber kann sich nicht darauf berufen, dass „das Betriebsratsmitglied XY doch Mitglied im Projektteam war“, und auch nicht darauf, dass der Betriebsrat informiert worden ist, aber nicht reagiert hat. Der gerne angeführte *Zeitdruck* rechtfertigt nicht, ein mitbestimmungspflichtiges System ohne die vorherige Zustimmung des Betriebsrats einzuführen.

Es liegt also im Interesse und in der Verantwortung des Arbeitgebers, die rechtlich korrekte Zustimmung einzuholen. Er ist derjenige, der bei der Projektplanung an die Schritte denken muss, die für die Durchführung der Mitbestimmung erforderlich sind. Auch muss er die Informations- und Verhandlungsphase mit dem Betriebsrat im Einführungsprozess berücksichtigen und professionell planen. Ein mitbestimmungspflichtiges System ohne Zustimmung des Betriebsrats zu betreiben, ist keine gute Idee. Und die Mitbestimmung in letzter Minute durchpeitschen zu wollen, weil erst drei Tage vor geplanter Produktivsetzung zufällig daran gedacht wurde, ist eine Zumutung, die leider allzu häufig vorkommt.

Was passiert nun, wenn Informations- und Verhandlungsphase zeitlich angemessen ablaufen und die Betriebsparteien die Modalitäten des Systembetriebs in einer Betriebsvereinbarung festhalten möchten?

#### Die Herausforderung

Betriebsvereinbarungen sind Verträge. Sie werden zwischen dem Arbeitgeber und dem Betriebsrat des jeweiligen Unternehmens geschlossen.

Grundprinzipien, die für andere Arten von Verträgen gelten, gelten für Betriebsvereinbarungen gleichermaßen. Dass man das betonen muss, liegt an dem häufig zu beobachtenden Phänomen, dass Arbeitgeber die Verbindlichkeit einer Betriebsverein-

barung leugnen. Zu diesen Grundprinzipien gehört, dass ein Vertrag nicht einseitig geändert werden kann. Sind die Unterschriften der Vertragspartner geleistet, dann gilt das Vereinbarte – und zwar als Ganzes und nicht nur in Teilen. Wenn Arbeitgeber das vereinbarte System anpassen möchten und Änderungen wünschen, die den Bestimmungen einer Betriebsvereinbarung widersprechen würden, müssen die Änderungen von beiden Vertragsparteien neu verhandelt und vereinbart werden. Erst danach dürfen sie gemäß der neuen Vereinbarung umgesetzt werden.

*Auch nach Abschluss eines Vertrags zur wöchentlichen Lieferung einer Biogemüsebox muss man sich nicht gefallen lassen, dass der Lieferant plötzlich einseitig bestimmt, nur noch konventionell angebautes Gemüse zu liefern.*

Eine Änderung der Vertragsbedingungen kann nur in beiderseitigem Einverständnis erfolgen; andernfalls muss der Vertrag gekündigt werden. Insofern unterscheidet sich eine Betriebsvereinbarung nicht von jedem anderen Vertrag.

Das bedeutet: Jede Änderung eines Vertrags, in diesem Fall der Betriebsbedingungen eines vereinbarten IT-Systems, macht die erneute Durchführung der Mitbestimmung nötig. Der Arbeitgeber muss, wie bei Einführung des Systems, mit dem Betriebsrat als Vertragspartei über die gewünschten Änderungen verhandeln und einen modifizierten, den neuen Rahmenbedingungen entsprechenden Vertrag schließen.

### Patches, Updates und der ganze Rest

Natürlich ist einem in der IT vertraut, dass sich das *Update-Karusell* heutzutage immer schneller dreht. Wo man früher Systeme auf Servern im eigenen Rechenzentrum betrieb und häufig jahrelang keine nennenswerten Änderungen an der Software vornahm, wird heute auf Cloud-Systeme gesetzt, die bei einem Dienstleister betrieben und manchmal fast jede Nacht durch alle Arten von Patches und Updates modifiziert werden. Es ist zudem häufig, dass der Kunde, also der Arbeitgeber, überhaupt keine faktische Einflussnahmemöglichkeit hat. Die Auswirkungen *ungefragt eingespielter Updates* werden den Nutzerinnen und Nutzern oft erst im Nachhinein bewusst.

Arbeitgeber als verantwortliche Systembetreiber haben also das große Problem, dass sie in vielen Fällen, selbst wenn sie wollten, die Konsequenzen von Änderungen auf Beschäftigtendaten und ganz allgemein die Mitbestimmung nicht überblicken. Sie können oft selbst im Nachhinein keine ausreichenden Informationen gem. § 80 (2) BetrVG zur Verfügung stellen, so dass dem Betriebsrat auch keine Bewertung möglich ist. Viele kleinere Änderungen finden sich allenfalls in *Release Notes* wieder, deren Aussagekraft schwankt.

Es zählt daher heute zu den traurigen Selbstverständlichkeiten, dass selbst bei geregelten Systemen, die also mit einer Zustimmung des Betriebsrats betrieben werden, große Diskrepanzen zwischen dem zum Zeitpunkt der Prüfung formal vereinbarten Stand und dem tatsächlichen Systemstand bestehen. Die aus grundsätzlichem Regelungsunwillen des Arbeitgebers nicht geregelten Systeme sind hier explizit nicht angesprochen. Es soll

nur das Problem der durch nicht nachgehaltene Änderungen vom Soll-Stand (der Betriebsvereinbarung) abweichenden Systeme betrachtet werden.

### Verbrauch der Mitbestimmung

Die Schwierigkeit, ein System initial durch Betriebsvereinbarung zu regeln und dann über längere Zeit Vereinbarung und Systembetrieb kongruent zu halten, ergibt sich aus folgenden wesentlichen Aspekten:

1. Die initiale Vereinbarung muss vor der Produktivsetzung geschlossen sein. Andernfalls würde der Arbeitgeber das System rechtswidrig ohne Zustimmung betreiben.
2. Zum Zeitpunkt der initialen Vereinbarung sind wesentliche Charakteristika des Systemeinsatzes häufig weder dem Arbeitgeber noch dem Betriebsrat in allen für die Beschäftigten überwachungsrelevanten Auswirkungen bekannt. Selbst wenn, was selten genug vorkommt, eine sachgerechte technische Dokumentation vorliegt, werden unerwünschte Nebenwirkungen oft erst im Praxisbetrieb deutlich.

Wird eine Betriebsvereinbarung vor dem Systemeinsatz geschlossen, so gilt die Mitbestimmung zu diesem System(stand) als ausgeübt. Ein Nachbessern, beispielsweise aufgrund späterer Erkenntnisse, ist nur in zwei Fällen möglich: Entweder wird durch eine gewünschte Systemänderung ohnehin eine Anpassung der bestehenden Regelung nötig, oder eine der beiden Seiten kündigt die Vereinbarung. Außerhalb dieser beiden Fälle spricht man von „verbrauchter“ Mitbestimmung: Der Betriebsrat kann für einzelne Aspekte keine Nachbesserungen verlangen, solange die Bestimmungen der einmal abgeschlossenen Betriebsvereinbarung vom Arbeitgeber eingehalten werden.

Betriebsräte unterliegen also auch bei unübersichtlichen Systemführungen dem Zwang, dass *der erste Wurf sitzen* muss. Das gilt, obwohl die Identifizierung vieler Gefährdungen durch das System für die Persönlichkeitsrechte Beschäftigter einem Blick in die Glaskugel gleicht.

Wie soll man vor Inbetriebnahme eines Systems wissen, welche Aspekte zukünftigen Änderungen unterworfen sein werden und welche Auswirkungen das auf die Persönlichkeitsrechte Beschäftigter haben wird? Selbst wenn der Arbeitgeber – was selten genug der Fall ist – für eine aussagekräftige Dokumentation gesorgt hat, der man arbeitnehmerunfreundliche Funktionen, unnötige Datenspeicherungen und überbordende Auswertungen entnehmen kann, ist das keine Garantie dafür, dass durch spätere Updates keine zusätzlichen unerwünschten Auswirkungen entstehen.

Die Systembeschreibung als Grundlage der Betriebsvereinbarung bildet jedoch die *einzig juristisch belastbare* Grundlage für einen später eventuell erforderlichen Rechtsstreit. Bei diesem wird dann *nur* geprüft, ob der Arbeitgeber sich an den vereinbarten Stand gehalten hat. In der Betriebsvereinbarung nicht erfasste Sachverhalte und Regelungen können später nicht gerügt werden, auch wenn sich dadurch aus Sicht des Betriebsrats im Nachhinein schädliche Auswirkungen ergeben.

Um diesem Dilemma wenigstens teilweise zu entkommen, haben sich in der Praxis der Gestaltung von Betriebsvereinbarungen Vorgehensweisen etabliert, die das bei Abschluss Unbekannte kompensieren sollen.

### Praktikable Änderungsmodalitäten und bisherige Lösungsansätze

Eine weit verbreitete Methode, die technischen und funktionalen Änderungen unterworfenen Einzelheiten eines IT-Systems in einer Betriebsvereinbarung zu fassen, besteht in der strukturierten Auslagerung bestimmter Aspekte in sogenannte *Anlagen*. Juristisch gesehen besteht kein Unterschied in der Verbindlichkeit, ob eine Aufzählung zulässiger Auswertungen im Haupttext der Betriebsvereinbarung steht oder in einer eigens geschaffenen Anlage *Auswertungen*. Solange derartige Anlagen als Teil der Betriebsvereinbarung definiert werden, sind sie gleichermaßen verbindlich. Allerdings hat die Nutzung von Anlagen zwei wesentliche Vorteile: Die Betriebsvereinbarung ist in ihren Kernpunkten wesentlich besser verständlich (lesbar) und später gewünschte Änderungen (der Arbeitgeber möchte eine zusätzliche Auswertung einführen) sind nach Verhandlung durch Aktualisierung und Unterzeichnung der entsprechenden Anlage relativ einfach zu bewerkstelligen.

In der Vergangenheit waren Betriebsvereinbarungen weit verbreitet, in denen die detaillierten Anlagen mehrere Ordner füllten.

Diese in der Praxis kaum zu handhabende Textflut entsprang der Notwendigkeit, möglichst alle Eventualitäten vorzubedenken und eine *abschließende Positivliste des Erlaubten* zu erstellen, damit bei einer eventuellen gerichtlichen Auseinandersetzung einfach festzustellen ist, ob eine bestimmte Nutzung, Auswertung o. ä. zulässig war oder nicht.

Sowohl zwecks Handhabung eventueller Änderungen als auch wegen der Schwierigkeiten bei der dauerhaften Befolgung derartiger Vorgaben werden immer wieder Wege zur Vereinfachung der Anlagen und der Änderungssystematik gesucht, die keine Einschränkung der Mitbestimmungsrechte und des Beschäftigtenschutzes bewirken.

### Mitbestimmung in Phasen

Um aus dem Zwang zum Glaskugellesen bei der Erstvereinbarung eines teilweise unbekanntes, eventuell nicht ausreichend dokumentierten Systems zu kommen, teilen manche Betriebsparteien die *Mitbestimmung in verschiedene Phasen* auf und bestimmen, dass sie erst dann als vollständig ausgeübt gelten soll, wenn alle Phasen wie vorgesehen durchlaufen wurden. Das erfordert allerdings eine gute Organisationsstruktur, eine gute Projektplanung sowie eine gute Kooperation der Betriebsparteien. Aus einem Zusammenspiel von vorbereitenden Maßnahmen, Systemeinführung, Untersuchungs- und Dokumentationsaufgaben, regelmäßigen Mitbestimmungsbesprechungen der Betriebsparteien, schrittweiser Lieferung erarbeiteter Konzepte und dem jederzeitigen Vetorecht des Betriebsrats (solange die Mitbestimmung nicht gemäß Phasenmodell abgeschlossen wurde) kann so ein sinnvoller und praxisorientierter Mitbestimmungsprozess stattfinden.

Ist ein derart aufwändiges Modell nicht durchsetzbar, bestehen auch andere Möglichkeiten der Vereinfachung von Systemdokumentation und Änderungsverfahren.

### Mitbestimmung nach Kategorie

In Bestimmungen einer Betriebsvereinbarung können beispielsweise *Änderungsprozesse mit unterschiedlicher Beteiligungstiefe* festgelegt werden. Dazu kategorisieren die Parteien Anlagen oder Vereinbarungsteile in Abhängigkeit von der erwarteten Eingriffstiefe in das Persönlichkeitsrecht der Beschäftigten. So können in einer Anlage bestimmte verwendete Beschäftigendaten als hoch sensibel eingestuft werden, dagegen eine Anlage mit der topologischen Darstellung des Systems als wenig kritisch. Aus der gewählten Kategorisierung ergeben sich dann Änderungsverfahren mit unterschiedlich hoher Komplexität und Tiefe (beispielsweise mit vorheriger Information und Verhandlung, mit kombinierter Information und späterem Vetorecht, mit reiner Information ohne Vetorecht). In Bezug auf die Persönlichkeitsrechte als weniger kritisch eingestufte Änderungen verursachen so weniger Aufwand beim Umsetzen von Systemänderungen.

Die Schwierigkeit aus Sicht des Betriebsrats besteht in der Vorausschau, welche Sachverhalte in welchen Anlagen er für so sensibel hält, dass er Änderungen klassisch mitbestimmen will und an welchen Stellen er das Risiko einer Lockerung im Änderungsprozess tragen kann.

### Nachlaufende Mitbestimmung

Arbeitgebern und Betriebsräten stehen gleichermaßen bei heutigen Cloud-Systemen vor dem Problem, dass ein dienstleistender Betreiber ungefragt Änderungen durchführt, die der Fehlerbehebung oder der Schließung von Sicherheitslücken dienen. Eine eventuell gebotene vorherige Beteiligung des Betriebsrats ist dann nicht mehr möglich. Bei Sicherheits-Patches besteht regelmäßig auch kein vernünftiger Grund, die Zustimmung zu verweigern. Ganz im Gegenteil: Zur Erhaltung der Sicherheit, auch der im System verarbeiteten Beschäftigendaten, hat auch der Betriebsrat ein Interesse an deren schneller Schließung.

Diese Erkenntnis mündet manchmal in Bestimmungen, die den Umgang mit derartigen Sicherheitspatches oder als vergleichbar harmlos eingeschätzten Patches in eine eigene Kategorie einordnen, für die ein Änderungsmechanismus mit *automatischer Bereitstellung von Release Notes* des Dienstleisters mit *nachträglichem Vetorecht* vereinbart wird für den Fall, dass der Sicherheits-Patch im Hintergrund unerwartet beschäftigtenrelevante Änderungen zur Folge hat, die erst später entdeckt werden.

Mit dieser Vereinfachung wird eine *nachlaufende Mitbestimmung* etabliert. Sie basiert auf der grundsätzlichen Erkenntnis, dass weder eine detaillierte vollständige Positivliste erstellt werden kann noch jede kleinste Änderung tatsächlich für den Schutz der Beschäftigten relevant ist – so wie das bei vielen Fehlerbehebungs-Patches der Fall ist. In derartigen Modellen ist der Detaillierungsgrad von Anlagen nur dort hoch, wo auch der

Schutzbedarf der Beschäftigten offensichtlich hoch ist. Erscheinen Anlageninhalte zum Zeitpunkt des Vertragsschlusses wenig heikel oder können deren Auswirkungen nicht abschließend beurteilt werden, so einigt man sich vorläufig auf eine einfache Darstellung, die überblicksartig, *grob, auf großer Flughöhe* ausfallen kann.

Entscheidend für dieses Zugeständnis des Betriebsrats ist jedoch, dass er zu einem beliebigen späteren Zeitpunkt eine detaillierte Darstellung verlangen und anschließend die Wiederaufnahme der Mitbestimmung zu diesem Aspekt verlangen kann. Auslöser für eine solche Nachbewertung können Beschwerden Beschäftigter, eigene Erkenntnisse des Betriebsrats oder Vorfälle sein, die zwar nicht *expressis verbis* den Bestimmungen der Betriebsvereinbarung widersprechen, aber so nicht erwartbar waren und für Beschäftigte nachteilig sein können.

Für derartige Fälle kann in einer Betriebsvereinbarung bestimmt werden, dass die Mitbestimmung für die beschriebenen vereinfachten Fälle/Anlagen nicht als *verbraucht* gilt und der Arbeitgeber sich daher nicht aufgrund bestehender Betriebsvereinbarung weigern kann, erneut zu verhandeln. Das garantiert ein Auffangnetz für Fälle, in denen der Betriebsrat, auch zur Erleichterung des Arbeitgebers, auf eine genaue Darstellung verzichtet und später offenbar werdende Implikationen übersehen hat.

### Grenzen durch das BetrVG

Die oben dargestellten Lösungsansätze zur praktikableren Handhabung von Änderungen können in jeder Betriebsvereinbarung vereinbart werden. Allerdings gelingt das nur, wenn zwischen den Verhandlungspartnern ein gewisses Maß an *Vertrauen* und an *gutem Willen* vorhanden ist.

Verhandlungssituationen, in denen Arbeitgeber und Betriebsrat bei jeder noch so kleinen technischen Festlegung darüber streiten, ob sie mitbestimmungspflichtig ist, sind für alle Arten nach-

laufender Mitbestimmung nicht geeignet, weil sie im BetrVG nicht explizit vorgesehen sind und im Streitfall die Wirksamkeit des vereinbarten Konstrukts davon abhängt, wie das Arbeitsgericht diesen Teil der Vereinbarung bewertet.

Auch für konfliktäre Verhandlungen, die schließlich in einer Einigungsstelle münden, verbieten sich vereinfachte Systembeschreibungen oder vereinfachte Änderungsmodalitäten. Derartige Vereinbarungsteile wären wohl regelmäßig nicht spruchfähig, weil ihnen keine interessenausgleichende Regelung innewohnt. In solchen Konfliktfällen ist die klassische, aufwändige und möglichst weit vorausschauende Dokumentation der Systemcharakteristika die einzige zur Verfügung stehende Möglichkeit, einen Vertragsschluss ohne Aufgabe von Mitbestimmungsrechten herbeizuführen. Die Änderungssystematik des BetrVG für Betriebsvereinbarungen ist auch für den Arbeitgeber die einzige Möglichkeit, ein System dauerhaft rechtskonform zu betreiben: Jede technische oder funktionale Änderung, die die Betriebsvereinbarung, also den geschlossenen Vertrag, falsch werden lässt, muss *vor* Umsetzung eine vertragsändernde Mitbestimmung durchlaufen.

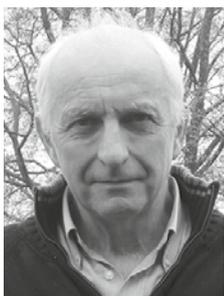
Diese Situation stellt letztlich keine der beiden Seiten zufrieden. Aufgrund der praktischen Umsetzungsprobleme wird so zumeist auch nicht die gewünschte Schutzwirkung für die Beschäftigten entfaltet. Ist für den Betriebsrat die Aufrechterhaltung einer Betriebsvereinbarung aufgrund der erfolgten Veränderungen nicht mehr tolerierbar, so bleibt ihm nur die Möglichkeit, der vollständigen Kündigung, die, soweit nichts anderes vereinbart wurde, nach einer Frist von drei Monaten wirksam wird (§ 77 (5) BetrVG). Das hat regelmäßig zur Folge, dass sämtliche – auch die die Beschäftigten schützenden – Regelungen unwirksam werden, soweit sie nicht einer Nachwirkung unterfallen, also bis zum Abschluss einer neuen Betriebsvereinbarung zum gleichen Thema weitergelten. Auch wenn durch die (in Teilen) nachwirkende Betriebsvereinbarung kein völlig rechtsfreier Zustand entsteht, ist die Schutzwirkung für Beschäftigte stark eingeschränkt. Das führt im günstigsten Fall zu Neuverhandlungen und bei fehlender Einigung zur Einrichtung einer Einigungsstelle.

### Karin Schuler und Thilo Weichert



**Karin Schuler** ist Diplom-Informatikerin und seit über dreißig Jahren selbständige Beraterin und Sachverständige für Datenschutz und IT-Sicherheit. Sie berät Betriebsräte und Datenschutzbeauftragte. Karin ist FfF-Mitglied, ehemalige Vorständin der Deutschen Vereinigung für Datenschutz und Gründungsmitglied des Netzwerks Datenschutzexpertise.

*Foto: Christian Liepe*



**Thilo Weichert**, gelernter Jurist und Politologe, war 2004 bis 2015 Landesbeauftragter für Datenschutz in Schleswig-Holstein. Er ist heute als Datenschutzberater tätig und engagiert sich im Vorstand der Deutschen Vereinigung für Datenschutz und im Netzwerk Datenschutzexpertise.

## Änderungsbedarf: BetrVG

Das Problem, dass durch äußere Veränderungen die Grundlagen einer Betriebsvereinbarung nicht mehr passen, kann bei jeder Betriebsvereinbarung auftreten. Das spezifische von Vereinbarungen nach § 87 (1) Nr. 6 BetrVG, also der Mitbestimmung bei überwachungsgeeigneten IT-Maßnahmen, ist, dass diese meist eine langfristige Wirkung haben und dass mitbestimmungsrelevante Systemveränderungen eher die Regel als die Ausnahme sind. Daher sind die folgenden Überlegungen zu einer „nachlaufenden Mitbestimmung“ zwar generell relevant, vor allem aber von hoher praktischer Bedeutung für die hier erörterten Fragen zu IT-Systemmodifikationen.

Die bisherige Darstellung hat gezeigt, dass eine *nachlaufende Mitbestimmung* im IT-Bereich sinnvoll ist, aber in der Praxis davon abhängt, dass hierzu beide Betriebsparteien und damit der Arbeitgeber auch tatsächlich bereit sind. Rechtlich erzwingbar ist sie nicht und verbindlich gestaltbar nur unter erschwerten Bedingungen. Im Rahmen eines konfliktären Einigungsstellensverfahrens ist es eher ein glücklicher Umstand, wenn ein solches Verfahren vom Einigungsstellenvorsitzenden durchgesetzt wird.

Insofern liegt es nahe, dass hier der Gesetzgeber tätig wird. Dieser kann die in der Praxis bereits erprobten Konzepte der Klassifizierung von Vereinbarungsteilen mit unterschiedlichen Mitbestimmungsabläufen und der nachlaufenden Mitbestimmung auf eine verbindliche, d. h. auch rechtlich erzwingbare Grundlage stellen. Angesichts des Umstands, dass das Bundesarbeitsministerium derzeit ohnehin an einem Beschäftigtendatenschutzgesetz arbeitet, bei dem auch das Verhältnis von Datenschutz- und Arbeitsrecht neu austariert wird, sollte eine entsprechende Änderung im BetrVG möglich sein. Sie ist zumindest diskussionsbedürftig.

Die *Abkehr vom Prinzip der abschließenden Regelung* und die Hinwendung zu einem System der Mitbestimmung, bei dem bereits vereinbarte Systeme unter bestimmten Voraussetzungen erneut der Mitbestimmungspflicht unterfallen, bedürfen einer ausdrücklichen gesetzlichen Grundlage. Diese Pflicht sollte unabhängig davon vorgesehen werden, ob durch eine geplante Änderung die bestehende Betriebsvereinbarung unzutreffend würde. Derartige Änderungen würden auch heute schon – so zumindest der rechtliche Anspruch der Regelung – eine erneute Mitbestimmungspflicht auslösen. Tatsächlich besteht aber eine hohe Rechtsunsicherheit, ob und wenn ja inwieweit, die gegebenen Veränderungen die Voraussetzungen des § 87 (1) Nr. 6 BetrVG zur Folge haben.

Um diese Unsicherheit aufzulösen, könnten als Auslöser einer *nachlaufenden Mitbestimmung* explizit geregelt werden:

- Beschäftigte beschwerten sich, weil sie durch bestimmte, nicht klar vereinbarte Auswertungen Nachteile erleiden.
- Der Betriebsrat hat konkrete Erkenntnisse zu nachteiligen Auswirkungen für Beschäftigte, die beim Abschluss der Betriebsvereinbarung nicht erkennbar waren und daher in der Betriebsvereinbarung nicht thematisiert wurden.
- Nach Abschluss der Betriebsvereinbarung werden Funktionen und Verarbeitungen ergänzt, die formal aufgrund *großer Flughöhe in der Betriebsvereinbarung*, also wegen ihrer abstrakten Beschreibung von der Zustimmung der Betriebsvereinbarung

erfasst sind, im Detail jedoch entweder nicht erkennbar waren oder durch eine Kombination mit anderen zulässigen Verarbeitungen zu neuen personenbezogenen Erkenntnissen führen.

- Der Arbeitgeber kann oder will bezüglich erfolgter Änderungen keine ausreichenden Unterlagen beibringen und beruft sich auf Nichtwissen („der Hersteller macht nächtliche Updates, da weiß ich selbst nicht, was da passiert“).

### Die Vorteile

Die gesetzliche Festlegung der Rahmenbedingungen nachlaufender Mitbestimmung brächte für Beschäftigte und Arbeitgeber gleichermaßen Vorteile:

- Der Zwang zum Glaskugellesen und die damit verbundene Furcht, wesentliche Aspekte einer sinnvollen Regelung zu vergessen oder übersehen zu haben, verlören an Schärfe. Das Vertrauen der Belegschaft in die eingesetzten Systeme würde gesteigert. Die Arbeitgeber erhielten eine arbeitsrechtliche Handhabe gegenüber ihren Dienstleistern, die erforderlichen Informationen über das System und über Systemänderungen jederzeit nachzufordern.
- Die Bereitstellung (noch) nicht vorhandener Dokumentationssteile, Konzepte oder Einsatzszenarien könnten unter einem Regime der *nachlaufenden Mitbestimmung* auf einen späteren Zeitpunkt verlegt werden.
- Die Experimentierbereitschaft hinsichtlich von IT-Innovationen würde sowohl für den Arbeitgeber als auch für die Belegschaft erhöht, was letztlich der Innovationsfähigkeit der Unternehmen zugutekäme.
- Die unsägliche Diskussion, ob bestimmte Systeme überwachungsgeeignet sind oder nicht und inwieweit Datenschutzfragen Einfluss auf das Mitbestimmungsrecht haben, würde von der theoretischen auf die praktische Ebene verlagert, was eine konkrete Verbesserung des Persönlichkeitsschutzes für Beschäftigte zur Folge hätte.

Das Mitbestimmungsrecht des § 87 (1) Nr. 6 BetrVG stammt aus der Frühzeit der betrieblichen Mitbestimmung, ohne dass bisher eine wesentliche gesetzliche Anpassung erfolgt ist. Das Bundesarbeitsgericht zog angesichts der technischen Entwicklung von Verhaltens- und Leistungskontrollen schon in den 1970er-Jahren den Schluss, dass die „Bestimmung zur Überwachung“ objektiv und unabhängig von den tatsächlichen Praktiken und Absichten des Arbeitgebers zu bewerten ist. Angesichts der inzwischen für alle Beteiligten bestehenden Intransparenz der im Einsatz befindlichen und erst recht der in Zukunft absehbaren Systeme zur Verhaltens- und Leistungskontrolle *ist das Mitbestimmungsrecht nicht mehr zeitgemäß*. Es ist daher nicht abwegig, generell die Wahrung des Datenschutzes von Beschäftigten der Mitbestimmung zu unterwerfen. Sollte sich dies politisch nicht durchsetzen lassen, so wird mit dem vorliegenden Vorschlag ein Weg aufgezeigt, mit dem in praktikabler und rechtssicherer Form der Persönlichkeitsschutz erheblich verbessert und der Technikentwicklung angepasst werden kann, ohne das bestehende Gesamtgefüge der informationstechnischen Mitbestimmung in Frage zu stellen.



Daniel Leisegang

## Algorithmen rütteln kaum an politischen Einstellungen

31. Juli 2023 – *Die Empfehlungsalgorithmen sozialer Plattformen sind mitverantwortlich für die Verbreitung von Falschinformationen und die gesellschaftliche Polarisierung, heißt es gemeinhin. Aber stimmt das tatsächlich? Vier aktuelle Studien kommen zu einem gemischten Ergebnis.*

Es ist ein wissenschaftliches Großprojekt: In insgesamt 16 Studien untersuchen Wissenschaftler:innen, welchen Einfluss Algorithmen auf die Verbreitung von Falschinformationen und Polarisierung haben. Zwei der weltweit größten sozialen Netzwerke stehen dabei im Fokus: Facebook und Instagram. Beide Plattformen gehören dem Tech-Konzern Meta an.

Die ersten vier Studien wurden nun in den wissenschaftlichen Zeitschriften *Science*<sup>1</sup> und *Nature*<sup>2</sup> veröffentlicht. Die Ergebnisse zeichnen – zumindest für den Zeitraum der zweiten Jahreshälfte 2020 in den Vereinigten Staaten – ein differenziertes Bild davon, wie US-Amerikaner:innen soziale Netzwerke nutzen, und welche Auswirkungen dies auf ihre politischen Überzeugungen und den gesellschaftlichen Diskurs hat.

Demnach spielen die beiden Plattformen eine entscheidende Rolle dabei, die Nutzer:innen zu Inhalten zu leiten, denen sie wahrscheinlich zustimmen. Die Studien bezweifeln jedoch, dass Meta die politischen Überzeugungen der Nutzer:innen signifikant beeinflussen kann.

Aus Sicht von Meta bestätigen die Studien damit frühere Aussagen des Unternehmens<sup>3</sup>, wonach es für die Nutzer:innen von Vorteil sei, wenn Algorithmen deren Feeds sortierten. Kritiker:innen entgegnen, dass die Studien vor allem zeigten, dass das Problem mannigfaltiger sei als bislang zumeist angenommen und der Konzern nicht aus der Verantwortung entlassen sei.

### Verschärfen Algorithmen die gesellschaftliche Polarisierung?

Die Forschungen starteten offiziell Ende August 2020, also noch vor der US-Präsidentenwahl im November des gleichen Jahres. Ein Team aus rund zwei Dutzend Facebook-Forscher:innen und externen Wissenschaftler:innen untersuchte über mehrere Monate konkret<sup>4</sup>, welche Auswirkungen Facebook und Instagram unter anderem auf die politische Beteiligung, die politische Polarisierung und die Verbreitung von Falschwahrnehmungen hat.

An drei der vier Studien nahmen für einen Zeitraum von knapp drei Monaten mehrere zehntausend US-amerikanische Facebook- und Instagram-Nutzer:innen teil. Ihre Identifikationsdaten wurden dabei unkenntlich gemacht. Für die vierte Studie

stellte Meta den Forschenden anonymisierte Daten von rund 208 Millionen Facebook-Nutzer:innen zur Verfügung.

In der ersten Studie mit dem Titel *Wie beeinflussen Algorithmen von Social-Media-Feeds die Einstellungen und Verhalten in Wahlkampagnen?*<sup>5</sup> sahen die teilnehmenden Facebook- und Instagram-Nutzer:innen drei Monate lang einen chronologisch statt algorithmisch sortierten Nachrichten-Feed. Die Beiträge waren also nach ihrer zeitlichen Aktualität sortiert und nicht vorrangig nach den persönlichen Interessen der einzelnen Nutzer:innen.

Die Forschenden fanden heraus, dass die chronologische Anzeige – eine Option, die Facebook erst wieder seit Kurzem seinen Nutzer:innen anbietet – „den Anteil von Inhalten aus als nicht vertrauenswürdig eingestuften Quellen um mehr als zwei Drittel im Vergleich zum algorithmischen Feed erhöhte, während er die Anzeige von unhöflichen Inhalten um fast die Hälfte reduzierte“. Zugleich aber habe die chronologische Anzeige keinen nennenswerten Einfluss auf die „Polarisierung“ oder das politische Wissen der Menschen gehabt.

Algorithmen übten zweifelsohne einen großen Einfluss darauf aus, wie Menschen Inhalte auf den Plattformen wahrnehmen, so Talia Jomini Stroud, eine der Studienleiter:innen und Direktorin des Center for Media Engagement an der University of Texas gegenüber dem britischen *Guardian*<sup>6</sup>, „wir haben aber auch festgestellt, dass [...] eine Änderung der Algorithmen in den sozialen Medien die politischen Einstellungen nicht beeinflusst haben.“

Und noch etwas habe sich gezeigt: Ohne algorithmisch vorsortiertem Feed verringert sich die Zeit, die Nutzer:innen auf Facebook oder Instagram verbringen, deutlich. Zugleich verbrachten sie mehr Zeit auf TikTok, YouTube und anderen Plattformen. Aus Sicht der Forschenden zeige dies, wie wichtig die algorithmische Sortierung für Meta aus wirtschaftlichen Gründen ist.

### Facebook und Instagram sind mächtige Kuratiermaschinen

Die zweite Studie<sup>7</sup> nahm das Nutzungsverhalten von etwa 208 Millionen US-amerikanischen Facebook-Nutzer:innen rund um die US-Präsidentenwahl im Jahr 2020 – konkret von September 2020 bis Februar 2021 – in den Blick.

Die Forschenden kommen zu dem Schluss, dass Facebook „ideologisch stark gespalten ist – weit mehr als frühere Forschungen zum Nachrichtenkonsum im Internet auf der Grundlage des Surfverhaltens ergeben haben.“ Die Spaltung werde aber weniger durch die Inhalte geprägt, die Freund:innen posten, sondern erheblich mehr durch Facebook-Seiten und -Gruppen. Aus Sicht der Forschenden veranschaulicht dies, dass Facebook-Seiten und -Gruppen eine „sehr mächtige Kurations- und Verbreitungsmaschine“ seien.

Außerdem stellt die Studie fest, dass weit mehr Konservative Falschinformationen sehen und teilen würden als Liberale. Demnach waren 97 % der als Falschinformation deklarierten Meldungen bei Konservativen beliebter als bei Liberalen. Konservative neigten zudem dazu, vergleichsweise mehr politische Nachrichten zu lesen, die auch von anderen Konservativen gelesen würden, so die Studie.

### Wenig Auswirkungen auf die Überzeugungen

Allerdings habe das Verhalten der Nutzer:innen nicht unmittelbarer Einfluss auf deren Überzeugungen, so das Ergebnis der dritten Studie<sup>8</sup>. Für diese Studie reduzierten die Forschenden die Menge jener Inhalte, die von „gleichgesinnten“ Verbindungen in die Feeds von mehr als 23.000 Facebook-Nutzer:innen gepostet wurden, um rund ein Drittel.

Dies führte zwar dazu, dass die Nutzer:innen deutlich mehr Inhalte aus Quellen jenseits des eigenen politischen Spektrums sahen. Allerdings interagierten sie weniger mit diesen Inhalten als mit jenen, die ihren eigenen Überzeugungen entsprachen. Vor allem aber hatten die anderen Inhalte keine messbaren Auswirkungen auf die Überzeugungen der Studienteilnehmer:innen, ihre Einschätzung von politischen Kandidat:innen oder die politische Polarisierung im Allgemeinen.

Zu einem ähnlichen Schluss kommt die vierte Studie<sup>9</sup>, die das Verhalten von 27.000 Facebook- und Instagram-Nutzer:innen untersuchte. Hier nahmen die Forschenden den Teilnehmer:innen die Möglichkeit, Beiträge erneut zu teilen. Diese sogenannten Reshares sorgen maßgeblich dafür, dass Inhalte bei Facebook „viral gehen“.

Fehlt die Möglichkeit, hat dies vor allem Auswirkungen darauf, wie sich politische Nachrichten sowie Meldungen aus nicht-vertrauenswürdigen Quellen verbreiten. Infolgedessen nimmt zwar

unter anderem das Wissen der Studienteilnehmer:innen über politische Nachrichten ab. Allerdings habe auch dies keine signifikanten Folgen auf die Überzeugungen der Nutzer:innen, so die Studie.

### Ergebnisse sind nur eingeschränkt verallgemeinerbar

Meta sieht sich durch die bisherigen Forschungsergebnisse bestätigt. Laut Nick Clegg, Leiter von Metas Unternehmenskommunikation, zeigen die Forschungen<sup>10</sup>, „dass es wenig Beweise dafür gibt, dass zentrale Funktionen von Metas Plattformen allein eine schädliche ‚affektive‘ Polarisierung verursachen oder irgendeinen bedeutenden Einfluss auf wichtige politische Einstellungen, Überzeugungen oder Verhaltensweisen haben.“

Indes sagen selbst die Forschenden<sup>11</sup>, die die erste Studie erstellt haben, dass die Ergebnisse der Studien nur mit Einschränkungen verallgemeinbar seien. Demnach sei der Studienzeitraum von drei Monaten zwar länger als der vieler anderer Untersuchungen zur politischen Kommunikation. Allerdings wären die Ergebnisse „möglicherweise anders ausgefallen, wenn diese Studie nicht während eines polarisierten Wahlkampfes [...] und in einem spezifischen politischen Kontext (den Vereinigten Staaten) durchgeführt worden wäre.“

Die Whistleblowerin Frances Haugen kritisiert ebenfalls den Zeitpunkt der Untersuchungen. Die ehemalige Facebook-Produktmanagerin, die 2021 Tausende von internen Facebook-Dokumenten an die Börsenaufsichtsbehörde weitergab<sup>12</sup>, weist unter anderem darauf hin<sup>13</sup>, dass Meta bereits in den Monaten vor der damaligen US-Präsidentenwahl einige „seiner aggressivsten Wahlschutzmaßnahmen“ eingeführt habe, um gegen extreme Beiträge vorzugehen. Viele der Maßnahmen habe der Konzern nach der Wahl wieder zurückgenommen, so Haugen.

### Meta darf sich nicht der Verantwortung entziehen

Ohnehin sollte sich Meta nicht allzu kräftig auf die Schulter dafür klopfen, dass es Forschenden seine Daten zur Verfügung stellt. Denn die laut Clegg „beispiellose Forschungspartnerschaft zwischen Meta und externen Akademikern“ wird wohl gezwungenermaßen fortgesetzt: Laut Digital Service Act (DSA), der im November vergangenen Jahres in Kraft trat, müssen Internet-Plattformen<sup>14</sup> künftig deutlich rigider gegen „systemische

Daniel Leisegang

**Daniel Leisegang** ist Politikwissenschaftler und seit August 2022 Co-Chefredakteur bei *netzpolitik.org*. Zuvor war er Redakteur bei den *Blättern für deutsche und internationale Politik*<sup>17</sup>. 2014 erschien von ihm das Buch *Amazon – Das Buch als Beute*<sup>18</sup>; 2016 erhielt er für seinen Beitrag *Facebook rettet die Welt* den Alternativen Medienpreis<sup>19</sup> in der Rubrik *Medienkritik*. Daniel gehört dem Board of Trustees von Eurozine<sup>20</sup> an.

**Kontakt:** E-Mail<sup>21</sup> (OpenPGP<sup>22</sup>), Mastodon<sup>23</sup>, Twitter<sup>24</sup>, Telefon: +49-030-577148228 (Montag bis Freitag, jeweils 8 bis 18 Uhr).

Risiken“ vorgehen als in der Vergangenheit. Zu diesen Risiken zählen unter anderem die systematische Verletzung der Privatsphäre oder des Rechts auf freie Meinungsäußerung. Damit diese Risiken aber überhaupt entdeckt werden können, räumt das EU-Gesetz Forscher:innen umfassenden Zugang zu den Daten der großen Plattformen ein – ob diese wollen oder nicht.

Bürgerrechtsorganisationen weisen zudem darauf hin, dass die Studienergebnisse Meta und andere Technologieunternehmen keineswegs aus jener Verantwortung entlassen würden, die sie für gesellschaftliche Spaltungen, politische Umwälzungen und der Verbreitung von Verschwörungsmäthen trügen. „Die von Meta unterstützten Studien, die sich mit kleinen Stichprobenzeiträumen befassen, sollten nicht als Entschuldigung dafür dienen, dass Lügen verbreitet werden können“, sagt Nora Benavidez<sup>15</sup>, Senior Counsel bei der Bürgerrechtsgruppe Free Press gegenüber der Washington Post<sup>16</sup>. „Social-Media-Plattformen sollten sich im Vorfeld von Wahlen stärker engagieren und nicht neue Pläne aushecken, um sich der Verantwortung zu entziehen“, so die Juristin.

Ähnlich sieht das der Journalismus-Professor Michael W. Wagner, der an der Universität von Wisconsin lehrt und als unabhängiger Beobachter der Forschungskoooperation fungiert. Aus seiner Sicht belegen die Studien keineswegs, dass soziale Plattformen wie Facebook „kein Problem sind“. Vielmehr dienen die Studien aus seiner Sicht als „ein guter wissenschaftlicher Beweis dafür, dass es nicht nur ein Problem gibt, das leicht zu lösen ist.“

Quelle: <https://netzpolitik.org/2023/studien-zu-facebook-und-instagram-algorithmen-ruetteln-kaum-an-politischen-einstellungen/>

## Anmerkungen

- 1 [https://www.science.org/doi/10.1126/science.adj7023?adobe\\_mc=MC MID%3D61761082717445531361376140889916401782%7CMCORGID%3D242B6472541199F70A4C98A6%2540AdobeOrg%7CTS%3D1690482061](https://www.science.org/doi/10.1126/science.adj7023?adobe_mc=MC MID%3D61761082717445531361376140889916401782%7CMCORGID%3D242B6472541199F70A4C98A6%2540AdobeOrg%7CTS%3D1690482061)
- 2 <https://www.nature.com/articles/s41586-023-06297-w>
- 3 <https://www.washingtonpost.com/technology/2021/11/13/facebook-news-feed-algorithm-how-to-turn-it-off/>
- 4 [https://medium.com/@2020\\_election\\_research\\_project/a-proposal-for-understanding-social-medias-impact-on-elections-4ca5b7aae10](https://medium.com/@2020_election_research_project/a-proposal-for-understanding-social-medias-impact-on-elections-4ca5b7aae10)
- 5 <https://www.science.org/doi/10.1126/science.abp9364>
- 6 <https://www.theguardian.com/technology/2023/jul/27/meta-facebook-algorithm-2020-election>
- 7 <https://www.science.org/doi/10.1126/science.ade7138>
- 8 <https://www.nature.com/articles/s41586-023-06297-w>
- 9 <https://www.science.org/doi/10.1126/science.add8424>
- 10 <https://about.fb.com/news/2023/07/research-social-media-impact-elections/>
- 11 <https://www.science.org/doi/10.1126/science.abp9364>
- 12 <https://netzpolitik.org/2021/facebook-leaks-whistleblowerin-erhebt-schwere-vorwuerfe-gegen-facebook/>
- 13 <https://www.washingtonpost.com/technology/2023/07/27/social-media-research-meta-political-views/>
- 14 <https://netzpolitik.org/2023/digital-services-act-wie-die-forschung-den-datenschatz-der-plattformen-heben-kann/>
- 15 <https://www.freepress.net/about/staff/nora-benavidez>
- 16 <https://www.washingtonpost.com/technology/2023/07/27/social-media-research-meta-political-views/>
- 17 <https://www.blaetter.de/>
- 18 [http://www.schmetterling-verlag.de/page-5\\_ isbn-3-89657-068-4.htm](http://www.schmetterling-verlag.de/page-5_ isbn-3-89657-068-4.htm)
- 19 <https://www.alternativer-medienpreis.de/preistraeger-2016/daniel-leisegang/>
- 20 <https://www.eurozine.com/>
- 21 <mailto:daniel@netzpolitik.org>
- 22 <https://keys.openpgp.org/search?q=daniel@netzpolitik.org>
- 23 <https://mastodon.social/@dleisegang>
- 24 <https://twitter.com/dleisegang>



Anna Biselli

## Der Verteidigungsminister besucht den Cyberraum

31. Juli 2023 – *Ein Teil der Bundeswehr hackt, öffentliche Diskussion darüber gibt es kaum. Konkrete Antworten liefert auch der Verteidigungsminister nicht, wie sein Antrittsbesuch beim Kommando Cyber- und Informationsraum zeigt.*

Bundesverteidigungsminister Boris Pistorius hat heute die Tomburg-Kaserne in Rheinbach besucht<sup>1</sup>. Dort sitzen viele Angehörige des Kommandobereichs Cyber- und Informationsraum (CIR). Die Cybertruppe der Bundeswehr hat viele Aufgaben: Sie kümmert sich um die IT-Sicherheit der Armee, um digitale Infrastruktur bei Einsätzen, um digitale Lagekarten.

Ein Teil des CIR besteht aber auch aus Hacker:innen, vor allem im sogenannten *Zentrum Cyber-Operationen* (ZCO). Sie gingen aus der früheren Einheit *Computer Netzwerk Operationen* hervor, die sich schon 2015 in die Netze eines afghanischen Mobilfunkbetreibers gehackt<sup>2</sup> haben sollen. Doch genau wie eine offizielle Stellungnahme zu diesem Einsatz bleiben Bundeswehr und die Bundesregierung viele Antworten auf offene Fragen seit Jah-

ren schuldig. Dazu passt der Leitspruch des ZCO<sup>3</sup>: „Keiner hört uns, keiner sieht uns, keiner kennt uns.“

2015 haben wir die Strategische Leitlinie Cyber-Verteidigung<sup>4</sup> im Geschäftsbereich des Bundesministeriums der Verteidigung veröffentlicht. Das war noch vor der Einrichtung des CIR. Damals erklärte das Verteidigungsministerium unter Ursula von der Leyen den sogenannten Cyberraum zum Operationsraum neben den bisherigen militärischen Domänen Land, Luft, See und Weltraum. Und es war die Rede von offensiven „Cyber-Fähigkeiten“ oder „Wirkmöglichkeiten“. Es ging also nicht mehr nur um die Resilienz und Verteidigung der eigenen Systeme, sondern um Angriffe auf die Systeme eines möglichen Gegners und Hackbacks.

## Viele ungeklärte Fragen

Wenn die Bundeswehr digital nicht nur verteidigen, sondern auch hacken oder zurückhacken soll, entstehen mehrere Probleme: Wie kann man sich sicher sein, dass „der Richtige“ getroffen wird? Bits und Bytes tragen keine Uniform<sup>5</sup>, Angreifer verschleiern regelmäßig ihre Herkunft oder legen falsche Fährten. So kann es schnell zu Eskalationen kommen.

Wenn die Bundeswehr Schadsoftware entwickelt und dafür Sicherheitslücken nutzt, steht das im Widerspruch zum Koalitionsvertrag. Der besagt, dass gefundene Sicherheitslücken gemeldet und schnellstmöglich geschlossen werden sollen. Wer sie offen lässt, um vielleicht im Ernstfall einen gegnerischen Mobilfunkanbieter oder eine Radarstation hacken zu können, gefährdet in der Regel die IT-Sicherheit vieler anderer Systeme und Menschen. Und hält ein System am Laufen, das in staatlich geförderter IT-Unsicherheit mündet.

Ab wann soll die Bundeswehr digitale Gegenangriffe starten dürfen? Hackbacks schließt der Koalitionsvertrag aus. Die Verteidiger:innen der „offensiven Cyberverteidigung“ berufen sich immer wieder darauf, dass für die Parlamentsarmee die üblichen Regeln gelten. Aber wann tritt im Digitalen ein Verteidigungsfall ein? Wie wägt man potenzielle Kollateralschäden ab?

### Das meiste bleibt geheim

Um hacken zu können, egal ob beim Erst- oder Gegenschlag, braucht es Vorbereitung. Soll die Bundeswehr die IT-Systeme anderer Staaten ausspähen, um nach Verwundbarkeiten zu suchen? Wer im Ernstfall die Schwachstellen anderer ausnutzen will, muss sich dieses Wissen meist schon vorher ansammeln und Möglichkeiten entwickeln, sie auszunutzen – in Friedenszeiten<sup>6</sup>. Und mühsam gesammeltes Wissen kann schnell mit einem Update oder einer Änderung obsolet werden.

Ein Gutachten der wissenschaftlichen Dienste des Bundestages<sup>7</sup> nennt digitale Gegenmaßnahmen daher treffend „wartungsaufwändige Einmal-Wirkmittel mit hohem Proliferationsrisiko“. Das meint, die Werkzeuge müssen immer wieder aktualisiert werden; sobald die Gegenseite sie kennt und sich dagegen verteidigen kann, werden sie wertlos; sie können jederzeit auch anderen in die Hände fallen.

Viele Fragen zu den digitalen heutigen und künftigen Fähigkeiten, Kompetenzen und Grenzen der Bundeswehr sind seit Jahren in der Öffentlichkeit ungeklärt. Parlamentarische Anfragen dazu beantwortet die Bundesregierung oberflächlich<sup>8</sup> oder unter Vorbehalt der Geheimhaltung gar nicht<sup>9</sup>. Der jährliche Sachstandsbericht „Cyber- und Informationsraum“ – unter Verschluss<sup>10</sup>. Währenddessen gliedert sich die Cybertruppe wieder um, in CIR 2.0<sup>11</sup>.

Die vielen offenen Fragen sind ein Problem, nicht nur wegen der IT-Sicherheit. Das dringt auch in den Antworten des aktuellen Verteidigungsministers Boris Pistorius in einer Pressekonferenz<sup>12</sup> durch, als er gefragt wird, ob es eine klarere Linie braucht, um offensive Angriffe im Ernstfall zu ermöglichen. Es brauche Regelungen, im Cyberraum seien die Grenzen „bisweilen fließend“. „Ich würd's mal so formulieren“, sagt Pistorius. „Was wir nicht gebrauchen können, ist, nicht zu wissen, was wir tun dürfen, wenn wir's tun müssen.“

Der Koalitionsvertrag schließt nicht nur Hackbacks aus und will Schwachstellen konsequent schließen, er will auch die „parlamentarische Kontrolle über den Einsatz von Cyber-Fähigkeiten der Bundeswehr“. Ohne Diskussion, was die Armee dürfen soll, wird das nicht gehen. Und ohne Antworten darauf, was die Armee schon heute tut, erst recht nicht.

Quelle: <https://netzpolitik.org/2023/bundeswehr-der-verteidigungsminister-besucht-den-cyberraum/>

### Anmerkungen

- <https://www.bmvg.de/de/presse/antrittsbesuch-verteidigungsminister-pistorius-bei-cybertruppe-5660586>
- <https://www.spiegel.de/politik/ausland/cyber-einheit-bundeswehr-hackte-afghanisches-mobilfunknetz-a-1113560.html>
- [https://ga.de/region/voreifel-und-vorgebirge/rheinbach/cyberabwehr-beginnt-in-der-rheinbacher-tomburg-kaserne\\_aid-43724989](https://ga.de/region/voreifel-und-vorgebirge/rheinbach/cyberabwehr-beginnt-in-der-rheinbacher-tomburg-kaserne_aid-43724989)
- <https://netzpolitik.org/2015/geheime-cyber-leitlinie-verteidigungsministerium-erlaubt-bundeswehr-cyberwar-und-offensive-digitale-angriffe/>
- <https://netzpolitik.org/2015/im-cyber-raum-gibt-es-keine-uniform-bundesregierung-ignoriert-voelkerrechtliche-probleme-zum-cyberkrieg/>
- <https://netzpolitik.org/2020/militaerische-cyber-operationen-staatliches-hacking-entscheidet-keinen-krieg/>
- <https://netzpolitik.org/2019/geheimes-bundestagsgutachten-attackiert-hackback-plaene-der-bundesregierung/>
- <https://dserver.bundestag.de/btd/20/055/2005597.pdf>
- <https://dserver.bundestag.de/btd/19/103/1910336.pdf>
- [https://media.frag-den-staat.de/files/foi/103576/2018-09-05\\_ifg-bmvg-sachstandsbericht-cyber-informationsraumcleaned.pdf](https://media.frag-den-staat.de/files/foi/103576/2018-09-05_ifg-bmvg-sachstandsbericht-cyber-informationsraumcleaned.pdf)
- <https://www.bundeswehr.de/de/organisation/cyber-und-informationsraum/aktuelles/cir-2-0-cir-gliedert-sich-neu-5092642>
- <https://nextcloud.auf.bundeswehr.de/s/yAx8wwEgpMYx8LW?dir=undefined&openfile=19938554>
- <mailto:anna@netzpolitik.org>
- <https://keys.openpgp.org/search?q=anna@netzpolitik.org>
- <https://mastodon.social/@annskaja>



Anna Biselli

Anna Biselli ist Co-Chefredakteurin bei netzpolitik.org. Sie interessiert sich vor allem für staatliche Überwachung und Dinge rund um digitalisierte Migrationskontrolle.

**Kontakt:** E-Mail<sup>13</sup> (OpenPGP<sup>14</sup>), Mastodon<sup>15</sup>, Telefon: +49-1579-2498920 (Montag bis Freitag jeweils 8 bis 18 Uhr).

## Das sind 650.000 Kategorien, in die uns die Online-Werbeindustrie einsortiert

8. Juni 2023 – Ein Dokument gibt einen einmaligen Einblick in den globalen Datenhandel für die digitale Werbung. Erstmals können wir im Detail nachvollziehen, wie invasiv und kleinteilig die Werbefirmen und Datenhändler uns kategorisieren. Das Bild ist erschreckend, auch zahlreiche deutsche Firmen sind beteiligt.

Was auch immer wir im Internet tun, wird aufgezeichnet und ausgewertet, um uns zielgerichtet Werbung anzuzeigen. Das ist eine Realität, an die viele Menschen sich inzwischen gewöhnt haben – im Gegenzug sind schließlich viele Internetangebote kostenlos. Wo genau unsere Daten landen, wenn wir Websites aufrufen oder Apps nutzen, das können die wenigsten nachvollziehen. Auch daran haben wir uns gewöhnt. Die Wege des Targeted Advertising sind unergründlich.



*Hunderte Firmen arbeiten daran, uns für Werbezwecke digital zu vermessen – CC BY 4.0 netzpolitik.org mit Midjourney*

Die Werbeindustrie tut viel dafür, damit das so bleibt: Die Netzwerke der Datensammler sind selbst für Branchenkenner:innen kaum zu überschauen. Jetzt präsentieren *netzpolitik.org* und das US-Medium *The Markup* einen einmaligen Einblick in das Geschäft mit unseren Daten. Wir haben die Angebotsliste von Xandr ausgewertet, einem der größten Datenmarktplätze der Werbewelt. Sie enthält mehr als 650.000 unterschiedliche Kategorien, in die die Industrie Menschen einsortiert, um sie mit gezielter Werbung erreichen zu können.

Umfang und Detailtiefe dieser Datensammlung sind erschreckend. Es gibt kaum eine menschliche Eigenschaft, die Werbetreibende nicht für Werbung ausnutzen wollen. Sie wollen Menschen aus Dänemark erreichen, die einen Toyota gekauft haben? Kein Problem. Sie wollen Menschen erreichen, die gerade finanzielle Probleme haben? Oder keine Krankenversicherung? Kein Problem. Minderjährige? Schwangere? Homosexuelle? Depressive? Politiker:innen? Alles kein Problem.

„Diese Liste ist das gewaltigste Dokument über den globalen Datenhandel, das ich je gesehen habe“, sagt der Wiener Tracking-Forscher Wolfie Christl. Er hat die Datei aufgestöbert und mit *netzpolitik.org* sowie *The Markup* geteilt. Das US-Medium berichtet heute<sup>1</sup> unter anderem über die zahlreichen sensiblen Daten und macht sie mit einem interaktiven Tool einfach durchsuchbar. Xandr hat auf mehrere Presseanfragen nicht reagiert. Die Liste ist auf Mai 2021 datiert, sie stand bis zu unserer Anfrage auf einer Dokumentationsseite von Xandr offen im Netz. Heute ist sie nicht mehr erreichbar, aber beim Internet Archive gibt es eine archivierte Version der Seite<sup>2</sup> und der Datei [23

MBJ<sup>3</sup>. Laut von uns befragten Jurist:innen zeige die Liste, dass das derzeitige Werbegeschäft strukturell unvereinbar mit Datenschutzerfordernungen ist.

### Riesige Container für Gruppen mit bestimmten Eigenschaften

Zielgerichtete Werbung ist das dominante Geschäftsmodell des Internetzeitalters. Das Versprechen: Mehr verkaufen, indem man genau die Menschen mit Werbung erreicht, die davon besonders angesprochen werden. Es ist ein Versprechen, das zieht. Die Umsätze der Branche klettern seit Jahren in immer größere Höhen, 2022 betragen sie weltweit mehr als 550 Milliarden Dollar<sup>4</sup>. Tendenz: rasant steigend. Denn je mehr man über einen Menschen weiß, desto besser man kann ihn beeinflussen, lautet die Annahme. Die Werbeindustrie arbeitet deshalb seit langem hart daran, immer mehr über uns zu erfahren.

Wann immer wir Websites besuchen, Apps nutzen, an Umfragen teilnehmen oder etwas mit Kreditkarte bezahlen, hinterlassen wir Daten. Hunderte Firmen weltweit sammeln diese Informationen und formen daraus sogenannte „Audience Segments“. Man muss sich diese Segmente als riesige Container für Gruppen von Menschen vorstellen. Jedes Segment enthält eine Liste mit digitalen Identifikationscodes, denen bestimmte Eigenschaften zugeschrieben werden. Das sind zum Beispiel demografische Merkmale, Interessen, Konsumverhalten oder Charaktereigenschaften. Auch Informationen darüber, welche Apps und Websites wir nutzen, welche Orte wir besuchen, woran wir vermeintlich glauben, welche Krankheiten wir haben, und unsere sexuelle Orientierung landen in Segmenten.

Diese Segmente werden in der Industrie gehandelt wie ganz normale Waren. Einer der wichtigsten Umschlagplätze dafür ist der Xandr Marketplace. Das US-Unternehmen gilt als eine der ersten Adressen für alle, die sich bei der Online-Werbung nicht vollständig den geschlossenen Systemen von Google, Meta und Amazon ausliefern wollen. 2022 hat Microsoft Xandr vom US-Telekommunikationsanbieter AT&T übernommen, Branchengerüchten zufolge für etwa eine Milliarde Dollar<sup>5</sup>.

Für den ehemaligen Software-Konzern, der vielen noch primär als Hersteller von Produkten wie Windows und Office ein Begriff ist, war die Übernahme ein weiterer Schritt bei der Transformation zu einem Daten- und Werbekonzern. Während das Unternehmen den Konkurrenten Google lange für seinen Datenhunger kritisierte, erhielt Microsoft kürzlich selbst vom Forbes-Magazin den informellen Titel *Next Ad-Tech Giant*<sup>6</sup>. Neben dem Ausbau des eigenen Werbenetzwerkes kaufte Microsoft unter der Führung von CEO Satya Nadella Startups wie den Werbedienstleister PromoteIQ auf. Auch die Übernahme des Sozialen Netzwerks LinkedIn und des Spieleherstellers Activision Blizzard

passt in diese Strategie. Beide Unternehmen liefern nicht nur umfassende Daten von Nutzer:innen, sondern bringen auch gefragte Werbeplätze mit sich, die Microsoft vermarkten kann.

## Eine Excel-Liste mit 650.000 Zeilen

Mit Xandr integriert Microsoft nun einen zentralen Player der Branche nach und nach in die eigenen Werbedienste. Xandr richtet sich mit seinen Plattformen sowohl an Werbetreibende, die Zielgruppen zuschneiden wollen, als auch an Unternehmen, die selbst Werbeplätze zur Verfügung stellen und dafür Käufer suchen – also zum Beispiel Nachrichtenwebsites, Spiele-Apps oder Smart-TV-Angebote. Außerdem bietet Xandr eine Infrastruktur für den Handel von Daten zwischen Anbietern von Werbeplätzen, Werbetreibenden und anderen Datenhändlern sowie Dienste, mit denen sich die IDs von Menschen synchronisieren lassen, um sie über mehrere Geräte hinweg tracken zu können.

Kurzum: Microsofts Neuerwerbung ist eine wichtige „Infrastruktur für das globale Werbe-Ökosystem“<sup>7</sup>, wie das Branchenmagazin Adzine schreibt. Auf dem Marktplatz können andere Datenhändler ihre Daten zur Nutzung anbieten. Dass Xandr in der Vergangenheit offenbar keine Probleme hatte, hier auch fragwürdige Daten handeln zu lassen, davon zeugt die Angebotsliste. Die Datei enthält hunderte Segmente zu sensiblen Themen wie Gesundheit, sexueller Orientierung, Politik, Religion, Ethnizität und offenbar auch Segmente mit Minderjährigen.

Welche Kunden die bei Xandr angebotenen Daten nutzen, ist nicht bekannt. Wir haben Xandr und Microsoft eine umfangreiche Presseanfrage geschickt. Trotz mehrfachen Nachhakens im Verlauf von fast zwei Wochen haben weder wir noch The Markup Antworten erhalten. Laut Adzine<sup>8</sup> arbeiten in Deutschland beispielsweise die Medienunternehmen Axel Springer und Burda mit Xandr. Die Zahl der Kunden ist vermutlich groß, 650.000 Segmente dürften für viele als Pluspunkt gelten. In der digitalen Werbebranche ist schließlich allen bekannt, dass das derzeitige System auf granularen Datensammlungen basiert.

An dieser Stelle ist wichtig zu betonen, dass Xandr und ähnliche Plattformen ihren Werbekunden die Daten nicht direkt überlassen. Die Werbetreibenden zahlen nicht für die Rohdaten mit den pseudonymen IDs einzelner Personen. Sondern dafür, über spezielle Plattformen zielgerichtet Menschen in bestimmten Zielgruppen-Segmenten zu erreichen. Viel davon findet automatisiert statt, die Branche spricht deshalb von Programmatic Advertising. Damit dieses System funktioniert, müssen die personenbezogenen Daten von Internetnutzer:innen durch die Hände zahlreicher Unternehmen gehen. Immer wieder kommt zudem ans Licht, dass einige Firmen doch auch Rohdaten verkaufen, zum Beispiel an das FBI<sup>9</sup>.

## So sieht die Angebotsliste des Datenmarktplatzes aus

Die Angebotsliste des Xandr Marketplace enthält die Metadaten zu insgesamt 651.463 Segmenten. Für jedes Segment gibt es in der Excel-Datei eine eigene Zeile, die den Namen sowie eine ID des Anbieters enthält, den Namen des Segments und eine Segment-ID. Das sieht beispielsweise so aus:

*Adsquare (Data Provider) | 2711 | 25423859 | Adsquare Data Alliance > DE > Place Visits > By Category > Eat and Drink > Restaurant > Fast Food (adsquare)*

Das bedeutet, der Datenhändler Adsquare bot Daten zu Personen in Deutschland an, die Fast-Food-Restaurants besucht haben. Die Segmentnamen geben häufig Hinweise auf die Unterkategorien. Hier ist es *Place Visit* für besuchte Orte, andere Beispiele sind *Purchase* für getätigte Einkäufe, *Top Mobile App* für die meistgenutzte App oder *Travel intent* für Orte, an die jemand mutmaßlich reisen wird.

Viele Unternehmen der Branche lehnen den Begriff *Datenhändler* für ihr Geschäft ab. Sie nennen sich lieber Technologieplattformen, Infrastrukturdienstleister oder *Location Intelligence Platform*. Schließlich geben sie die Rohdaten meist nicht direkt an die Endkunden weiter. Aber die Firmen beziehen Daten aus unterschiedlichen Quellen, organisieren und verpacken sie neu, helfen, Personen über unterschiedliche Geräte hinweg zu tracken und zu erreichen, und bieten sie anderen Firmen gegen Geld oder andere ökonomische Vorteile zur Nutzung an. Das ist Datenhandel.

Insgesamt 93 dieser Firmen werden in der Xandr-Datei explizit als *Data Provider* gelistet, also als Datenlieferanten. Das sind Unternehmen, die die von ihnen vertriebenen Daten offenbar direkt bei Xandr angeboten haben. Sieben davon sind deutsche Datenhändler. Sie erhalten ihre Daten aus unzähligen Quellen. Dutzende Firmen sind in der Datei durch die Bezeichnung der Segmente als ursprüngliche Quelle für die Daten zu erkennen, auch darunter zahlreiche deutsche Unternehmen. Man findet zum Beispiel diverse Segmente, die offenbar auf Daten beruhen, die von der Kreditkartenfirma Mastercard stammen. Die Namen der Segmente lesen sich so, als würde die Kreditkartenfirma Informationen über Einkäufe und daraus abgeleitet über besuchte Orte und Interessen zur Verfügung stellen:

- Eyeota – DE Mastercard – Top Spending Geography – Casino and Gambling Activities
- Eyeota – DE Mastercard – Top Spending Geography – Children's Apparel
- Eyeota – DE Mastercard – Top Spending Geography – College or University Education
- Eyeota – DE Mastercard – Top Spending Geography – Computer and Software Stores

Auf Anfrage sagt Mastercard, man biete Dienstleistungen auf Basis der „Erstellung aggregierter, geografischer Einblicke in die Ausgaben aus anonymen Transaktionsdaten“ an. Auf unsere Frage, wie genau die Daten mit Bezug zu Mastercard-Transaktionen in der Segmentliste von Xandr landen, antwortete Mastercard nicht.

Keine Informationen enthält die Datei von Xandr darüber, wie viele IDs von Menschen die einzelnen Segmente enthalten und wie die IDs aussehen. Bekannt ist allerdings, dass dies hunderttausende oder gar Millionen unterschiedliche IDs sein können. Allein Oracle, in der Xandr-Liste mit mehr als 200.000 Segmenten der größte Anbieter, behauptet, Daten über mehr als fünf Milliarden Personen zu haben<sup>10</sup>.

## Mercedes, Mütter und Militär

Die Angebotsliste in ihrem ganzen Ausmaß zu erfassen, ist schwer. Das liegt nicht nur am Umfang der Segmentsammlung, sondern auch daran, dass die Kategorienamen je Datenhändler sehr unterschiedlich aufgebaut sind. Die Liste enthält außerdem viele Segmente, die speziell für einzelne Werbekunden erstellt wurden. Mehr als 50.000 Segmente tragen die Bezeichnung *custom*, was sich mit *maßgeschneidert* übersetzen lässt. Laut Xandr-Dokumentation<sup>11</sup> werden so Segmente bezeichnet, die nicht für alle Werbekunden nutzbar sind. Stattdessen müssen sie erst von den Anbietern freigeschaltet werden.

Trotz dieser Unübersichtlichkeit hat The Markup eine Datenanalyse vorgenommen, die zumindest eine grobe Häufigkeitsverteilung einiger übergeordneter Kategorien zeigt. Demzufolge sind Segmente mit Verweisen auf den Automobilsektor die größte Gruppe. Werbetreibende können mit den Xandr-Daten beispielsweise Fans oder Besitzer:innen einer bestimmten Automarke ansprechen oder Leute, in deren Haushalt es mehr als zwei Autos gibt und die jährlich mehr als 32.000 Kilometer fahren. Allein zum Schlagwort *Mercedes* findet man mehr als 1.000 Segmente.

Die zweitgrößte Gruppe sind demografische Merkmale. Werbetreibende können nicht nur nach Geschlecht oder Alter auswählen, sondern etwa Eltern von Teenagern, alleinerziehende Mütter mit Kleinkindern oder Menschen, die gerade vor der Scheidung stehen. Oft fließen Informationen über den Lebensstil mit ein, etwa *konservative Rentner*, *urbane Eliten* oder auch *multikulturelle Familien*. Mütter scheinen eine besonders interessante Gruppe zu sein, es gibt Segmente mit *Soccer Moms*, *Big City Moms*, *Busy Moms* oder auch *Moms who shop like crazy*.

Der groben Analyse zufolge basiert die drittgrößte Segmentgruppe auf Informationen zu Beruf oder Branche von Perso-

nen. Segmente tragen dann Namen wie *Beauty Center Owner*, *Lawyer* oder *Policy Makers*. Fast immer sind die Segmente auf Englisch benannt, auch wenn sie sich explizit auf Länder wie Deutschland oder Frankreich beziehen. Auf Deutsch betitelte Segmente findet man selten, etwa *KMU Entscheider kleine & mittelständische U*. Häufig geht es in dieser Kategorie auch um Mitarbeiter:innen ganz bestimmter Unternehmen, etwa *aldi competitor* oder *Volvo SUV Konkurrenz*. Selbst Angehörige von Militär und Polizei kann man als Zielgruppe auswählen.

## Brustkrebs, Depressionen und Schwangerschaftsabbrüche

Hunderte Segmentbezeichnungen deuten auf hochsensible Daten wie Gesundheitsinformationen. Die Werbetreibenden können aus Kategorien wie Brustkrebs, Blasenkrebs oder Depression wählen. Einige Segmentnamen verweisen sogar auf Besucher:innen einzelner Kliniken. Hier einige Beispiele des US-Anbieters Liveramp:

- LiveRamp Data Store | 8082 | 16237485 | HealthRankings > BPD
- LiveRamp Data Store | 8082 | 16237395 | HealthRankings > BPH
- LiveRamp Data Store | 8082 | 16237478 | HealthRankings > Breast Cancer
- LiveRamp Data Store | 8082 | 24900788 | HealthRankings > Breast Cancer Caregivers
- LiveRamp Data Store | 8082 | 16237416 | HealthRankings > Cholesterol



- LiveRamp Data Store | 8082 | 16237450 | HealthRankings > Cough/Cold
- LiveRamp Data Store | 8082 | 16237432 | HealthRankings > Diabetes
- LiveRamp Data Store | 8082 | 16237508 | HealthRankings > Diabetes Type II
- LiveRamp Data Store | 8082 | 16237498 | HealthRankings > Eating Disorder

Zu den Segmenten mit Gesundheitsbezug gesellen sich zahlreiche Segmente, die auf die Religion verweisen, etwa *muslim* oder *jewish*, und solche, die auf die sexuelle Orientierung von Menschen oder ihre Herkunft und Ethnizität schließen lassen. Auch viele Segmente zu politischen Themen enthält die Liste: Wer ist für Donald Trump und wer gegen ihn? Wer ist für oder gegen *Black Lives Matter* und wer gegen das Recht auf Schwangerschaftsabbruch?

Tracking-Forscher Wolfie Christl sieht in derart umfassenden Datensammlungen gleich mehrere Probleme. Da sei zum Beispiel das Eindringen in die Privatsphäre aller Menschen, die in den Segmenten landen. „Kein Mensch kann damit rechnen, dass alle Informationen über das eigene Verhalten bei hundert Firmen landen.“ Christl nennt das eine „Verletzung der kontextuellen Integrität“. Der Kontext, in dem Menschen ihre Daten preisgegeben haben – möglicherweise sogar bewusst, indem sie an einer Umfrage teilgenommen oder bei einer App ihr Gehalt angegeben haben –, sei ein anderer Kontext als der, in dem die Informationen genutzt werden.

Das führe nicht nur zu Kontrollverlust, sondern habe auch ganz konkrete Auswirkungen auf das Leben von Menschen. „Die

Werbefirmen nutzen das Wissen über uns, um uns zu manipulieren und unser Verhalten zu formen“, so Christl. Ganz bewusst würde manche Online-Werbung persönliche Schwächen ausnutzen. Es sei etwa an der Tagesordnung, dass verletzte Gruppen gezielt angesprochen werden, Menschen mit Spielsucht, Depression oder anderen Krankheiten. Die Segmente werden zudem im Marketing auch immer wieder genutzt, um Menschen gezielt von Informationen auszuschließen. Bei Werbung für Wohnungen oder Jobs zum Beispiel kommt es immer wieder zu Diskriminierung, auch Fälle von irreführender oder betrügerischer Werbung, sogenanntes *Scamvertising*, kann mit den Daten befeuert werden.

„Wie die Datenfirmen uns einsortieren, entscheidet darüber, was wir zu sehen bekommen oder und sogar welche Preise oder Rabatte wir bekommen. Das prägt unsere Welt und es prägt auch unser Selbstbild, wenn wir von Werbung immer auf eine bestimmte Art und Weise angesprochen werden“, sagt Christl. Der auf Datenschutz spezialisierte Rechtsanwalt Tillmann Herbrich sieht darin insgesamt eine Gefahr für das demokratische Gemeinwesen. „Durch die unmittelbare Beeinflussung der subjektiven Nutzerwahrnehmung lässt sich auch die Wahrnehmung des Weltgeschehens beeinflussen“, so Herbrich. Dabei gehe es auch um die Platzierung von Fake News und die datengetriebene Verstärkung von Desinformation.

## Deutsche Nutzer, die abnehmen wollen

Wieviele IDs von Menschen aus Deutschland und anderen EU-Ländern sich in den Segmenten befinden, lässt nicht genau sagen. Klar ist: Die Xandr-Liste enthält Segmente mit Bezug zu allen Weltregionen. Ein Schwerpunkt liegt auf den USA und Europa, doch auch zu Ländern in Südamerika und Asien finden sich Segmente in der Liste. Das ist erkennbar an Länderkennzeich-

NETZPOLITIK.ORG

Advertising-ID 74CD18DE2-38KD-01D2  
Demographic: Conservative Retiree  
Browser Language: Arabic  
Disease Propensity by Type > Cancer: Lung (Adstra)  
Personal Finance: Insurance Seekers  
Health and Wellness: Pessimistic Future  
Health Outlook and Stressed

Advertising-ID 6AB01CF8-45C5-44B1  
Age: 40-49  
Household Income: High  
Composite Segment: Moms Who Shop Like Crazy  
Brand affinity: Gucci  
Travel intention: Paris

Advertising-ID: 4DB78KD8-77CF-63F3  
Age: 30-39  
Income: 2.500 - 4.000 €  
Marital Status: Divorced  
GfK-Shopper: Alcoholic Beverages  
Mastercard - Top Spending Geography: Casino and Gambling Activities

Advertising-ID 7EF08HK1-72AB-7331  
Age: 50-59  
Gender: Male  
Auto - Ownership: Toyota Landcruiser  
Top Mobile Apps: Grindr - Gay chat  
Churches and Religious Groups: 1st and 2nd Management Level

Microsofts Datenmarktplatz Xandr  
**Mehr als 650.000 Kategorien für Werbezielgruppen im Angebot**

Die gezeigten Zielgruppensegmente entstammen einer Angebotsliste des Xandr Marketplace von Mai 2021. Die Zuordnung zu einzelnen Personen erfolgt hier beispielhaft, die Advertising-IDs sind fiktiv. Microsoft hat Xandr im Jahr 2022 gekauft.



CC BY-SA 4.0 netzpolitik.org

nungen in den Segmentnamen, etwa *US*, *DE* oder *Germany*. Mehr als 3.000 Kategorien mit explizitem Bezug zu Deutschland sind in der Liste, angeboten von Datenhändlern aus Deutschland und den USA.

Von den Segmenten mit sensiblen Daten haben zahlreiche einen klar erkennbaren Bezug zu den USA. Dort lassen sich offenbar selbst Menschen im Umkreis von Militärbasen oder Besucher:innen konkreter Wahlkampfveranstaltungen targeten. Viele kritische Segmente haben jedoch keine klare Länderzuordnung. Wir wollten von Xandr wissen, ob es ausschließen kann, dass sich hierin auch IDs von EU-Bürger:innen befinden. Wir erhielten keine Antwort.

Einige sensible Segmente haben im Namen jedoch auch einen klaren Bezug zu Deutschland und anderen EU-Ländern. Dazu zählen Segmente mit Informationen über Kasino-Besuche, ein Faible für Sportwetten oder gar Spielsucht. Die deutsche Firma Adsquare bot mit Deutschlandbezug Segmente wie *poor* oder zur Altersgruppe *13-18 years*. Die Berliner Firma gehört mit einem Angebot von mehr als 15.000 unterschiedlichen Segmenten zu den größeren Anbietern in der Datei. Dazu zählen auch Kategorien wie *Period Trackers*, *Conservative Retiree*, *Multicultural Families* oder Besuche bei den Geldautomaten von verschiedenen Banken, etwa der Sparkasse. Auf Anfrage sagt Adsquare, die hier genannten Segmente nicht mehr anzubieten.

Auch die ProSiebenSat.1-Tochter The ADEX hatte Segmente mit Gesundheitsbezug im Angebot, etwa *Alternative Medicine*, *Menopause* oder *Sleep Disorders*. Außerdem bot sie offenbar ein Segment zu Angestellten von Kirchen und religiösen Gruppen an. Die Telekom-Tochter emtriq wiederum bot offenbar Segmente wie *Hobbies and Interests: Erotic* und *pregnant* an. Die deutsche Firma Zeotap hatte ein Segment *LGBT* mit der

Länderkennzeichnung *ES* für Spanien im Programm. Was die deutschen Datenhändler dazu sagen, lest ihr in unserem Artikel *Wie deutsche Firmen am globalen Geschäft mit unseren Daten verdienen*<sup>12</sup>.

## Wie Xandr und das Programmatic Advertising System funktionieren

Wie die Segmente im Einzelnen entstehen, also aus welchen Quellen die Daten stammen und wie die Zusammensetzung berechnet wird, wollte uns keine der von uns angefragten deutschen Firmen erklären. Gesichert ist: Neben Segmenten mit harten Fakten wie dem Alter oder besuchten Orten und Websites gibt es zahlreiche Kategorien, die auf Schlussfolgerungen und Schätzungen beruhen. Es gibt berechtigte Zweifel an der Genauigkeit und Qualität der Daten, erklärt Tracking-Forscher Wolfie Christl. Dennoch würden sie genutzt, um zu entscheiden, was Menschen zu sehen bekommen.

Welchen Weg manche Daten gehen, lässt sich gut am Beispiel der Ratgeber-Seite *gutefrage.net* erklären. Das Unternehmen hat uns gegenüber bestätigt, seit 2019 Daten über die Interessen seiner Nutzer:innen an die ProSiebenSat.1-Tochter The ADEX weiterzugeben. „Wir erheben ‚Interest‘-Daten basierend auf unseren Frage-Themen auf *gutefrage.net* für Targeting-Zwecke, die wir an Werbetreibende weitergeben.“ Es sind Segmente, die Auskunft über Interessen der Nutzer:innen geben. Wer sich für Bücher und Literatur interessiert zum Beispiel, wer für die Zubereitung von Desserts, wer für Motorräder, wer für Autos allgemein und wer für Autoteile. Auch die Interessen *Alternative Medizin* und *Abnehmen* finden sich in der Liste, mit *gutefrage.net* als Quelle.

Wir wollten von gutefrage.net wissen, in welcher Form die Daten weitergegeben werde. Das Unternehmen antwortet, bei dem Prozess würden „keine personenbezogenen oder soziodemografischen Daten erhoben“. Das allerdings steht im Widerspruch zum Cookie-Banner der Website. Auf diesem heißt es „Wir erheben personenbezogene Daten (z.B. IP-Adresse, Pseudonymisierte Identifier, Emailadresse) und übermitteln diese auch Drittanbietern wie The Tradedesk und Liveramp, die uns helfen, unser Webangebot zu verbessern und zu finanzieren.“ Übersetzt: Wer die Website besucht und nicht 2,99 Euro monatlich zahlt, sondern beim Cookie-Banner auf *Akzeptieren* klickt, landet wohl mit einer ID in einer dieser Interessenkategorien. Der Datenhändler The ADEX bietet diese Segmente dann auf verschiedenen Plattformen zur Nutzung an. Etwa bei Xandr und auf sogenannten *Demand Side Platforms* (DSP). Werbekunden müssen sowohl der DSP als auch dem Datenbieter eine Gebühr zahlen.

Auf der anderen Seite des Werbekreislaufes sind die Anbieter freier Werbefläche, zum Beispiel auf Nachrichtenwebsites. Sie tun dies über eine *Supply Side Platform* (SSP). Die freien Werbefläche und die Informationen über die gewünschten Zielgruppe kommen über Auktionen auf AdExchange-Plattformen zusammen. Die DSP übermittelt die gewünschte Zielgruppe an die AdExchange. Wenn nun jemand besagte *Nachrichtensite* besucht, sendet die Seite die IDs des Besuchers an die SSP. Diese wiederum leitet eine Auktion ein, indem sie einen Bid-Request mit der ID des Nutzers und Informationen über den Werbefläche an die AdExchange schickt. Dort wird erkannt, ob der Besucher der Website zur gewünschten Zielgruppe des Kunden passt, sich zum Beispiel für das Thema Abnehmen interessiert. Wenn dies der Fall ist, bietet die DSP auf den Werbefläche. Die Meistbietenden gewinnen, sie können ihre Werbung bei der gewünschten Zielgruppe anzeigen.

## Die Berliner Datenschutzbeauftragte ist skeptisch

Die Erkennung der passenden Nutzer:innen und die Auktionen finden in wenigen Millisekunden statt. Technisch ist dieses System eine Meisterleistung, rechtlich steht es auf tönernen Füßen. Erst Anfang Juni kam das Bundeskartellamt am Ende einer mehrjährigen Untersuchung zu dem Ergebnis<sup>13</sup>, dass der Markt des Programmatic Advertising in hohem Maße intransparent ist. Das gelte sowohl gegenüber beteiligten Unternehmen und Werbekunden als auch gegenüber Nutzer:innen. Die Marktdominanz vor allem von großen Digitalkonzernen wie Google verschärfe die Problematik. „Wir sollten uns ernsthaft fragen, ob wir quasi den ‚gläsernen‘ Internetnutzer wollen, allein aus dem Grund, dass wir bestimmte Produkte oder Serviceleistungen kaufen sollen“, schlussfolgert Kartellamtspräsident Andreas Mundt.

Auch der Datenschutz macht der Branche Probleme. Denn die Werbeflächen brauchen laut Datenschutzgrundverordnung (DSGVO) eine rechtliche Grundlage für ihr Geschäft. Dafür komme nur die informierte und freiwillige Einwilligung der Betroffenen in Frage, sagt die Chefin der Berliner Datenschutzaufsichtsbehörde Meike Kamp zu netzpolitik.org. Lange wurde darum gerungen, ob auch andere Rechtsgrundlagen in Frage kommen, inzwischen sind sich die meisten Aufsichtsbehörden einig: Nein.

Die Datenhändler sagen, dass für das Einholen der Einwilligung ihre Datenpartner zuständig sind. Auf Websites nimmt das heute fast immer die Form von Cookie-Bannern an, auch bei der Installation von Apps muss man meist die Datenschutzbestimmungen akzeptieren. Wer hier zustimmt, muss damit rechnen, dass die Daten bei dutzenden oder gar hunderten Werbefirmen landen. Und genau das ist das Problem, sagt Jurist Marco Blocher von der gemeinnützigen Organisation None of Your Business (NOYB). „Kein Mensch kann da noch überschauen, wo die eigenen Daten landen und was dann genau damit passiert, vor allem weil diese Werbefirmen die Daten abermals an eine Vielzahl von Empfängern weitergeben.“

Es fehle bei den gängigen Einwilligungs-Bannern an der notwendigen Transparenz, pflichtet Rechtsanwalt Tilman Herbrich bei: „Aufgrund der Komplexität der Infrastruktur, der hohen Anzahl von involvierten Drittanbietern, der Vielzahl an abgefragten Verarbeitungszwecken und dem Umfang der Nutzerprofile werden die Anforderungen an die Informiertheit einer Einwilligungserklärung kaum erreicht.“

Auch die Berliner Aufsichtschefin Meike Kamp kommt zu diesem Schluss: „Die komplexe Struktur der beteiligten Akteure sowie die spezifischen Datenflüsse“ nachzuvollziehen, sei für den Einzelnen schwierig. Die Auswirkungen der Einwilligungentscheidung seien kaum nachvollziehbar. „Eine tatsächlich selbstbestimmte und informierte Einwilligung wird somit praktisch unmöglich.“ Der Transparenzgedanke der DSGVO bleibe auf der Strecke.

Hinzu käme, dass viele Einwilligungsbanner<sup>14</sup> noch immer mit Tricks und manipulativem Design<sup>15</sup> arbeiten, um an Einwilligungen zu kommen, ergänzen Herbrich und Blocher. „Viele Consent-Ausgestaltungen auf Websites und in Apps gewähren Nutzern keine gleichwertigen Auswahlmöglichkeiten, vor allem in Form und Farbgebung, für die Erteilung und Ablehnung einer Einwilligung“, so Herbrich. Deshalb fehle es an der Freiwilligkeit.

## Wer ist eigentlich verantwortlich?

Wir haben die deutschen Datenfirmen Adsquare, emtriq und The ADEX nach der Rechtsgrundlage für ihre Datenverarbeitung gefragt. Sie alle verweisen auf die Einwilligung und darauf, dass ihre Datenlieferanten dafür zuständig sind, sie einzuholen. Man lasse sich vertraglich zusichern, dass dies korrekt geschehe, gelegentlich gebe es stichprobenartige Überprüfungen.

Doch damit machten es sich die Datenhändler zu leicht, sagt Marco Blocher von NOYB. „Sie müssten die eingeholten Einwilligungen dokumentieren, verwalten, Widerrufe berücksichtigen.“ Sich vom Datenlieferanten eine vertragliche Zusicherung einzuholen, dass alles DSGVO-konform sei“, reiche nicht aus. Man könne schließlich als Arzneimittelhersteller auch nicht sagen: „Der Lieferant hat vertraglich gesagt, das ist diese und jene Chemikalie, ich habe das deshalb ungeprüft in der Medikamentenherstellung verwendet.“ Dass die Anforderungen im derzeitigen Werbesystem umgesetzt werden, bezweifelt Blocher.

Bei den Daten, die derzeit von Websites gesammelt und an Werbefirmen weitergegeben werden, setzen fast alle Firmen auf

einen technischen Standard. Das sogenannte *Transparency and Consent Framework* (TCF) wird von der Branchenorganisation IAB bereitgestellt. Es soll sicherstellen, dass die Einwilligung standardisiert und rechtsicher eingeholt wird und dass Daten im Werbesystem nicht ohne Einwilligung verarbeitet werden. Doch ob der Standard rechtssicher ist, daran gibt es große Zweifel.

Die belgische Datenschutzaufsicht hatte den TCF-Standard 2022 für ungültig erklärt<sup>16</sup>, unter anderem, weil er die Anforderungen an die Transparenz nicht erfüllt. Inzwischen nutzen Firmen eine angepasste Version des TCF, doch auch diese werde den Anforderungen der DSGVO nicht ohne Reduktion der Verarbeitungszwecke und Empfänger gerecht, so Herbrich. Seiner Meinung nach müssten Nutzer:innen beispielsweise darüber aufgeklärt werden, in welche Kategorien sie einsortiert werden, wie viele Firmen ihre Daten erhalten und wie genau diese genutzt werden.

Die Werbe- und Datenbranche sieht das anders. Höchststrichterliche Urteile zu Datenhandel und dem System des Programmatic Advertising gibt es noch nicht.

### Keine sensiblen Informationen?

Für die datenschutzrechtliche Beurteilung dürfte besonders wichtig sein, ob auch sensible Daten nach Artikel 9 der DSGVO<sup>17</sup> verarbeitet werden. Dazu zählen Gesundheitsdaten und solche, aus denen sich Schlüsse über politische Überzeugungen, sexuelle Orientierung, Religion und Herkunft ziehen lassen. Bei solchen Daten gelten für die Einwilligung nochmal erhöhte Anforderungen, erklärt Behördenchefin Meike Kamp. Diese müsse ausdrücklich erfolgen, also mit ausführlicheren Informationen über die Daten, sagt Tilman Herbrich. „Eine ausdrückliche Einwilligung kann man nicht mit einem Klick in einem Cookie-Banner auf ‚Alle akzeptieren‘ erklären.“ Auch die Daten von Minderjährigen unter 16 Jahren sind besonders geschützt, in vielen Ländern braucht es die Einwilligung der Eltern.

Die Datenhandelsfirmen stellen sich auf den Standpunkt, sie würden keine sensiblen Daten verarbeiten, sondern lediglich allgemeine Interessen. So erklärt etwa die Pressesprecherin der ProSiebenSat.1-Tochter The ADEX, dass sie zu den Details einzelner Segmente nichts sagen könne, weil die Verträge mit Datenpartnern Verschwiegenheitsklauseln enthielten. Wir hatten unter anderem gefragt, wie sich die ADEX-Segmente *Menopause*, *Corona Virus*, *Sleep Disorders* und *Churches and Religious Groups – 1st and 2nd Management Level – GPS Data* zusammensetzen.

Die Sprecherin antwortet auf einer grundsätzlichen Ebene: Die Bezeichnung der Segmente spiegele „ausschließlich die Kategorienart der besuchten Website wider, ohne dass damit eine Aussage über zum Beispiel die körperliche oder geistige Gesundheit eines Betroffenen verbunden ist“. Als Beispiel nennt die Sprecherin das Segment *Sports Cars*, in das man einsortiert werde, wenn man einen Medienbericht über einen Ferrari aufgerufen habe. „Hieraus die Aussage, der Betroffene wäre Eigentümer eines Ferraris oder sonstigen Sports Car, abzuleiten, ist erkennbar unzutreffend und auch nicht intendiert.“

Dieser Widerspruch begegnet einem immer wieder in der Branche: Einerseits sollen die Daten so genau sein, dass sie perfektes Targeting garantieren, andererseits sollen die Daten so allgemein sein, dass es keine Probleme mit dem Datenschutz gibt. Auch die Telekom-Tochter Emetriq schreibt uns, das Unternehmen verarbeite keine Daten, die zu den besonderen Kategorien nach Artikel 9 DSGVO gehören: „Unsere Segmentdaten beruhen nicht zum Beispiel auf Gesundheitsdaten, wie etwa medizinischen Diagnosen, Einkaufsdaten zu Medikamenten, sondern auf das jeweiligen Surf-Verhalten, das ein allgemeines Interesse zu bestimmten Themen abbildet.“

Adsquare wiederum schreibt auf unsere Frage, ein Segment namens *Multicultural Families* basierte lediglich „auf ortsbezogenen Daten von Drittanbietern“ und sei nicht mehr im Angebot. Zu einem Segment namens *Viagra – Unhealthy Place Visits* sagt Adsquare, der Name der Segmente sei von Kunden frei wählbar und müsse „nicht notwendigerweise mit dem Inhalt des Segments übereinstimmen. Er kann sich auch auf das beworbene Produkt beziehen oder ein Code-Name für ein Projekt sein.“ Grundsätzlich erklärt das Unternehmen, „dass Adsquare keine Segmente mit Bezug zum Gesundheitszustand oder Medikamenten oder Ähnlichem von Personen im Angebot hat.“

Tilman Herbrich hält diese Argumentation für problematisch und verweist auf eine Entscheidung des Europäischen Gerichtshofes (EuGH)<sup>18</sup>. „Es reichen Angaben in einem Formular, aus denen sich Ableitungen über die sexuelle Orientierung entnehmen lassen, um in den Anwendungsbereich von Artikel 9 zu kommen“, so Herbrich. Der EuGH verdeutliche damit, dass bereits Ableitungen als Resultat eines Denkvorgangs ausreichend seien, um eine Verarbeitung über besondere Kategorien personenbezogener Daten anzunehmen. Für Werbeplattformen folge hieraus, dass Segmentbildungen mit Bezügen zur sexuellen Orientierung, Religion, politischen Überzeugung, Gewerkschaftszugehörigkeit oder Gesundheit nur mit ausdrücklicher Einwilligung zulässig sind.

### Politische Lösungen sind nicht in Sicht

„Der wichtigste Trick im Datenhandel ist die Behauptung aller Beteiligten, man wäre nur im Auftrag von anderen tätig und diese anderen wären verantwortlich“, konstatiert Aktivist und Forscher Wolfie Christl. „Jede Verantwortlichkeit verschwindet ins Nirvana. Das ist natürlich lächerlich, aber trägt dazu bei, die Durchsetzung der DSGVO zu verzögern.“ Dieser gordische Knoten könne nur durchschlagen werden, „wenn die Aufsichtsbehörden einen systemischen Blick auf den gesamten Echtzeit-Datenhandel werfen.“

Dass das bislang noch nicht geschehen ist, liegt auch an einem Konstruktionsproblem der Datenschutzaufsicht. Die Behörden werden in der Regel dann aktiv, wenn es Beschwerden von Bürger:innen gibt. Doch über Firmen, die man nicht kennt, kann man sich schlecht beschweren. „Die geringe Anzahl an Beschwerden zu der Thematik verweist auch auf die Intransparenz des Geschäfts der Datenhändler“, erklärt deshalb die Berliner Datenschutzbeauftragte Meike Kamp. Ihr Haus lege den Fokus bisher eher auf Websites, die ungültige Einwilligungen einholen, als auf die Firmen im Hintergrund.

Wolfie Christl hat deshalb einen dringenden Appell für die Aufsichtsbehörden: „Die Öffentlichkeit weiß zwar jetzt mehr als zuvor, aber ich vermute, auf den Servern dieser Firmen schlummern noch viel mehr Belege für möglicherweise illegale Datenpraktiken“, vermutet er. An solche Informationen kämen nur die Aufsichtsbehörden ran. „Die Firmen, die alles über uns wissen wollen, berufen sich auf ihre Geschäftsgeheimnisse. Die deutschen Datenschutzbehörden müssen endlich aktiv werden und diese unverantwortlichen Datenpraktiken effektiv beenden.“

Die Berliner Datenschutzbeauftragte Meike Kamp sieht bei dem Thema auch politischen Handlungsbedarf. „Die geschilderten Beispiele zeigen, dass über flankierende gesetzliche Maßnahmen nachgedacht werden sollte, um ein gesetzliches Umfeld zu schaffen, in dem die Einwilligung auch tatsächlich Wirkung entfalten kann.“ Zum Beispiel könnten der Grad an Individualisierung der angesprochenen Personen und Gruppen sowie die Anzahl der erlaubten Datenkategorien eingeschränkt werden.

Die Datenschutzprobleme des wohl wichtigsten Internet-Geschäftsmodells sind seit langem bekannt. Politische Lösungen, wie sie Meike Kamp fordert, sind jedoch nicht in Sicht. Eine Initiative von EU-Abgeordneten, „Überwachungswerbung“ grundsätzlich zu verbieten, ist im letzten Jahr gescheitert<sup>19</sup>.

Dabei machen die Kritiker des Targeted Advertising immer wieder deutlich: Es gibt Alternativen<sup>20</sup>. Kontextbasierte Werbung zum Beispiel. Sie ist nicht auf das Profil der Nutzer:innen zugeschnitten, sondern auf das Umfeld, in dem sie erscheint. Besucht jemand eine Website mit einem Ratgebertext zu Autoreifen, bekommt er entsprechende Anzeigen ausgespielt. Das würde nicht nur dem Datenschutz helfen, sondern auch die Umweltbilanz von digitaler Werbung verbessern: Weniger Daten bedeutet weniger benötigte Rechenkapazität und somit weniger Stromverbrauch.

Doch ein freiwilliger Paradigmenwechsel scheint in der Branche nicht in Sicht. Dass alles, was wir im Internet tun, aufgezeichnet und ausgewertet wird, bleibt vorerst Standard. Nutzer:innen bleiben nur zwei Dinge: Einerseits Selbstschutz durch Maßnahmen wie das Löschen von Cookies. Andererseits Datenauskunftsanfragen an Datenhändler, um Grundlagen für Beschwerden bei Aufsichtsbehörden zu schaffen. Wie das geht, erklären wir in einem Tutorial: *So findest du heraus, was Datenhändler über dich gespeichert haben*.<sup>21</sup>

*Weißt du mehr über fragwürdige Datenpraktiken in der Marketing-/Adtech-Industrie? Wir würden uns freuen, von dir zu hören. Auf folgendem Weg kannst du uns sicher kontaktieren: [ingo.dachwitz@netzpolitik.org](mailto:ingo.dachwitz@netzpolitik.org) / Signal-Kontakt auf Anfrage. Bitte nicht vom Dienstrechner/-telefon oder dienstlichem Account schreiben.*

Mitarbeit: Anne-Lena Schmierer<sup>22</sup> und Jan Lutz<sup>23</sup>.

Quelle: <https://netzpolitik.org/2023/microsofts-datenmarkt-platz-xandr-das-sind-650-000-kategorien-in-die-uns-die-online-werbeindustrie-einsortiert/>

## Anmerkungen

- 1 <https://themarkup.org/privacy/2023/06/08/from-heavy-purchasers-of-pregnancy-tests-to-the-depression-prone-we-found-650000-ways-advertisers-label-you>
- 2 [https://web.archive.org/web/20230525225839/https://docs.xandr.com/en-US/bundle/monetize\\_monetize-standard/page/topics/data-marketplace-buyer-overview.html](https://web.archive.org/web/20230525225839/https://docs.xandr.com/en-US/bundle/monetize_monetize-standard/page/topics/data-marketplace-buyer-overview.html)
- 3 [https://web.archive.org/web/20230525225541mp\\_/https://xandr-be-prod.zoominsoftware.io/bundle/monetize\\_monetize-standard/page/attachments/data-marketplace-buyer-overview/data\\_marketplace\\_public\\_segments\\_pricing\\_05212021.xlsx](https://web.archive.org/web/20230525225541mp_/https://xandr-be-prod.zoominsoftware.io/bundle/monetize_monetize-standard/page/attachments/data-marketplace-buyer-overview/data_marketplace_public_segments_pricing_05212021.xlsx)
- 4 <https://www.oberlo.com/statistics/digital-ad-spend>
- 5 <https://www.businessinsider.com/microsoft-potential-adtech-acquisitions-that-could-build-advertising-business-2022-6?op=1>
- 6 <https://www.forbes.com/sites/daviddoty/2022/01/27/microsoft-has-quietly-become-the-next-ad-tech-giant-marketers-ignore-its-reach-at-your-peril/?sh=6cb6495e5312>
- 7 <https://www.adzine.de/techfinder/xandr/>
- 8 <https://www.adzine.de/techfinder/xandr/>
- 9 <https://www.theverge.com/2020/6/25/21303190/congress-data-smartphone-tracking-fbi-security-privacy>
- 10 <https://techcrunch.com/2022/08/22/oracle-us-privacy-class-action/>
- 11 <https://docs.xandr.com/bundle/data-providers/page/managing-public-and-private-custom-segments.html>
- 12 [https://netzpolitik.org/2023/adsquare\\_theadex\\_emetriq\\_werbettracking-wie-deutsche-firmen-am-geschaefft-mit-unsere-daten-verdienen/](https://netzpolitik.org/2023/adsquare_theadex_emetriq_werbettracking-wie-deutsche-firmen-am-geschaefft-mit-unsere-daten-verdienen/)
- 13 [https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Sektoruntersuchungen/Sektoruntersuchung\\_Online\\_Werbung\\_Abschlussbericht.html](https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Sektoruntersuchungen/Sektoruntersuchung_Online_Werbung_Abschlussbericht.html)

Ingo Dachwitz

**Ingo Dachwitz** ist Kommunikationswissenschaftler und seit 2016 Redakteur bei netzpolitik.org. Er schreibt und spricht über Datenpolitik, Überwachungskapitalismus und den digitalen Strukturwandel der Öffentlichkeit. Ingo moderiert Diskussionen, gibt Workshops für junge und ältere Menschen in digitaler Selbstverteidigung und lehrt gelegentlich an Universitäten zur politischen Ökonomie digitaler Medien. Er ist Ko-Autor der Studie *Medienmäzen Google*<sup>24</sup> über Journalismusförderungen des Konzerns. Ingo ist Mitglied des Vereins Digitale Gesellschaft.

**Kontakt:** E-Mail<sup>25</sup> (OpenPGP<sup>26</sup>), Mastodon<sup>27</sup>, Twitter<sup>28</sup>, FragDenStaat<sup>29</sup>

- 14 <https://netzpolitik.org/2022/cookie-banner-der-meistbesuchten-websites-miese-tricks-und-fiese-klicks/>
- 15 <https://netzpolitik.org/2022/edpb-dark-patterns-datenschutzbehoerden-warnen-vor-manipulativem-design-im-netz/>
- 16 <https://netzpolitik.org/2022/datenschutzgrundverordnung-wichtiger-baustein-fuer-cookie-banner-ist-illegal/>
- 17 <https://dsgvo-gesetz.de/art-9-dsgvo/>
- 18 <https://curia.europa.eu/juris/document/document.jsf?text=&docid=263721&pageIndex=0&doclang=DE&mode=req&dir=&occ=first&part=1&cid=1837628>
- 19 <https://netzpolitik.org/2022/digitale-dienste-gesetz-wie-die-datenindustrie-ein-verbot-von-ueberwachungswerbung-verhinderte/>
- 20 <https://netzpolitik.org/2022/tracking-und-targeting-wir-brauchen-ein->

*umdenken-bei-der-online-werbung/*

- 21 <https://netzpolitik.org/2023/auskunftsfragen-so-findest-du-heraus-was-datenhaendler-ueber-dich-gespeichert-haben>
- 22 <https://netzpolitik.org/author/anna-lena-schmierer/>
- 23 <https://netzpolitik.org/author/jan-lutz/>
- 24 <https://www.otto-brenner-stiftung.de/wissenschaftsportal/informationsseiten-zu-studien/studien-2020/medienmaezen-google/>
- 25 <mailto:ingo.dachwitz@netzpolitik.org>
- 26 <https://keys.openpgp.org/search?q=ingo.dachwitz@netzpolitik.org>
- 27 <https://mamo.fr/@roofjoke>
- 28 <https://twitter.com/roofjoke>
- 29 <https://fragdenstaat.de/profil/i.dachwitz/>



Bianca Kastl

## Effizienz ist ein konservatives Wesen

13. August 2023 – *Damit Verwaltung moderner wird, soll sie effizienter werden, heißt es oft und gerne. Damit aber ist es nicht getan. Denn Effizienz erfordert mehr Fachwissen, konserviert den Status quo und verliert die Nutzer:innen aus dem Blick.*

Die heutige Episode von Digitalisierung leiten wir am besten mit einer total unerwarteten Headline ein: „Internetdienste deutscher Behörden schneiden international schlecht ab“<sup>1</sup>. Ach. Echt? Schon wieder? Das ist die wesentliche Erkenntnis des *Digital Government Citizen Survey* der Boston Consulting Group, die die Zufriedenheit der Bürger:innen mit digitalen Behördendiensten bewertet hat. Deutschland ist da aktuell eher so semi-gut gelegen und befindet sich laut der Studie auf dem drittletzten<sup>2</sup> Platz von insgesamt 41 untersuchten Ländern.

Glücklicherweise liefert die Unternehmensberatung gleich die vermeintliche Lösung für alle Probleme der Verwaltungsdigitalisierung mit, acht konkrete Punkte, um genau zu sein. Darunter so wenig überraschende Aspekte wie eine stärkere Nutzer:innen-Zentrierung für eine bessere Nutzendenerfahrung. Das kennen wir ja schon aus dem Problemfeld Digitalisierung des Gesundheitswesens<sup>3</sup>.

Darüber hinaus empfehlen sie gleich mehrfach, die Effizienz der Verwaltung zu optimieren. Da sollen Potenziale „zur Effizienzsteigerung durch Automatisierung“ gehoben werden und der Einsatz sogenannter Künstlicher Intelligenz (KI) soll den Zu-

gang zu Informationen verbessern – natürlich „unter Beachtung strenger Regeln einer Verantwortungsvollen KI“.

Das alles klingt zunächst nachvollziehbar und könnte auch den Fachkräftemangel beheben und den drohenden Personalkollaps verhindern<sup>4</sup>.

Allerdings ist das mit der Effizienz so eine Sache. Denn Effizienz ist leider ein sehr konservatives Wesen. Und das Streben allein nach Effizienz ist nicht selten der Anfang einer ganzen Kette von Problemen. Paradox eigentlich, oder?

### Die Ironie der Automation

Im Jahr 1983 veröffentlichte die Kognitionspsychologin Lisanne Bainbridge eine Studie mit dem Titel „Ironies of Automation“, die auf ein gewisses Automationsparadox hinweist.

Darin<sup>5</sup> heißt es:

*By taking away the easy parts of [the] task, automation can make the difficult parts of the human operator's task more difficult.*

Nehmen wir an, es gibt ein System mit einem starken Automatisierungsgrad. Dann gibt es in einem Prozess am Ende zwar weniger einfache Aufgaben, die Menschen übernehmen können. Die noch verbleibenden werden dann aber umso schwieriger und benötigen unter Umständen Spezialwissen.

Das zeigt sich auch an Werkzeugen wie der Robotic Process Automation, die heute oftmals als vermeintliche Lösung gefeiert werden. Sie sind so eine Art automatisierte Clickworker-Bots, die bestimmte Tätigkeiten mit bestehenden Programmen nach einem vorher definierten Ablauf abarbeiten.

Im Verwaltungsbereich können diese Bots etwa bei veralteter, aufwändig zu bedienender Fachsoftware zum Einsatz kommen,



*New york times journal, Presseraum, Foto: janeb13, pixabay*

die dazu dient, Bescheide für Sozialleistungen zu erstellen. Mit Hilfe von RPA-Tools kann diese Software, so zumindest das Versprechen, effizienter gemacht werden. Was vorher ein Mensch durchklicken musste, kann dann ein „Roboter“ im Hintergrund abarbeiten. Der ganze Prozess wird damit, trotz schlechter Softwarebasis, einfach und effizient. Das gelingt aber nur, wenn auch alles gut geht.

### Fachkräftemangel<sup>3</sup>

Geht jedoch etwas schief oder muss etwas verändert werden, was bei digitalen Systemen bekanntlich stets vorkommen kann, braucht es gleich zwei oder drei Expert:innen, die das automatisierte System anpassen.

Zunächst braucht es jemanden, der die Fehler des RPA-Systems behebt. Eine weitere Person wird dann gegebenenfalls benötigt, die das Grundsystem, das die Bescheide von Sozialleistungen erstellt, technisch versteht. Und am Ende braucht es eine dritte Person, die das Ergebnis – also zum Beispiel die Berechnungen der Bescheide – validiert. Solange diese Personen das Problem nicht gelöst haben, geht nichts mehr. Der ganze Prozess steht erstmal still. Unterm Strich wird der menschliche Eingriff in das System also insgesamt schwieriger und benötigt mehr Expertise.

Bei Systemen, die auf sogenannter KI basieren, kommt möglicherweise noch eine weitere Komplexitätsebene hinzu. Denn KI-Systeme bzw. deren Trainingsdaten und Ergebnisse sind nicht so einfach zu debuggen und deren Output wohl auch nie so vorhersehbar<sup>6</sup> wie das von „einfachen“ automatisierten Systemen.

Wenn Verwaltung nun in der Automatisierung die vorrangige Lösung sieht, sollte sie sich auch die Frage stellen, wie sie mit den neuen kritischen Abhängigkeiten von Spezialwissen und -systemen umgehen will. Schon heute ist es schwer, IT-Expertise für allgemeinere Tätigkeiten wie Systemadministration zu finden. In Zukunft wird es noch weitaus schwieriger sein, Expert:innen zu finden, die automatisierte Spezialsysteme aufrechterhalten oder neu entwickeln. Neue Lock-in-Effekte, die eine vereinfachte Betrachtung von Effizienz oftmals ignoriert.

### Never run an unchangeable system

Doch RPA oder sogenannte KI-Systeme weisen noch ein weiteres Problem auf: Sie konservieren den Zustand der Ausgangssysteme beziehungsweise den Stand der Ausgangsdatenbasis, mit denen KI-Systeme meist aufwändig trainiert werden.

Aber wer wollte das schon leichtfertig ändern? Verwaltungs-IT funktioniert eher nach der Prämisse „never change a running sys-

tem“ – bloß nichts anpassen, was einmal läuft. Mehr oder weniger gut laufende Systemen nicht verändern zu können oder zu wollen, ist in der Verwaltung leider zu einem Mantra geworden, das in Kombination mit dem Leitmotto „Haben wir schon immer so gemacht“ zur fast vollständigen Veränderungsstarre geführt hat.

Aufgesetzte Effizienztechnologien wie RPA oder sogenannte KI tun ihr Übriges, um tiefergehende Probleme an Systemen und Prozessen nicht angehen zu müssen. Alt-Systeme können ja – Achtung: Verwaltungsdeutsch! – „ertüchtigt“ werden. Keine medienbruchfreie Übermittlung von Antragsdaten? Egal, ertüchtigen wir mit KI, soll die doch irgendwie PDF-Dateien übersetzen.

Wie schwer sich Verwaltung mit Veränderungen tut, zeigt auch der mehr oder weniger beschränkte Erfolg, Bürokratie zu reduzieren. Die Bilanz der „One in, one out“-Regel, die vor Einführung einer neuen gesetzlichen Maßnahme eigentlich das Abschaffen an anderer Stelle vorsieht, ist in ihrer Wirkung laut Nationalem Normenkontrollrat<sup>7</sup> – Achtung: Wortwitz! – nur beschränkt effizient.

Stattdessen sollte die Devise für das System Verwaltung lauten: Never run an unchangeable system.

### Drucksache 9/2442

Effizienz ist per se nichts Schlechtes, nur scheint sie oftmals wichtige andere Aspekte wie Nutzer:innenzentrierung zu verdrängen.

Wer das jetzt aber für eine gänzlich neue Erkenntnis hält, dem sei die Lektüre eines Zwischenberichts der Enquete-Kommission „Neue Informations- und Kommunikationstechniken“ empfohlen, der ironischerweise aus dem Jahr 1983 stammt. Unter Drucksache 9/2442<sup>8</sup> des Deutschen Bundestages ist dort Folgendes zu lesen:

*Bislang erfolgte der Einsatz der IuK-Techniken in der öffentlichen Verwaltung in der Hauptsache zum Zweck der Effizienzsteigerung (Rationalisierung des Verwaltungshandelns). Andere Ziele, die sich an den Bedürfnissen und Interessen der Verwaltungsbediensteten und der Bürger sowie an einer Verbesserung der Qualität der Verwaltungsleistung insgesamt orientieren (Mitarbeiterzufriedenheit, Bürgernähe etc.), wurden zu wenig berücksichtigt.*

Verwaltung ist in den vergangenen dreißig Jahren in gewisser Hinsicht effizienter geworden. Dafür sorgte unter anderem die Einführung von Faxgeräten und Personal Computern, die Abschaffung von Schreibmaschinen sowie das Aufkommen des Internets.

**Bianca Kastl**

**Bianca Kastl** ist Entwicklerin und unterstützt seit Beginn der Corona-Pandemie Gesundheitsämter bei der Digitalisierung. Von dort aus schaut sie kritisch auf die digitalen Infrastrukturen, die im öffentlichen Gesundheitswesen genutzt werden – vor allem auf deren Schwachstellen.

Allerdings sind beim Einsatz neuer Technologien andere Aspekte zu wenig berücksichtigt worden. In gewisser Weise haben wir einen effizienten Bürokratie-Automatismus geschaffen, der aber auch im Jahr 2023 noch mit den Mitteln und Methoden von 1983 agiert.

Schwer nachvollziehbare Besuche im Amt gibt es immer noch, etwa wenn in Australien der Führerschein verloren geht<sup>9</sup>. Und noch immer sind auch heute noch viele Ausdrücke erforderlich<sup>10</sup> und finden sich etliche Medienbrüche an jeder Ecke. Auch gibt es weiterhin hierarchische Strukturen mit vielen Regeln und Vorschriften, obschon an der ein oder anderen Stelle Menschen im Home Office sitzen und Telex<sup>11</sup> und Teletex<sup>12</sup> der Vergangenheit angehören. Und damals wie heute ist es – wie schon der Zwischenbericht der Enquete-Kommission festhält – für „Außenstehende oft unüberschaubar, wo Informationen eingeholt werden können“.

Mit Effizienz allein ist es also nicht getan. Denn Effizienz für sich ist ein konservatives Wesen.

Quelle: <https://netzpolitik.org/2023/degitalisierung-effizienz-ist-ein-konservatives-wesen/>

## Anmerkungen

- 1 <https://www.spiegel.de/wirtschaft/digitale-verwaltung-deutsche-behoerden-schneiden-international-schlecht-ab-a-45da51fd-692a-467c-879a-f3d3cf01af86>
- 2 <https://www.bcg.com/publications/2023/germany-digital-government-citizen-survey>
- 3 <https://netzpolitik.org/2023/degitalisierung-vertrauen-ist-ein-weg-aus-vielen-kleinen-schritten/>
- 4 <https://www.spiegel.de/wirtschaft/fachkraeftemangel-gewerkschaften-warnen-vor-personalkollaps-im-oeffentlichen-dienst-a-5e9b4775-5763-4196-86f6-9494fac7e664>
- 5 <https://www.complexcognition.co.uk/2021/06/ironies-of-automation.html>
- 6 <https://fortune.com/2023/08/01/can-ai-chatgpt-hallucinations-be-fixed-experts-doubt-altman-openai/>
- 7 <https://www.tagesschau.de/inland/innenpolitik/buerokratie-bremse-100.html>
- 8 <https://dserver.bundestag.de/btd/09/024/0902442.pdf>
- 9 [https://www.youtube.com/watch?v=\\_S9jbe0kMrE](https://www.youtube.com/watch?v=_S9jbe0kMrE)
- 10 <https://www.rbb24.de/panorama/beitrag/2022/12/bafoeg-studenten-studierende-studium-uni-berlin-.html>
- 11 <https://de.wikipedia.org/wiki/Telex>
- 12 <https://de.wikipedia.org/wiki/Teletex>



## Wissenschaft & Frieden 3/2023

### Gesellschaft in Konflikt

#### Was es zu thematisieren gilt

Gesellschaften sind durchzogen von Konflikten – so auch die deutsche. In diesem Heft richten die Autor:innen einen kritischen Blick auf die vielschichtigen Konfliktlagen in unserer eigenen Gesellschaft. Konflikte im Inland zu thematisieren ist notwendig und gleichermaßen oft verschwommen – viele Konflikte verlaufen latent, eine Eskalation passiert oft nicht direkt sichtbar, Beachtung finden sie oft wenig.

Diese Ausgabe soll helfen, sich mit der Frage zu beschäftigen: Welche Konflikte wollen wir, aber welche können wir auch sehen – und wie können wir als Gesellschaft neue Betrachtungswinkel erlernen, um Konflikte frühzeitiger zu erkennen und präventiv zu intervenieren?

Mit Beiträgen von Lotta Mayer, Axel Salheiser und Christoph Richter, Ulrika Mientus, Maraike Henschel und Joschka Dreher und weiteren

Weitere Beiträge im Heft: Bieß – Konfliktsensible Netzwelt? | Bläsius/Siekmann – Gefährliche KI? | Bader – Strukturelle Nichtverteidigbarkeit

W&F 3/23 | August | 68 Seiten | 12 € (druck) / 9 € (ePUB+PDF) als Einzelheft oder im Abonnement  
[www.wissenschaft-und-frieden.de](http://www.wissenschaft-und-frieden.de)



### 40 Jahre W&F

Jubiläum W&F: 6. und 7. Oktober 2023, Bonn  
Symposium **Wissenschaft für den Frieden**  
Informationen:

[wissenschaft-und-frieden.de/projekt/40-jahre-jubilaem](http://wissenschaft-und-frieden.de/projekt/40-jahre-jubilaem)



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

Vierteljährlich erhalten Mitglieder die Fachzeitschrift Fiff-Kommunikation mit Artikeln zu aktuellen Themen, problematischen

## Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung

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

### FIFF online

#### Das ganze FIFF

[www.fiff.de](http://www.fiff.de)

Twitter FIFF e.V. – @Fiff\_de

#### Cyberpeace

[cyberpeace.fiff.de](http://cyberpeace.fiff.de)

Twitter Cyberpeace – @Fiff\_AK\_RUIN

#### Faire Computer

[blog.faire-computer.de](http://blog.faire-computer.de)

Twitter Faire Computer – @FaireComputer

#### Mitglieder-Wiki

<https://wiki.fiff.de>

### FIFF-Mailinglisten

#### FIFF-Mailingliste

An- und Abmeldungen an:

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

Beiträge an: [fiff-L@lists.fiff.de](mailto:fiff-L@lists.fiff.de)

#### FIFF-Mitgliederliste

An- und Abmeldungen an:

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

### FIFF-Beirat

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

### FIFF-Vorstand

**Stefan Hügel** (Vorsitzender) – Frankfurt am Main  
**Rainer Rehak** (stellv. Vorsitzender) – Berlin  
**Michael Ahlmann** – Kiel / Blumenthal  
**Maximilian Hagner** – Jena  
**Alexander Heim** – Berlin  
**Sylvia Johnigk** – München  
**Benjamin Kees** (Mitgliedschaft ruhend) – 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

### FIFF-Geschäftsstelle

**Ingrid Schlagheck** (Geschäftsführung) – Bremen

## Impressum

|                             |  |
|-----------------------------|--|
| <b>Herausgeber</b>          | Forum InformatikerInnen für Frieden und gesellschaftliche Verantwortung e. V. (FIfF)   |
| <b>Verlagsadresse</b>       | FIfF-Geschäftsstelle<br>Goetheplatz 4<br>D-28203 Bremen<br>Tel. (0421) 33 65 92 55<br><a href="mailto:fiff@fiff.de">fiff@fiff.de</a>   |
| <b>Erscheinungsweise</b>    | vierteljährlich  |
| <b>Erscheinungsort</b>      | Bremen   |
| <b>ISSN</b>                 | 0938-3476  |
| <b>Auflage</b>              | 1 300 Stück  |
| <b>Heftpreis</b>            | 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. |
| <b>Hauptredaktion</b>       | Dagmar Boedicker, Stefan Hügel (Koordination), Sylvia Johnigk, Hans-Jörg Kreowski, Dietrich Meyer-Ebrecht, Ingrid Schlagheck   |
| <b>Schwerpunktredaktion</b> | Dagmar Boedicker, Klaus Heß, Katharina Just  |
| <b>V.i.S.d.P.</b>           | Stefan Hügel   |
| <b>Retrospektive</b>        | Beiträge für diese Rubrik bitte per E-Mail an <a href="mailto:redaktion@fiff.de">redaktion@fiff.de</a>   |
| <b>Lesen, SchlussFIfF</b>   | Beiträge für diese Rubriken bitte per E-Mail an <a href="mailto:redaktion@fiff.de">redaktion@fiff.de</a>   |
| <b>Layout</b>               | Berthold Schroeder, München  |
| <b>Cover</b>                | Midjourney - A roboticist at work in her lab, 1950s  |
| <b>Druck</b>                | Meiners Druck, Bremen<br>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](http://www.fiff.de))

### 40 Jahre W&F in Bonn

Symposium Wissenschaft für den Frieden  
6. + 7. Oktober, German Institute of Development and Sustainability (IDOS)

### FIfF-Konferenz 2023 in Berlin

„Wir möchten gemeinsam mit euch allen in die Zukunft schauen!“  
3. – 5. November, Jugendherberge Berlin-Ostkreuz

### FIfF-Kommunikation

4/2023 „Wissenschaft für den Frieden“

Hans-Jörg Kreowski

Redaktionsschluss: 4. November 2023

1/2024 „FIfF-Konferenz 2023“

FIfF

Redaktionsschluss: 2. Februar 2024

2/2024 „40 Jahre FIfF“

FIfF

Redaktionsschluss: 3. Mai 2024

### Zuletzt erschienen:

4/2022 100 Jahre Joseph Weizenbaum

1/2023 #FIfFKon22

2/2023 Mensch – Gesellschaft – Umwelt ... und Informatik

### W&F – Wissenschaft & Frieden

1/23 Jenseits der Eskalation

2/23 Klimakrise

3/23 Gesellschaft in Konflikt

### vorgänge – Zeitschrift für Bürgerrechte und Gesellschaftspolitik

#239/240 Krieg und Frieden

#241 Demokratie und Rechtsstaat verteidigen

#242 Künstliche Intelligenz

### DANA – Datenschutz-Nachrichten

4/22 Beschäftigtendatenschutz

1/23 Europäische Entwicklungen

2/23 Europäische Entwicklungen, Teil 2

## 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](mailto:fiff@fiff.de)

Die Bürozeiten finden Sie unter [www.fiff.de](http://www.fiff.de)

### Bankverbindung

Bank für Sozialwirtschaft (BFS) Köln

Spendenkonto:

IBAN: DE79 3702 0500 0001 3828 03

BIC: BFSWDE33XXX

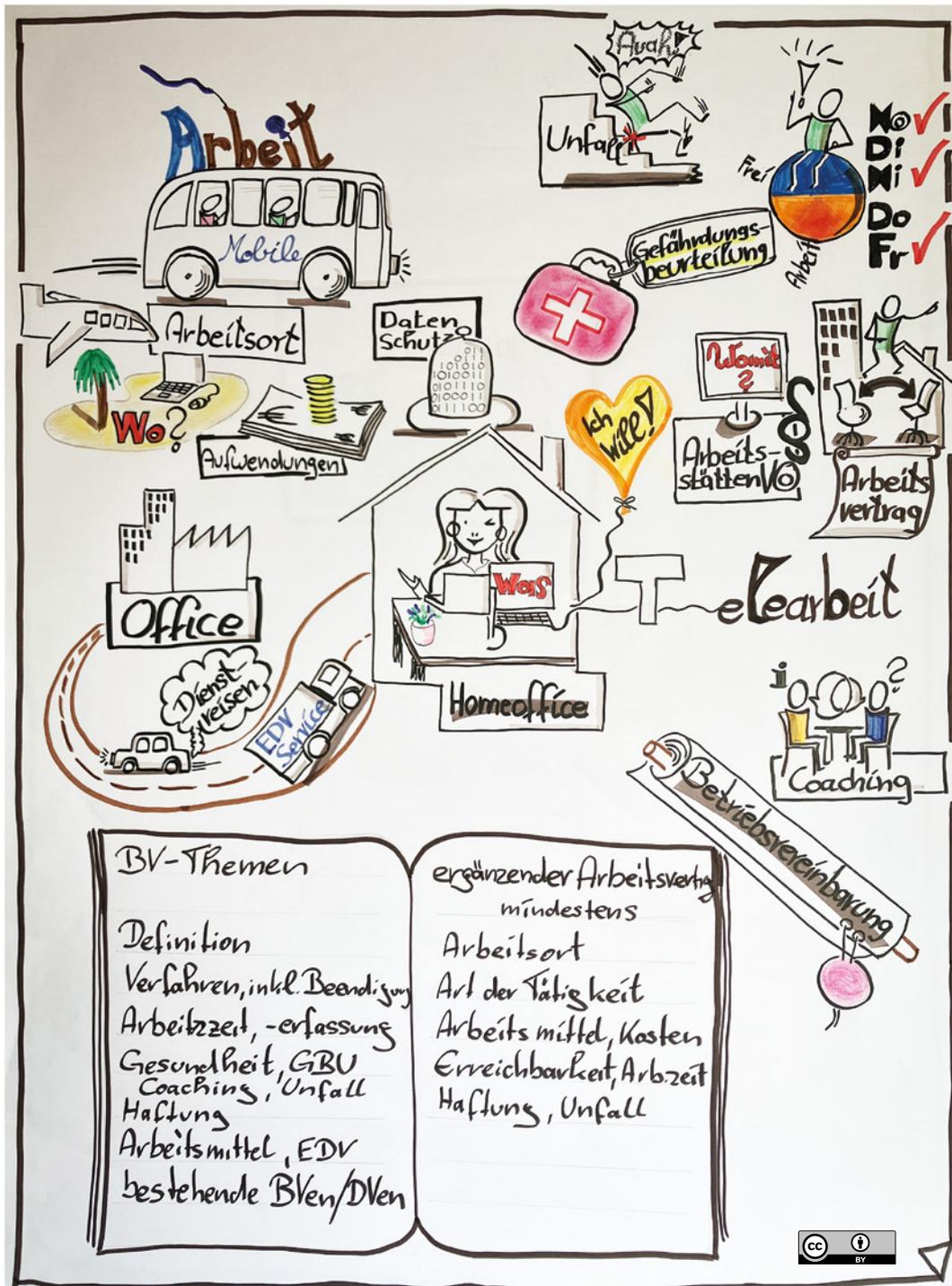
### Kontakt zur Redaktion der FIfF-Kommunikation:

[redaktion@fiff.de](mailto:redaktion@fiff.de)

# Schluss F..I..f..F..

Katharina Just

## Mobile Arbeit gestalten



Geeignete Texte für den SchlussFiff bitte mit Quellenangabe an [redaktion@fiff.de](mailto:redaktion@fiff.de) senden.