

WENN KI,  
DANN  
FEMINISTISCH  
IMPULSE AUS  
WISSENSCHAFT  
UND  
AKTIVISMUS

Hrsg. von netzforma\* e.V.  
Berlin 2020





Wenn KI, dann feministisch  
—  
Impulse aus Wissenschaft und Aktivismus  
Hrsg. von netzforma\* e.V.  
—  
Berlin 2020







# DIE GESCHICHTE MEINER SCHRIFT

Zu Beginn des Jahres entdeckte meine Mutter eine neue Leidenschaft: Sticken. Sie hatte schon vorher Freude an kreativen Praktiken gehabt (wie Weben und Stricken). In ihrer neuen Aktivität fand sie etwas befriedigendes, was sie sehr erfüllt. Ihre ersten Stickereien waren Buchstaben, vor allem aus dem Sajou-Alphabet. Als ich die sah, kam mir eine Idee. Gemeinsam dachten wir über eine fruchtbare Zusammenarbeit nach, die unsere beiden jeweiligen Hobbys vereinen könnte. Was wäre, wenn sie meine eigenen Schriften stickte? Um diese Idee umzusetzen, entwickelte ich für meine Schrift Arthemys Display Light eine neue Version, die es meiner Mutter ermöglichte, die Formen auf ein spezielles Gewebe (namens Aïda) zu übertragen. Dieser Prozess erfordert eine binäre Sprache, genau wie Pixel auf einem Computer. Die Buchstaben werden also Punkt für Punkt gedacht und akribisch zusammengesetzt. Das geht weg von dem Gedanken einer vektorisierten Form, die nur durch ihre Umrisse definiert wird. Vielmehr muss die Form in ihrer gesamten Fläche aufgebaut werden, wie mit Legosteinen! Auf dieser Grundlage konnte meine Mutter die verschiedenen Buchstaben bauen, indem sie einfach dem Raster des Gewebes folgte. Sicherlich war dieser Prozess ziemlich aufwendig und verlangte eine Menge Geduld ab, aber meine Mutter fand die Idee, mithilfe eines mathematischen Systems ein zierliches Kunstwerk zu schaffen, sehr befriedigend.

Morgane Vantorre  
12.10.2020  
Paris

# EINLEITUNG

Wenn KI, dann feministisch

netzforma\* – Verein für feministische Netzpolitik

1–318

7

# ÜBERWACHUNG

1 — Überwachung und Künstliche Intelligenz

Wer überwacht hier eigentlich wen?

Francesca Schmidt & Johanna Luise Mellentin — 319–878

15

# DIGITALE GEWALT

2 — Kontrollverlust und (digitale) Entmündigung

– Das Gewaltpotential Künstlicher Intelligenz

Leena Simon — 879–1335

31

# ALGORITHMISCHE ENTSCHEIDUNGS- SYSTEME

3 — Effiziente Ungleichheit

Victoria Guijarro-Santos

1336–1805

47

4 — Automatisierte Ungleichheit

Wie algorithmische Entscheidungssysteme

gesellschaftliche Machtverhältnisse (re)produzieren

Phillip Lücking — 1806–2240

65

5 — KI und Beschäftigung: Das Ende menschlicher Voreingenommenheit  
oder der Anfang von Diskriminierung 2.0?

Deniz Erden — 2241–2866

77

# ROBOTIK

6 — Feminismus und automatisierte Waffensysteme

Ray Acheson

2867–3381

91

7 — Harmony's Future | No Future w/o Harmony

Constanze Erhard

3382–3788

107

# FILM

- 8 — Droid in Distress oder Fembot Fatale?  
Ein Kulturhistorischer Überblick der Darstellung  
weiblich konstruierter Künstlicher Intelligenzen  
in Science-Fiction-Filmen  
Josephine D'Ippolito — 3789–4282 119

# KUNST

- 9 — PosterdesAln  
Katharina Nejdli — 4283–4335 133
- 10 — Re-präsentation verweigern  
Dr. Tiara Roxanne — 4336–4703 147
- 11 — AI: Queer Art  
Dr. Katrin Köppert — 4704–4970 159

# FEMINIST FUTURES

- 12 — K.rüppel Intelligenz  
Katharina Klappheck — 4971–5403 169
- 13 — Feministische Zukünfte schreiben.  
Katrin Fritsch & Helene von Schwichow — 5404–5479 181
- 14 — Use your technical Monsters – create them, nurture  
them and use them for creating nurture and care  
Annika Kühn — 5480–5684 185
- 15 — KI can't care.  
Mütterlichkeit im Zeitalter Künstlicher Intelligenz  
Hannah Lichtenthäler — 5685–6091 193

# FAZIT

- Ein feministischer Ausblick  
netzforma\* – Verein für feministische Netzpolitik — 6092–6233 205





EINLEITUNG

WENN KI,  
DANN  
FEMINISTISCH



NETZFORMA\*



VEREIN FÜR  
FEMINISTISCHE  
NETZPOLITIK



1 Technik ist, war und wird nie neutral sein. Ebenso wenig wie  
2 die Menschen, die sie erschaffen. Sie ist damit auch nicht  
3 objektiv, sondern agiert in hohem Maße normativ. Diese Tat-  
4 sachen rücken gerade in Anbetracht der Ausbreitung neuarti-  
5 ger digitaler Technologien zunehmend in den Vordergrund,  
6 bilden das Momentum einer kritischen Analyse sozio-techni-  
7 scher Verhältnisse. Betrachtet man die Normativität von Tech-  
8 nik, so eröffnet sich ein Blick auf ihre Verstrickung mit bereits  
9 bekannten Macht- und Herrschaftsstrukturen. Neue Techno-  
10 logien, allen voran Künstliche Intelligenz (KI), haben im gesell-  
11 schaftlichen, wirtschaftlichen, wissenschaftlichen wie politi-  
12 schen Geschehen Handlungs- und Wirkungsmacht. Sie müssen  
13 daher machtkritisch analysiert werden, denn sie agieren nie  
14 selbstbestimmt, sondern folgen stets menschlich gesetzten  
15 normativen Vorstellungen.

16 Was schon für rudimentäre Technik galt, gilt umso  
17 mehr für komplexe Technologien. Die Auseinandersetzung  
18 mit KI reiht sich hier mit ihrer Geschichte ein. In der Phase ihrer  
19 Entstehung – diese wird gemeinhin in die 1950er Jahre in den  
20 USA datiert, als 10 Wissenschaftler (!) einen Sommer lang in  
21 Dartmouth zusammenkamen, um die Möglichkeiten Künstlicher  
22 Intelligenz auszuloten – stand eine philosophische Her-  
23 angehensweise im Mittelpunkt. Sie rückte das Verhältnis  
24 Mensch-Maschine in den Fokus: es galt, Gemeinsamkeiten,  
25 aber auch grundlegende Unterschiede zu verstehen, die die  
26 Natur des Menschen mit seinen genuinen und unverwechsel-  
27 baren Eigenschaften der Leistungsfähigkeit der Maschine  
28 gegenüber stellten.

29 Heute sehen wir große technologische Fortschritte  
30 im Bereich des maschinellen Lernens (eine Form schwacher  
31 KI) – allen voran in deep learning Verfahren, also besonders  
32 komplexen algorithmischen Systemen. Zunehmend rücken  
33 technisch-materielle Aspekte in den Fokus, eng gebunden an  
34 gesellschaftliche Strukturen. Mit Blick auf bewusst zu tref-  
35 fende Entscheidungen erhielten in diesem Kontext insbeson-  
36 dere Themen des Datenschutzes bzw. der Schutz der indi-  
37 viduellen Privatsphäre im Zuge von Big Data, wie auch Trans-  
38parenzstandards – also der Einblick, das Verständnis oder  
39 auch die Verbesserung der technologischen Verarbeitung von  
40 Daten, prominenten Status.

41 Neue Technologien bringen immer auch neue Fragen  
42 mit sich, schauen wir nur beispielhaft in den Bereich der bewaff-  
43 neten Drohnen und die einhergehende Diskussion um Verant-  
44 wortung und Haftung. Vielfach handelt es sich bei diesen Dis-  
45 kussionen um Themen im Bereich von Moral, Ethik und Recht,  
46 die schon länger, wenn auch auf anderer Ebene diskutiert,  
47 aber nie final beantwortet, entschieden oder gelöst wurden.  
48 Oft auch deshalb, weil gesellschaftlichen Dynamiken die Auf-  
49 merksamkeit wieder weg von diesen Themen lenkten oder  
50 aufgrund einer vermeintlichen (philosophischen und gesell-  
51 schaftlichen) Unlösbarkeit der Probleme. Der Einsatz von KI  
52 zwingt uns nun diese Themen ganz oben auf die Liste zu  
53 schreiben. Das gilt für Drohnen, das gilt aber auch für auto-  
54 nom fahrende Autos und die (Haftungs-)Entscheidung darü-  
55 ber, wie diese im Falle eines unausweichlichen Unfalls reagie-  
56 ren soll, kann und rechtlich muss, wenn es um die Abwägung  
57 von Menschenleben geht.

58 Letztlich geht es um die grundlegenden gesellschaft-  
59 lichen Fragen von Gleichstellung, Gleichbehandlung, Zugang  
60 und Ausschluss. Rassistisch agierende Bots und Gesichtser-  
61 kennungssoftware, die noch immer bestimmte Gruppen dis-  
62 kriminiert, machen zunehmend deutlich, wie und wo beste-  
63 hende Diskriminierungsverhältnisse in digitale Systeme getragen

und dort verstärkt werden und nicht selten zu justiziablen Grundrechtsverletzungen führen. Mit Nachdruck müssen wir deshalb tradierte gesellschaftliche Muster aufdecken und aufbrechen und einen transparenten Zugang auf die Entwicklung algorithmischer Systeme einfordern. Technik ist am Ende schließlich immer nur so gut oder so schlecht, wie die Gesellschaft, die sie hervorbringt. Auch die Komplexität gegenwärtiger Diskriminierungsphänomene muss deshalb verstärkt in den Fokus rücken. Eine intersektionale Perspektive also auf vielschichtige, stets verschränkte und sich oft wechselseitig verstärkende Diskriminierungskategorien wie Geschlecht, sexuelle Identität, Hautfarbe und Herkunft, Religion oder Behinderung.

Im November 2019 veranstalteten wir unterstützt von der *Landeszentrale für politische Bildung Berlin* und in Zusammenarbeit mit dem *Institut für Geschichte und Theorie der Gestaltung der Universität der Künste Berlin* (UdK) und dem *Gunda-Werner-Institut für Feminismus und Geschlechterdemokratie* ein Barcamp sowie eine Kunstausstellung, kuratiert von Dr. Katrin Köppert, unter dem Titel „Wenn KI, dann feministisch“. Verschiedene Expert\*innen aus Politik und Zivilgesellschaft sowie Künstler\*innen begleiteten die Veranstaltung und brachten ihre Perspektiven über die Zukunft von KI ein.

In der vorliegenden Publikation haben wir unserer Barcamp-Teilnehmer\*innen und Expert\*innen erneut eingeladen, dieses Thema weiterzuentwickeln. Wir geben hiermit intersektionalen feministischen Perspektiven einen Raum in diesem sich rasant entwickelnden gesellschaftlichen Diskurs rund um Entwicklung und Nutzung intelligenter Technologien. Uns geht es auch um eine Bestandsaufnahme: Wer redet heute, wenn über Künstliche Intelligenz und algorithmische (Entscheidungs-)Systeme gesprochen wird? Welchen gesellschaftlichen Positionen gehört gegenwärtig die Bühne und die Aufmerksamkeit – und was bedeutet das für die künftige Entwicklung von intelligenten Systemen? Wie steht es um die Rolle von intersektionalen feministischen Positionen mit Blick auf die Entwicklung von KI? Welche gesellschaftlichen Positionen werden in den Forderungen nach einer *gerechten* KI berücksichtigt? Und wird dies überhaupt gehört?

Als Verein haben wir kollaborativ an dieser Publikation gearbeitet, Expert\*innen eingebunden und auch unsere eigenen Perspektiven weiterentwickelt. In den Überblickstexten, den Beiträgen unserer Mitglieder sowie in den übersetzten Texten haben wir uns bewusst für die folgenden Nomenklatur entschieden: Das Gendersternchen \* verweist auf die Konstruiertheit der gegenderten Kategorie (wie z. B. Frau\*) und ermöglicht es, Identitäten und Selbstpositionierungen jenseits der historisch gängigen und auch aktuell häufig zugeschriebenen Bezeichnungen mitzudenken. Die Großschreibung von Schwarz verweist auf die Strategie der Selbstermächtigung. Es zeigt das symbolische Kapital des Widerstandes gegen Rassismus an, welches rassistisch markierte Menschen und Kollektive sich gemeinsam erkämpft haben. BIPOC steht für die politische Selbstbezeichnung Black, Indigenous und People of Color, die eine geteilte Rassismuserfahrung aufgreift und diese in eine kollektive Positionierung einfließen lässt. Community verweist in der Beibehaltung der englischen Terminologie auf das Kollektive und das darin eingeschriebene Widerstandspotential. Dies in Abgrenzung zur herkömmlichen Bedeutung der deutschen Übersetzung „Gemeinde“ oder „Gemeinschaft“. Die Klein- und Kursivsetzung von *weiß* verweist auf die soziale Konstruiertheit von Diffe-

128 renzmarkierungen, wobei *weiß* üblicherweise unmarkiert bleibt.  
129 Da im Gegensatz zu Schwarz dieser Differenzmarkierung  
130 kein Selbstermächtigungs- und Widerstandspotential inne-  
131 wohnt, wird *weiß* auch nicht groß geschrieben.

# ÜBERSICHT ÜBER DIE BEITRÄGE

132 Diskriminierungsmuster werden durch den Einsatz von KI und  
133 ADM-Systemen zementiert, teils verstärkt und strukturell tradiert. Intersektionale Ausschlusssysteme, fußend auf den  
134 etablierten gesellschaftlichen Ausschlusskategorien von Gender und race, Klasse, Religion, Behinderung und Schichtung  
135 fließen auch in diesen Kontext unreflektiert und unkritisch in innovative Systeme ein. Dies ist die Grundlage, von der aus wir  
136 in dieser Publikation eine kritische Bestandsaufnahme wagen.

137 In verschiedenen Themenblöcken kommen dabei  
138 unterschiedliche Perspektiven und wissenschaftliche Disziplinen zu Wort – mit dem Ziel, einen möglichst breiten Zugriff  
139 auf das Thema zu erlauben. Technik und Gesellschaft beeinflussen sich wechselseitig auf viele Arten; besonders hervor-  
140 zuheben sind aber sicher die Kategorien, die wir als Schwerpunkt gewählt haben: Überwachung, Gewalt, Robotik und  
141 Bewertungsverfahren.

142 Francesca Schmidt und Johanna Luise Mellentin  
143 arbeiten in ihrem Beitrag *Überwachung und Künstliche Intelligenz: wer überwacht hier eigentlich wen?* heraus, wie der  
144 Einsatz von KI bestehende Formen der Überwachung auf die breite Gesellschaft ausweitet und so auch neue Formen von  
145 Diskriminierungen und Ausschlüssen verschärft. Außerdem stellen sie die Fragen, *wer* denn eigentlich nun wirklich über-  
146 wacht werden muss, Mensch oder KI?

147 Eng verknüpft mit dem Thema Überwachung ist das  
148 der digitalen (häuslichen) Gewalt. Neben dem bisher dominierenden Thema Hate Speech gehen mit KI auch neue, sich ver-  
149 ändernde Formen der vergeschlechtlichten Gewalt einher. Leena Simon betont in ihrem Beitrag *Kontrollverlust und (digi-  
150 tale) Entmündigung – Das Gewaltpotential Künstlicher Intelligenz*, dass mit steigenden Fähigkeiten der KI auch ihre Kom-  
151 plexität und Intransparenz zunimmt. In die formale Welt der mathematischen Berechnungen mischt sich, so Simon, eine  
152 gefährliche Melange aus hegemonialen Weltbildern, reproduzierten Stereotypen und Missverhältnissen unserer Vergan-  
153 genheit – und das vor allem leise und unbemerkt.

154 Unter der Überschrift *Algorithmische Entscheidungssysteme* widmen Deniz Erden, Phillip Lücking und Victoria  
155 Guijarro-Santos sich in ihren Beiträgen automatisierten Entscheidungsverfahren. Im ersten Beitrag stellt Victoria Guijar-



ro-Santos vordergründig die Allgegenwärtigkeit der auf Algorithmen basierenden Entscheidungssysteme heraus. Eingesetzt und legitimiert werden diese vor allem, weil ihnen eine besondere Effizienz zugeschrieben wird – doch rechtfertigt dieses Argument letztlich die durch sie entstehenden Ungleichheiten? Guijarro-Santos entgegnet dieser Sichtweise und macht deutlich, dass wir zur Bekämpfung der Manifestation von Ungleichheit bereits über ein geeignetes und ausbaufähiges Instrument verfügen: Unsere Grundrechte.

Phillip Lücking bildet in seinem Beitrag die technologisch-materialistische Ebene der KI ab. Bei seiner sozio-technischen Erläuterung grundlegender Begriffe wie KI und Daten, legt er wesentliche normativ-ökonomische Muster offen und demaskiert mit feministischen Denker\*innen die vermeintliche Neutralität technologischer Verfahren. Vor diesem Hintergrund fordert auch er eine politische Aneignung von algorithmischen Systemen, die unsere Realitäten bereits jetzt strukturieren. Darüber hinaus fordert er jedoch vor allem die Forcierung einer feministischen Technikentwicklung.

Deniz Erden beschäftigt sich in ihrem Beitrag *KI und Beschäftigung: Das Ende menschlicher Vorurteile oder der Beginn von Diskriminierung 2.0?* mit der Algorithmisierung von Bewerbungsverfahren in der Arbeitswelt. Vor dem Hintergrund des vermehrten Einsatzes eben dieser Technologien lotet sie deshalb die Potenziale und Grenzen des europäischen Antidiskriminierungs- und Datenschutzrechts sowie der DSGVO aus und vertritt die These, dass mit Blick auf das Vermeiden von (neuen) diskriminierenden Strukturen Aufsichtsorgane eine Agenda für Fairness fördern könnten, indem sie Beratung und Genehmigung von, auf dem Arbeitsmarkt einzusetzenden Tools übernehmen sowie Datenschutz-Folgenabschätzungen vornehmen.

Neben algorithmischen Systemen spielt im Technikzeitalter auch die *Robotik* eine besondere Rolle – etwa in Form von Pflege-, Sex-, und Arbeitsrobotern oder auch in Form von Drohnen als autonome Waffensysteme findet sie Eingang in die unterschiedlichsten gesellschaftlichen Bereiche.

In ihrem Beitrag *Feminismus und automatisierte Waffensysteme* verdeutlicht Ray Acheson, dass zunehmend Waffen auf Basis von KI und Machine-Learning-Verfahren entwickelt werden. Diese vollautonomen Waffensysteme können ohne menschliche Kontrolle entscheiden, wen sie als Ziel auswählen und gegebenenfalls auch wen sie töten. Dabei bleiben solche Systeme nicht unberührt von Stereotypen und Vorurteilen; Acheson zeigt davon ausgehend auf, welche Auswirkungen dies auf die zukünftige Kriegsführung haben wird.

Im zweiten Beitrag zum Thema Robotik wendet sich Constanze Erhard dem Thema der Sexbots zu, welche durch ihre Intimität ein besonders strittiges Feld der Mensch-Maschine-Interaktion darstellen. In ihrem Beitrag *Harmony's Future / No Future w/o Harmony* analysiert sie die damit (oft) verknüpfte (polarisierte) wissenschaftliche Diskussion und bietet eine deutlich differenziertere Analyse von Sexbots an. Erhard schlägt vor, im Bezug auf Sexbots von sexualisierter Care-Arbeit zu sprechen, um damit die (den Sexbots) zugrundeliegenden vergeschlechtlichten Herrschaftsverhältnisse besser fassen zu können und so zum Ausgangspunkt einer feministischen Analyse zu machen.

Aus einer filmwissenschaftlichen Perspektive widmet sich Josephine D'Ippolito dem Thema KI. In ihrem Artikel *Fembot Fatale vs. Droid in Distress: Entwicklung sexistischer Codes in Science Fiction Filmen* untersucht D'Ippolito die kulturhistorischen Entwicklungen der unterschiedlichen Darstel-

235 lungsformen Künstlicher Intelligenzen von 1927 bis 2019 im  
236 Genre der Science Fiction Filme. Dafür greift sie primär Filme  
237 des US-amerikanischen Mainstreams aus verschiedenen Epo-  
238 chen heraus und analysiert, wie die darin vorkommenden Künst-  
239 lichen Intelligenzen Geschlechterstereotype von Weiblichkeit,  
240 u.a. in ihrer intersektionalen Verwobenheit, reproduzieren.

241 Unter dem Kapitel Kunst findet sich ein künstlerischer  
242 Beitrag von Katharina Nejdil mit dem Titel *PosterdesAI*, in  
243 dem die von einem Neuronalen Netzwerk generierten Plakate  
244 anschließend von Grafik Designer\*innen reinterpretiert wur-  
245 den und durch die AR-App Artivive sichtbar sind – dazu gibt  
246 es auch eine Anleitung!

247 Katrin Köppert lädt die Lesenden in ihrem Text *AI:*  
248 *Queer Art* ein, Künstliche Intelligenz als queere Kunst zu ver-  
249 stehen. Sie zeigt, dass die Technologie der KI mit der Techno-  
250 logie des Geschlechts verbunden ist und macht so deutlich,  
251 wie Künstliche Intelligenz mit künstlerischen Konzepten ver-  
252 knüpft ist. Grundlage dafür ist der Test von Alan Turing, der  
253 anhand eines Geschlechter-Rate-Spiels untersucht, ob sich  
254 Computer in ihrem Verhalten von Menschen unterscheiden  
255 lassen. Dabei ergründet Köppert den Zusammenhang zwi-  
256 schen mathematischer Formalisierung und geschlechtlicher  
257 Unberechenbarkeit und zeigt, wie sich die Grenzen von Bere-  
258 chenbarkeit im Gründungstext Künstlicher Intelligenz festschrei-  
259 ben und somit das Potential queerer Kreativität von KI zeigen.

260 Im dritten Beitrag zum Thema Kunst befasst sich Dr.  
261 Tiara Roxanne in ihrem Beitrag *Repräsentation verweigern*  
262 mit KI im kolonialen Siedler-Kontext. Laut Roxanne führen  
263 Künstliche Intelligenz und Machine Learning Systeme Mechanis-  
264 men der kolonialen Siedlergewalt aus, indem sie indigene Bevöl-  
265 kerungen mittels rekonstituierter Bilder darstellen. Roxanne  
266 untersucht dabei Verweigerung als Überlebensmechanismus für  
267 indigene Bevölkerungen. Dabei analysiert sie die Ergebnisse  
268 von Style Transfer und GAN, zwei unterschiedlichen Machine  
269 Learning Modellen, die Ergebnisse als restrukturierte Informa-  
270 tionen liefern, die trotz Verwendung indigener Deskriptoren  
271 koloniale Praktiken fortschreiben. Ihr Artikel dient somit als  
272 eine Art Historisierung, die KI an bereits bekannte koloniale  
273 Praktiken rückkoppelt und zeigt, dass KI nicht neutral sein  
274 kann, da sie koloniale Praktiken fortschreibt.

275 Zu guter Letzt geben wir denen Raum, die nach vorne  
276 schauen, in eine Zukunft, in der intersektionale feministische  
277 Positionen zum *common sense* avanciert sind.

278 Unter der Überschrift *Feminist Futures* führen Katrin  
279 Fritsch und Helene von Schwichow in ihrem Beitrag *Femi-*  
280 *nistische Zukünfte schreiben* in den Themenbereich der  
281 Zukünfte und Utopien ein. Sie betonen, dass algorithmische  
282 Systeme uns vor allem eines gelehrt haben: Dass die Zukunft  
283 nicht ohne Gegenwart denkbar ist und auch, dass die Gestal-  
284 tung der Zukunft gegenwärtiger Gesellschaften vor allem von  
285 weißen Männern dominiert wird. Ausgehend davon muss die  
286 Frage lauten: Wie können wir Zukunft von Grund auf feminist-  
287 tisch denken und Werte wie Gemeinschaft oder soziale  
288 Gerechtigkeit grundlegend inkludieren?

289 Katharina Klappheck unternimmt den Versuch eines  
290 Entwurfs einer alternativen Vision für KI und orientiert sich  
291 dabei an den *disability studies*. Bisher wurde die Kategorie der  
292 Behinderung im Lernprozess Künstlicher Intelligenzen oft als  
293 Abgrenzungsfolie verwendet, um über die Abgrenzung des  
294 Unregelmäßigen zur Identifikation des Regelmäßigen, des „Nor-  
295 malen“ als Zielvorgabe zu gelangen. Dabei definiert Klapp-  
296 heck die Beziehung zwischen beiden als eine komplexe Kon-  
297 stellation, die selbst wiederum Hinweise darauf geben kann,

298 wie eine alternative KI aussehen kann und fragt: „Was wäre,  
299 wenn KI selbst ein Krüppel wäre?“

300 Annika Kühn beleuchtet in ihrem Beitrag, dass Erschaf-  
301 fer\*innen von Technologien zu selten ihren Einfluss auf die  
302 entsprechende Technologie und den Einfluss dieser Techno-  
303 logie auf sie selbst reflektieren. Ausgehend von einer „emo-  
304 tional carenden KI“ soll in dieser feministischen Utopie ein  
305 *empowernder* Raum für Macher\*innen und User\*innen von KI  
306 aufgezeigt werden.

307 Abschließend geht es im Text *KI can't care. Mütter-*  
308 *lichkeit im Zeitalter Künstlicher Intelligenz* von Hannah Licht-  
309 enthäler um Mutterschaft und Fürsorge im Zusammenhang  
310 mit KI. Lichtenthäler leitet mit bekannten mütterlichen Vor-  
311 bildern aus populärer Kultur ein und schaut kritisch auf die  
312 Konzepte der Mütterlichkeit und Mutterschaft. Sie wirft die  
313 Frage auf, weshalb Mutterschaft innerhalb der feministischen  
314 Diskurse eher ein Randthema bleibt und stellt die Frage in den  
315 Raum, wie eine Zukunft mit KI für Mutterschaft aussehen  
316 kann. In der Reproduktionsmedizin wird schließlich bereits auf  
317 KI gesetzt, um zukünftig erfolgreiche Schwangerschaften zu  
318 steigern.





1

ÜBERWACHUNG  
UND KÜNSTLICHE  
INTELLIGENZ

:

WER ÜBERWACHT  
HIER  
EIGENTLICH  
WEN?

✦

FRANCESCA  
SCHMIDT  
&  
JOHANNA LUISE  
MELENTIN

Künstliche Intelligenz (KI) und die automatisierte Auswertung von Daten ermöglichen die Ausweitung von Überwachung über ein bisher unbekanntes Maß weit hinaus. Zunehmende und beschleunigte Verknüpfung von Datensätzen bis tief in die Metadatenebene bringt eine neue Qualität des Überwachungsbildes mit sich – von Individuen, (marginalisierten) Bevölkerungsgruppen oder auch ganzen Gesellschaften. Zudem wird der Eingriff in das Grundrecht auf informationelle Selbstbestimmung intensiviert sowie in eine Vielzahl von weiteren Grundrechten bis hin zu Art. 1 GG: Der Menschenwürde. Nicht nur staatliche Überwachungsmaßnahmen steigen und werden in diesem Sinne zunehmend technisiert, auch die gegenseitige Überwachung (lateral surveillance) nimmt zu. Zeiten der Krise, wie wir sie derzeit mit der Corona-Pandemie erleben, können, aufgrund neu erwachsener Unsicherheiten und Ängste, diese Spirale verstärken. Die Vernetzung unterschiedlicher Technologien und Datensätze führt auch dazu, dass neue Technologien verstärkt nicht mehr nur zielgerichtet sondern kategorisch eingesetzt werden. Datenbestände wachsen ins Unermessliche, Künstliche Intelligenz kann helfen, zu durchforsten, clustern, zusammenzufügen und ganze Gesellschaftsgraphen zu erstellen (vgl. boyd & Crawford 2012). Heute werden also ganze Kontexte überwacht. Im Verhältnis Staat-Bürger\*innen verlagert diese Entwicklung vor allem staatliche Maßnahmen immer weiter ins Vorfeld. Mit Verweis auf die Verhinderung von Straftaten werden beispielsweise unzählige Videokameras auf öffentlichen Berliner Plätzen gefordert. Wenn diese Technologien mit Künstlicher Intelligenz gekoppelt werden, entsteht so ein Panoptikum, das in seinen Ausmaßen weit über das von Foucault erdachte hinausgeht.<sup>1</sup> Anlasslose Massenüberwachung wird im 21. Jahrhundert durch den Einsatz und die Entwicklung von KI mehr als jemals zuvor zur Realität. Eine Realität, die sich jeden Tag ein bisschen mehr, ein wenig intensiver und immer weiter ausdehnt. Für manche mehr, für manche weniger spürbar. Sichtbar? Für wen? Was? Zu welchem Zeitpunkt? Das bleibt intransparent bis unsichtbar.

Unsere grundlegenden demokratischen Freiheiten und Werte stehen mehr denn je auf dem Spiel. Nicht nur durch die aktuelle Corona-Pandemie wird der populistische Ruf nach Sicherheit, Ordnung und Kontrolle immer stärker. Auch diese verstärkten Rufe, die sich einer lähmenden Angstspirale bedienen sind es letztlich, die die Welt von Tag zu Tag unsicherer erscheinen lassen – mit verheerenden Folgen, denn dies steigert wiederum die Akzeptanz von sinnlosen und ineffektiven Maßnahmen der Überwachung. Mit ihrer Hilfe wird die zunehmende Ausstattung staatlicher Behörden mit rechtlichen und technischen Kompetenzen zur Überwachung gerechtfertigt. Vor allem die grundrechtlich gesicherte Privatsphäre wird dabei immer weiter eingeschränkt.

1 Das panoptische Gefängnis ist kreisrund, mit einem Wachturm in der Mitte, von dem aus die Wärter\*innen alle Gefangenen beobachten können, ohne, dass die Gefangenen die Wärter\*innen sehen können. Diese Anordnung unterteilt beide Gruppen in Sehende und zu Sehende und ist, anders als die öffentliche Videoüberwachung in ihrem Ausmaß begrenzt (vgl. Foucault 1977).

# FEMINISTISCHE PERSPEKTIVEN AUF ÜBERWACHUNG

Diese Einschränkungen treffen nicht alle Menschen gleich. Es lohnt sich ein feministischer, vor allem intersektionaler Blick auf den Komplex der Überwachung, auch über die Frage der Privatsphäre hinaus.

Schauen wir zurück, sehen wir: Die Überwachung von all jenen, die nicht der patriarchal geprägten männlichen heteronormativen Norm entsprechen, hat eine lange Tradition. Simone Browne, Professorin für African and African Diaspora Studies an der Universität Austin, Texas machte deutlich, dass Überwachung und deren Technologien rassifiziert sind und hauptsächlich dazu dienen, die *weiße*<sup>2</sup> Norm wiederherzustellen und sicherzustellen, wer dazugehört und wer nicht. Auch wird deutlich, dass viele Praktiken der Überwachung, die im transatlantischen Sklav\*innenhandel eingesetzt wurden, bis heute fortwirken, sowohl diskursiv, als auch real eingesetzt werden (Browne 2015). Zum Beispiel werden bis heute Praktiken der körperlichen Vermessung, vor allem zur Altersbestimmung von jungen Geflüchteten angewandt. Auch die biometrische Erfassung etwa von Fingerabdrücken folgt diesem Muster.

Auch der Frauenkörper ist, historisch betrachtet, geprägt von eben jener patriarchalen Überwachung. Zu denken wäre hier, vor allem im deutschen Kontext, nicht nur an die Einschränkung bürgerlicher Freiheiten, wie zum Beispiel die Aufnahme von Erwerbsarbeit, die für Frauen bis 1977 nur mit Zustimmung des Ehemanns möglich war, oder das Eröffnen eines eigenen Bankkontos, welches erst seit 1958 möglich ist (Bundeszentrale für politische Bildung 2012). Auch die derzeit gültige Regelung zu Schwangerschaftsabbrüchen ist eine Form der Überwachung des weiblichen Körpers bzw. in Teilen des Trans\*Körpers.

Es wird deutlich: Überwachung hat nicht für alle Menschen die gleichen regulierenden Eingriffe zur Folge. Überwachung trifft somit nicht alle gleich. Die Freiheit der Einen, innerhalb der *weißen* patriarchalen Norm, bedeutet die Überwachung der Anderen, also von all jenen, denen Zugehörigkeit verweigert wird, oder die vielleicht gar nicht Teil dieser Norm sein wollen.

2 In diesem Text wird *weiß* oder *weiß-sein* immer klein und kursiv geschrieben, da es im Gegensatz zu Schwarz kein Begriff für eine politisch empowernde Selbstdarstellung ist. Zudem verweist diese Schreibweise von *weiß* oder *weiß-sein* auf die üblicherweise unmarkierten Differenzmarkierungen sozialer Konstruktionen (vgl. Eggers et al. 2005).

# ABER WIR WERDEN DOCH ALLE ÜBERWACHT

407 Dass der beliebte Satz „Ich habe nichts zu verbergen“ – des-  
408 halb auch kein Problem mit massenhafter Überwachung – zu  
409 kurz greift, zeigen schon die Ausführungen weiter oben. Denn  
410 die Einschränkungen erleben eben in erster Linie deprivile-  
411 gierte und marginalisierte Menschen. Aber ganz grundsätz-  
412 lich sollten wir uns als Gesellschaft gegen eine durch KI zuneh-  
413 mend automatisierte Form der Überwachung wehren. Denn  
414 Privatsphäre zu wahren, heißt nicht unbedingt, Geheimnisse  
415 zu haben. Vielmehr lässt so ein Satz darauf schließen, der nur  
416 im Verhältnis zu einem Gegenüber stehen kann, dass bei die-  
417 sem die Überwachung wohl begründet zu sein scheint. Die  
418 *Anderen* (jenseits der Norm) haben also etwas zu verbergen.

## PRIVATSPHÄRE: FÜR WEN?

419 Überwachung wird sehr oft unter dem Aspekt der Privatsphäre  
420 diskutiert. Doch wir sollten hier zumindest in einem ersten  
421 Schritt unterscheiden, wer in die Privatsphäre eindringt: Der  
422 Staat, private Unternehmen oder andere Menschen (oft mit  
423 Hilfe von privaten Unternehmen, wie Facebook & Co).

424 Wenn wir von staatlichen Eingriffen in die Privat-  
425 sphäre sprechen, sehen wir, dass die Intensität dieser Ein-  
426 griffe mit dem Grad der Abhängigkeit vom Staat zunimmt. So  
427 müssen etwa Hartz-4-Empfänger\*innen, Menschen mit Behin-  
428 derung, Geflüchtete bzw. Asylsuchende mit wesentlich zahl-  
429 reicheren und tiefergehenden Eingriffen rechnen, als Men-  
430 schen die zu keiner dieser Gruppen oder anderen marginalisierten,  
431 stigmatisierten Gruppen gehören.

432 Menschen, die mit Hilfe von Bildung und technischem  
433 Wissen ihre Daten vor den privaten Unternehmen in größerem  
434 Umfang schützen können, oder, sofern das überhaupt mög-  
435 lich ist, selbstbestimmt ihre Daten zur Verfügung stellen kön-  
436 nen, können ihre Privatsphäre in größerem Umfang schützen.  
437 Es ist ein faktisches Privileg, denn rechtlich steht es allen glei-  
438 chermaßen zu.

439 Wenn wir also Überwachung unter dem Diktum der  
440 Privatsphäre diskutieren, wird deutlich, dass wir gesellschaft-  
441 liche Macht- und Herrschaftsstrukturen mitdenken müssen,  
442 denn die Freiheit, die einige wenige durch den Schutz der Pri-  
443 vatsphäre erhalten, ist zugleich die Unfreiheit der *Anderen*.



# BIG DATA, ALGORITHMEN UND KI SORTIEREN UNSERE GESELLSCHAFT „NEU“ UND SCHOTTEN SIE AB

Der große Datenhaufen, den wir jeden Tag aufs neue anhäufen, ist die Grundlage für zunehmende, auch staatliche Überwachung. Zumindest ermöglicht es dieser, auch an „Qualität“ stetig verbesserte Datenhaufen (neue) Korrelationen, Beziehungen von Sachverhalten herzustellen. Genutzt wird dies zum Beispiel beim *predictive policing*.

Die an unterschiedlichen Stellen und von unterschiedlichen Systemen gesammelten Daten werden durch automatisierte Prozesse in neue Zusammenhänge gebracht. Vereinfacht gesagt, sind es Algorithmen, die aufgrund einer Analyse von Falldaten Wahrscheinlichkeiten berechnen und so Aussagen treffen, wer wann und wo ein mögliches Verbrechen begehen könnte. Mit dieser Technologie wird der Einsatz von Polizeikräften gesteuert, mittlerweile auch von einigen Polizeibehörden in deutschen Bundesländern. In den USA jedoch ist es schon Teil des Alltags, dass potentielle Täter\*innen Besuch von der Polizei bekommen, mit dem Hinweis und der Warnung vor möglichen zukünftigen Straffälligkeiten (Gorner 2013; Merz 2016).

Doch was, wenn die davon Betroffenen nur durch neue, durch KI hergestellte Datenkorrelationen zu potentiellen Täter\*innen gemacht werden? Sie werden einem Risikocluster aufgrund persönlicher Gewohnheiten zugeordnet. Dagegen können sich die Betroffenen schwer wehren. Denn sie müssen sich gegen einen Verdacht zur Wehr setzen, der nur auf Grund einer Datenkorrelation bzw. dem, was KI daraus gemacht haben, entstanden ist (Gless 2016). Erschwerend kommt hinzu, dass Technologie Objektivität und Neutralität zugeschrieben wird, ihr deshalb häufig eine Fehlerlosigkeit attestiert wird. Doch aktuell werden diese Programme in der Regel weder nach transparenten Kriterien entwickelt noch einheitlich oder überhaupt überprüft. Wie die Wahrscheinlichkeiten genau berechnet werden, weiß die Polizei daher in den seltensten Fällen. Doch wie passt die daraus resultierende Notwendigkeit des Vertrauens in unser Wertesystem?

Wir werden also nicht nur in potentielle Täter\*innen und Opfer sortiert, auch andere Cluster werden mit Hilfe der Daten gebildet. Die dann wiederum über unsere Kreditwürdigkeit, die Höhe von Versicherungsbeiträgen, Jobangebote oder im Zweifel die Kosten unserer Krankenversicherung (was in Deutschland zum Glück noch nicht der Fall ist) entscheidet. Die Fachliteratur verwendet für dieses Clustering den Begriff *social sorting* (Lyon 2003), also die Einsortierung in bestimmte soziale Klassen und damit deren Verortung im vorherrschenden Machtgefüge.

Dieses Machtgefüge (z. B. Europa) schottet sich auch mit Hilfe von Künstlicher Intelligenz nach außen ab. Grenzkontrollen werden automatisiert, der elektronische Reisepass soll Zeit ersparen, stellt aber all jene vor eine Herausforderung bzw. unter ein besonderes Augenmerk, die ihn nicht haben, im Zweifel auch nie bekommen können und macht all diejenigen, die ihn besitzen zu noch durchsichtigeren Menschen. Und was passiert mit den bestehenden staatlichen Überwachungssystemen und gehorteten Datenmengen darüber hinaus, wenn sich einmal das derzeitige politische System ändern sollte? Ein Blick in Länder mit anderer Regierungsstruktur zeigt es.

Tödlich wird die Abschottung beim sogenannten Geofencing, wenn Künstliche Intelligenz herangezogen wird, um „digitale Zäune“ zu bewaffnen. Mittels Videoüberwachung, Bewegungssensoren oder Wärmebild werden Grenzgänger\*innen aufgespürt – an der türkisch-syrischen Grenze beispielsweise und dann den integrierten automatisierten Selbstschussanlagen überlassen. Doch auch, wenn wir uns als Gesellschaft gegen den Einsatz von tödlicher KI aussprechen, scheinen die moralischen Grenzen durchlässig: Eine Vermengung von Migrationspolitik, Wirtschaftspolitik und der Rüstungsindustrie führt dazu, dass „digitale Zäune“ auch aus Deutschland und Europa stammen, wie beispielsweise jene in Marokko. Oder der, der 2009 unter der Mitfinanzierung u. a. durch die deutschen Gesellschaft für Internationale Zusammenarbeit von EADS (heute Airbus), nach Saudi-Arabien geliefert wurde (Grieger & Schlindwein 2016).

# ÜBERWACHUNG DURCH UNTERNEHMEN

Unternehmen, v. a. soziale Netzwerke, überwachen ihre Nutzer\*innen und sammeln, speichern und verarbeiten massenhaft ihrer Daten. Auf ihrer Grundlage bieten sie beispielsweise maßgeschneiderte Werbung an, steuern Aufmerksamkeit und beeinflussen Kaufentscheidungen.

Die Nutzer\*innen selbst sind es, die sich mit Hilfe von Apps und anderen Technologien überwachen. Mit dem Ziel der Selbstoptimierung geben wir Daten über unsere sportlichen Leistungen, Menstruationszyklen, Arbeitsfortschritte, Schlafqualität oder Fokuszeiten an Unternehmen. Meist zu einem

sehr niedrigen Preis, nämlich um Apps, Technologien oder andere angebotene Services kostenlos nutzen zu können.

Die Daten, vornehmlich hoch sensible Gesundheitsdaten, sind auch für andere interessant. Der Staat könnte ebenfalls Regulierungsbestrebungen mit solchen Daten unterfüttern. Beispielsweise könnten Krankenkassen günstige Sporttarife für besonders aktive Menschen anbieten. Die Generali Vitality wirbt mit Sonderbelohnungen und vergünstigten Versicherungsprämien für ein gesundes Lebensverhalten (Dialog Versicherung AG). Ein Punkt für gesundes Essen auf dem Supermarktband, noch einen für den erfolgreichen Check beim Arzt, noch einen für eine Stunde im Fitnesscenter. Doch wer zahlt eigentlich den Preis? All diejenigen, die weniger gesund, weniger fit, weniger freizügig mit ihren Daten sind, müssen dann die teuren Tarife buchen. Auch hier werden die belohnt, die in der Norm bleiben und diese verfestigen.

# TECHNOLOGIE IST FEHLERANFÄLIG, ÜBERWACHUNG DAMIT AUCH

Dass Algorithmen und Künstliche Intelligenz Diskriminierung Vorschub leisten, weil sie in vielen Fällen selber diskriminieren, ist bekannt und in dieser Publikation ebenfalls sehr deutlich geworden. Automatisierte Gesichtserkennung, die nicht nur Menschen ihren „offiziellen“ Identitäten zuordnen soll, sondern in einigen Fällen sogar nach möglichen Affekten und Emotionen wie Wut, Aggression und Gewaltbereitschaft clustern soll, um so etwaigen terroristischen Attentaten zuvorzukommen, gilt als äußerst fehleranfällig. In Bezug auf *Affective Computing* wissen wir aus der Forschung, dass Emotionen und Affekte einerseits kulturell unterschiedlich codiert sind, andererseits werden damit auch Emotionen weiterhin stereotyp und vor allem binär vergeschlechtlicht werden: So wird oft Wut als etwas männliches und Hysterie als weiblich klassifiziert. Bleiben wir bei der Gesichtserkennung. Oftmals gelingt es den Programmen – vor allem bei Schwarzen Menschen<sup>3</sup> – nicht, Gesichter richtig oder gar als Menschen (Barr 2015) zu erkennen. Joy Buolamwini vom MIT Media Lab stellte beispielsweise fest, dass Gesichtserkennung eigentlich nur gut bei *weißen* Männern funktioniert. Schwarze Frauen und Schwarze Männer werden schlechter bis gar nicht erkannt. Bei Schwarzen Frauen ist knapp ein Drittel aller Zuordnungen schlicht falsch (Buolamwini & Gebru 2018). Die American Civil Liberties Union, übersetzt amerikanische Bürgerrechtsunion (ACLU), hat 2018 festgestellt, dass die Gesichtserkennung von Amazon 28 Mitglieder des amerikanischen Kongresses fälschlicherweise Fahndungsfotos zugeordnet hat. Die falschen Übereinstimmungen betrafen unverhältnismäßig viele Men-

570 schen of Color. Dabei waren unter anderem sechs Mitglieder  
571 des Congressional Black Caucus, einer Vereinigung afroame-  
572 rikanischer Mitglieder des Kongresses. Auch der kürzlich ver-  
573 storbene Kongressabgeordnete und Ikone der Schwarzen  
574 amerikanischen Bürgerrechtsbewegung John Lewis war  
575 unter denjenigen, die falsch zugeordnet wurden (Snow 2018).  
576 All die Mitglieder des Kongresses sind Personen öffentlichen  
577 Lebens, also auch mit einer Vielzahl von Bildmaterial in den  
578 unterschiedlichen Datenbanken vertreten, trotz allem hat die  
579 Gesichtserkennung versagt. Eingesetzt als Tool zur (staatli-  
580 chen) Überwachung fördert sie so Diskriminierung, indem  
581 beispielsweise polizeiliche Maßnahmen, sich dann vorwie-  
582 gend gegen Schwarze Menschen richten, da diese auf diese  
583 Weise immer wieder mit Verbrechen in Verbindung gebracht  
584 werden. Wird der Einsatz solch fehlerhafter Systeme durch  
585 staatliche Sicherheitsbehörden betrieben, hat dies aufgrund  
586 des Mächteverhältnisses und der Schutzpflicht des Staates  
587 besonders heftige Folgen – für die Sicherheit, die dadurch  
588 eben gerade nicht garantiert werden kann und für die betrof-  
589 fenen Individuen.

590 Leistungsfähige Algorithmen, wie sie beispielsweise  
591 bei Gesichts- und Spracherkennungstechniken eingesetzt  
592 werden, arbeiten mit lernenden Systemen statt mit einfachen  
593 regelbasierten Konditionalketten. Das bedeutet, die Daten-  
594 basis mit der das System gefüttert wird, hat eine direkte  
595 Auswirkung auf die späteren Entscheidungen des Systems.  
596 Beinhalten die unbereinigten Basisdatenbanken schon ein  
597 Ungleichgewicht oder Diskriminierungen, wird auch der Algo-  
598 rithmus diese fortschreiben und zu Ungleichbehandlung füh-  
599 ren bzw. diese sogar verstärken. Um beim Beispiel der Kon-  
600 gressabgeordneten zu bleiben: wenn in der Datengrundlage  
601 Schwarze Menschen nur bzw. mehrheitlich in Form von Ver-  
602 brecherfotografien abgebildet sind, wird der Algorithmus die  
603 Bilder Schwarzer Menschen automatisch entsprechend zuord-  
604 nen. Diese Diskriminierung muss nicht absichtlich geschehen.  
605 Doch zeigt sie ganz eindeutig fehlendes intersektionales Pro-  
606 blembewusstsein bezüglich Macht- und Hierarchiestrukturen  
607 und sie zeigt auch, wie institutionalisiert Diskriminierungs-  
608 formen wie Rassismus und Sexismus in unser Gesellschaft  
609 sind und fortgeschrieben werden.

3 Der Begriff Schwarz wird in diesem Artikel immer groß geschrieben, um das von People of Color und Schwarzen Menschen eingeschriebene Widerstandspotential zu verdeutlichen (vgl. Eggers et al. 2005).

# VERFASSUNGS- RECHTLICHE GRENZEN STAATLICHER ÜBERWACHUNG — UND WEITER?

Jetzt könnten wir letztlich konstatieren, dass es schlicht nur einer Bereinigung zum Beispiel der Gesichtserkennungstechnologie und ihres Datensatzes braucht, um diese gesellschaftlich und rechtlich akzeptabel zu gestalten. Unter der Ausgangssituation, dass es ebensowenig neutrale Technik wie neutrale Menschen gibt, kämen wir jedoch dann nicht darum herum zu fragen: Wie viel Diskriminierung soll „erlaubt“ sein und wie muss die Beurteilung anders ausfallen, wenn der Staat mit Überwachungsmaßnahmen agiert? Die Antwort gibt das Grundgesetz vor. Doch wenn rechtlich und faktisch gleichzusetzen wäre, gäbe es wohl weder diese Publikation noch bräuchte es Gerichte.

Klar ist, Überwachung mündet häufig in einer Vielzahl von Grundrechtsverletzungen, wie das Recht auf Privatheit, das Recht auf sexuelle und reproduktive Selbstbestimmung, Versammlungs- und Bewegungsfreiheit oder Diskriminierungsfreiheit. Letztendlich sprechen wir auch über die Würde des Menschen, Art. 1 GG. Mit Blick auf den Einsatz von Künstlicher Intelligenz, die die Schwere des Eingriffs um ein Vielfaches erhöhen kann, stellen sich hier ganz neue Fragen aber auch die alten Fragen mit nie dagewesener Dringlichkeit, bleiben aktuell. Mit der gesammelten und bevorrateten Datenmenge kann, so aufbereitet und analysiert, sehr kleinteilig nach verdächtigen Mustern und Korrelationen gesucht werden. Mit der Verknüpfung von Datensätzen lassen sich – vor allem mithilfe von Metadaten – sogar ganze Bevölkerungsprofile erstellen. Auch wenn bei deutschen Überwachungsmaßnahmen der Einsatz von Künstlicher Intelligenz noch weniger relevant ist, wohin diese Reise gehen kann, zeigt uns ein Blick auf das chinesische Social-Scoring-Programm. Massenhafte und tiefgreifende Überwachungsbefugnisse und Datenbevorratung können nicht nur der Demokratie auf vielfältige Weise schaden.

Das Verhältnis von Bürger\*innen und Staat ist stets verknüpft mit einhergehenden Fragen von Machtverteilung, auch im Spannungsverhältnis von Pflicht und Schutz. Mit letzterem wurden in den vergangenen Jahren immer mehr Maßnahmen begründet, die vermeintlich Sicherheit schaffen sollen, doch nicht nur faktisch das Gegenteil herbeiführen können, sondern mit Blick auf ihre Auswirkungen, zum Teil eine starke Dysbalance in der Verhältnismäßigkeitsprüfung aufweisen. Wir sehen eine lange Kette an schrill blinkenden Warnleuchten. Sie wurde sichtbar in den Veröffentlichungen von Edward Snowden über das weltumspannende Überwachungssystem. Parlamentarische Untersuchungsausschüsse haben aufge-

klärt und ihre Auswirkungen sind in wissenschaftlichen Gutachten belegt. Doch wer überwacht die Überwacher\*innen, wenn die Warnleuchten nicht für jede\*n gleichermaßen wahrnehmbar sind und das Internet nicht selten faktisch zum rechtsfreien Raum erklärt wird?

Die Ausweitung von staatlichen Überwachungsmaßnahmen, die zu einem kaum überschaubaren, undurchsichtigen Flickenteppich von Ermächtigungsgrundlagen anwächst, wurde lange Zeit rechtlich kaum in ihrer Gesamtheit betrachtet. Nun wurde sie flankiert von einer klaren Kante seitens der Judikative. Neben der formulierten Notwendigkeit von Konkretisierungen hinsichtlich der Differenzierung von Eingriffsbefugnissen je nach Eingriffsgewicht in der Entscheidung der Bundesverfassungsgerichts zur Bestandsdatenauskunft II (Beschl. v. 27.05.2020, BvR 1873/13, BvR 2618/13) hat erst jüngst eine Klage von zivilgesellschaftlichen Organisationen die Richter\*innen in Karlsruhe dazu gebracht, das nachrichtendienstliche Recht auf die Zulässigkeit anlassloser Überwachung des weltweiten Internetverkehrs durch den Bundesnachrichtendienst zu prüfen. Es sollte kein überraschendes Ergebnis sein: Die Menschenrechtswidrigkeit und Unvereinbarkeit von anlassloser Massenüberwachung mit der deutschen Verfassung wurde dabei festgestellt (BVerfG, Urtl. v. 19.05.2020, 1 BvR 2835/17). Das tatsächlich mit Ersterem einhergehende Novum: Eine Klarstellung hinsichtlich der bisherigen Ungleichbehandlungen von Staatsbürger\*innen und „Ausländern im Ausland“, deren Kommunikation zuvor „zum Abschuss freigegeben war“ wie es ein BND-Mitarbeiter einst im parlamentarischen Untersuchungsausschuss „NSA“ einmal betonte (Sattler 2014).

Doch neben diesem Meilenstein für den Rechtsschutz von Millionen von Menschen bleibt die Frage, ob das verfassungsrechtliche Korrektiv auch faktisch zur Grenzsetzung bzw. Rechtswahrung führt. Statt sich im Rahmen einer bereits 2010 verfassungsrechtlich geforderten unabhängigen Überwachungsgesamtrechnung (BVerfG, Urt. v. 02.03.2010 1 BvR 256/08, primär Rn. 218) der Frage zu widmen, wie die bestehende Überwachungspraxis deutscher Sicherheitsbehörden im Gesamten ausgestaltet ist und welche Folgen und Konsequenzen derzeitige Überwachung mit sich bringt, blicken wir auf eine sich fortschreibende Entwicklung des Überwachungssystems im Sinne der stetigen Ausweitung von Maßnahmen – trotz unzähligen höchstgerichtlichen klaren Entscheidungen zur Vorratsdatenspeicherung, präventiven Telekommunikationsüberwachung oder Online-Durchsuchung.

Erst dieser Tage wird versucht, auslaufende Überwachungsmaßnahmen (Entwurf eines Gesetzes zur Entfristung von Vorschriften zur Terrorismusbekämpfung, BT-Drs 19/23706), die ursprünglich in den sogenannten Schily-Katalogen mit einer Befristung und der Pflicht stetiger unabhängiger Evaluierungen versehen waren, nach zahlreichen Zwischenentfristungen nun endgültig zu entfristen. Unabhängige Evaluierungen bis dato: Nicht vorhanden. Im Gesetzespaket enthalten sind Rechtsgrundlagen für Überwachungsmaßnahmen, die längst vom Verfassungsgericht einkassiert wurden. Eine Überwachungsgesamtrechnung gibt es bis heute nicht, noch nicht mal einen Kriterienkatalog für eine dringende notwendige wissenschaftliche Evaluierung aller Überwachungsgesetze (Dolderer 2020). Es ist schlicht zu wenig bekannt über die von Sicherheitsbehörden eingesetzten Technologien, die Eignung oder Erforderlichkeit der mit ihr einhergehenden weitreichenden Grundrechtseingriffe. Und ohne

Technikfolgenabschätzung bleiben die Auswirkungen auf Individuum und Gesellschaft im Dunkeln.

Das stets vorgebrachte Argument: Die Sicherheit. Doch ohne sich der irreführenden Sicherheits-Freiheits-Dichotomie bedienen zu wollen: Es gibt keine Sicherheit, die ohne Freiheit, ohne Gleichberechtigung auskommt.

Daher mahnte bereits 1983 der Verfassungsgerichtshof in seiner immer noch hochaktuellen Entscheidung zum Volkszählungsurteil an, die Auswirkungen von Überwachung auf Individuen und Gesamtgesellschaft zu überprüfen:

Wer unsicher ist, ob abweichende Verhaltensweisen jederzeit notiert und als Information dauerhaft gespeichert, verwendet oder weitergegeben werden, wird versuchen, nicht durch solche Verhaltensweisen aufzufallen. Wer damit rechnet, daß etwa die Teilnahme an einer Versammlung oder einer Bürgerinitiative behördlich registriert wird und daß ihm dadurch Risiken entstehen können, wird möglicherweise auf eine Ausübung seiner entsprechenden Grundrechte (Art. 8, 9 GG) verzichten.

(Bundesverfassungsgericht 1983)

Eine visionäre Begründung, die in ihrer die Unsicherheiten aufzeigenden und mahnenden Weise, die (unterbewusst) handlungsverändernde Komponente von Überwachungsmaßnahmen (*Chilling Effects*) in den Mittelpunkt rückt. Eingeschränkt wird damit auch die Chance auf selbstbestimmte, individuelle Entfaltung. Aufgrund der Machtverhältnisse der Überwachungsmaßnahmen und bestehender Diskriminierungen ist die Beeinträchtigung dabei für einige weniger, für andere deutlicher intensiv. Die gesellschaftlichen Auswirkungen auf unsere freiheitlich-demokratische Gesellschaftsstruktur liegen auf der Hand, wenn Handlungs- und Mitwirkungsfähigkeit eingeschränkt werden. Wenn die Einschränkungen unterschiedlicher Art sind und uns als Gesamtgesellschaft unterschiedlich betreffen, verändert dies uns als vielfältige und diverse Gesellschaft grundlegend und verändert demokratische Prozesse wie politische Meinungsbildung. Dass solche sogenannten *Chilling Effects* nur schwer vor Gericht darstellbar sind, folglich eine rechtliche Überprüfung fast unmöglich erscheint, ist ein Fakt. Es ist also auch die grundlegende Frage, wie wir der Lösung dieser damit einhergehenden gesellschaftlichen Aufgabe nachkommen. In welchem Miteinander wollen wir leben? Wer ist das Wir und wer entscheidet? Ohne die Wahrnehmung der staatlichen Pflicht, diesen Entscheidungsprozess offen und transparent zu begleiten, interdisziplinär statt abgeschottet und elitär zu führen und zu entscheiden, werden wir die Lücke zwischen gesamtgesellschaftlichem Anspruch, unserem bestehenden Grundgesetz und der faktischen Lage von ebenfalls bestehenden Grundrechtsverletzungen aufgrund von Überwachung nie schließen und ohnehin nie aufdecken können.

Das, in Überwachungsmaßnahmen logischerweise selbst angelegte Problem stellt die in ihr verankerte Intransparenz dar. Während ein kleiner Teil der Gesellschaft immer mehr Informationen und damit Macht ansammelt, weiß der Großteil der Bevölkerung schlicht zu wenig über diese Vorgänge. Verstärkt wird dies durch den Einsatz von Technologien, wenn diese selbst mit fehlender Transparenz gebaut, mit Daten befüllt, eingesetzt und letztlich die Ergebnisse nicht evaluiert werden.

Bleibt also letztlich nur die digitale Selbstverteidigung als letzte Möglichkeit des Eigenschutzes? Klar ist, dass dies gerade nicht zu gleichberechtigter Teilhabe beitragen würde. Zynisch auch gerade mit Blick auf das Machtverhältnis



von Bürger\*in und Staat, denn das Verhältnis zur Verschlüsselung hat selbst der Staat bis heute nicht geklärt. Er nimmt oder gibt sich immer mehr das Recht staatlichen Durchgriffs, wie beispielsweise mit Hilfe von Staatstrojanern. Während zivilgesellschaftliche Organisationen wiederum durch gerichtliche Klärung einen Überwachungsstopp einfordern müssen (Mattes 2018), wird der Versuch der Ausweitung von staatlicher Seite fortgeführt, ja sogar soweit getrieben, Unternehmen zur Überwachungsunterstützung durch Verteilung von Malware und somit zum Hacken der Kund\*innen in die Pflicht zu nehmen (Meister & Biselli 2019).

# MENSCH ODER KI WER MUSS EIGENTLICH ÜBERWACHT WERDEN?

Und wer sollte diese Frage beantworten? Nicht nur, wenn es um anlasslose Massenüberwachung geht, können die grundlegenden Entscheidung wohl nur von der Gesamtgesellschaft getroffen werden. In welcher Welt wollen wir leben, in welchem Verhältnis, wie soll Macht verteilt sein, wer gehört dazu und wer ist draußen? Bei Massenüberwachung wissen wir nicht, wer von der Überwachung in welchem Umfang betroffen ist. Oder doch? Wir überwachen schlicht alle per Generalverdacht, um das einzelne Überwachungsobjekt erst später herauszudefinieren. Ein grundsätzliches *Nein* hierzu fällt leicht, doch die Grenzen zu ziehen wohl faktisch nicht.

Bei individueller Einzelüberwachung, Überwachung im privaten Bereich oder durch wirtschaftliche Unternehmen scheint die Frage, welche Formen von Überwachung wir wollen und welche nicht, einfacher zu beantworten zu sein. Nicht nur rechtlich, sondern auch faktisch. Doch bereits am proaktiven Ansatz zeigt sich das Scheitern.

Es geht nicht um ein Gefühl? Oder doch? Um der mit Verwendung von Technologien einhergehenden Verantwortung gerecht zu werden, müssen wir weg von der vielleicht unbewusst getroffenen Entscheidung, dass es ausreicht, Gefühle zur Legitimation von gesellschaftlichen Entscheidungen bis hin zu Rechtsgrundlagen heranzuziehen. Weder für die Entscheidung nach stärkeren und umfangreichen Überwachungsmaßnahmen zur Überdeckung von Angstgefühlen, noch für ein unreflektiertes Vertrauen in Technik. Dies gilt insbesondere für den zunehmenden Einsatz von Künstlicher



Intelligenz, die häufig zu falschen Entscheidungen führen und Diskriminierung verstärken kann, wie wir gezeigt haben. Black-boxes, also intransparente, in sich geschlossene Systeme, deren Aufbau und innerer Ablauf sich – wenn überhaupt – erst aus den Reaktionen auf eingegebene Signale erschließen lässt, sind zur Überwachung verwendet eine doppelte Bedrohung.

Das bedeutet, wir müssen die richtige überwatchen: Die Künstliche Intelligenz.

Damit bereits bestehende Diskriminierung und systemische Verzerrungen durch KI-Systeme nicht verstärkt, sondern minimiert werden können, braucht es eine grundlegende Transparenz bei ihrer Ausgestaltung. Von der offenen Kennzeichnung des Einsatzes solcher Systeme beginnend, müssen sie so nachvollziehbar und überprüfbar wie möglich entwickelt und ausgestaltet sein. *Explainable AI* ist hier das Stichwort.

Eine Dokumentierungs- und Protokollierungspflicht erscheint unausweichlich, um Fehler und Falschentscheidungen aufdecken und beheben zu können. Beim Einsatz von Künstlicher Intelligenz müssen nicht nur unbeabsichtigte Rückschlüsse auf einzelne Personen verhindert werden, es bedarf auch zwingend kontrafaktischer Erklärungen wie Auskunftsansprüche für Betroffene darüber, welche Faktoren zu einer negativen Entscheidung geführt haben.

Je höher das Schädigungspotential, desto höhere Anforderungen sind an die Kriterien von KI-Systemen anzulegen. Doch was bringen offene Standards, Datenqualität und -robustheit oder Datenschutzregeln, wenn eine KI zur heimlichen Überwachung eingesetzt wird? Was bringt es Betroffenen, wenn ihnen in Bezug auf Künstliche Intelligenz-Entscheidungen Rechte zustehen, wenn sie diese aber aufgrund von fehlender Kenntnis der Überwachung nicht durchsetzen können. Bedeutet das, dass beim Einsatz von Überwachung der Einsatz von Künstlicher Intelligenz gar ganz verboten werden muss? Klar scheint, dass bei Gesichtserkennungstechnik im öffentlichen Raum etwa von einem nicht vertretbaren Schädigungsrisiko gesprochen werden muss. Mit Blick auf die Sensibilität der meisten betroffenen Bereiche, die Eingriffsintensität, die Anzahl betroffener Personen oder die Irreversibilität von Entscheidungen, wäre das für viele Bereiche zu bejahen. An dieser Stelle muss gesagt werden, dass es durchaus Einsatzbereiche von Überwachung gibt, etwa im medizinischen Bereich, in denen die Abwägung eine andere wäre.

Aber wer übernimmt die Verantwortung, zu entscheiden, was uns Technologie in der Zukunft bringen soll und was nicht? Wer entscheidet, welche Weichen die richtigen sind und wo sie hinführen sollen und wer verantwortet sie letztlich? Wer entwickelt, zertifiziert oder standardisiert? Und welche gesellschaftlichen wie privaten Auswirkungen hat die Antwort? Hierfür braucht es eine grundlegende Entscheidung – persönlich und auch gesamtgesellschaftlich. In Deutschland, Europa, aber auch weltweit. Überwachung geht uns alle an. Sich aus dem Diskurs herausnehmen hat Auswirkungen, meist in größerem Maße für all jene, die besonders von Überwachung und Einschränkungen betroffen sind. Wir müssen diese Debatte intersektional feministisch gestalten und führen, also unter Einbeziehung von Macht- und Herrschaftsstrukturen, mit Blick auf die unterschiedlichen Diskriminierungsformen und Auswirkungen. Anders können wir die Fragen über das Ob – von Überwachungsmaßnahmen mit oder ohne KI – hin zum Wie und mit welchem Ergebnis und welchen Auswirkungen nicht gerecht für die Zukunft beantworten.

Francesca Schmidt arbeitet als Referentin für feministische Netzpolitik am Gunda-Werner-Institut für Feminismus und Geschlechterdemokratie und hat gerade ein Buch zu feministischer Netzpolitik im Budrich Verlag herausgegeben.

Johanna Luise Mellentin ist Ethnologin mit rechtswissenschaftlichem Background, auf kommunaler Ebene digitalpolitische Referentin im Stadtrat Fürstentfeldbruck, Sprecherin der Landesarbeitsgemeinschaften Digitales und Medien bzw. Netzpolitik von Bündnis 90/Die Grünen in Bayern als auch Berlin und arbeitet zu diesen Themen als wissenschaftliche Mitarbeiterin in einem grünen Bundestagsbüro.

# LITERATUR

- Barr, Alistair (2015)  
„Google Mistakenly Tags Black People as ‚Gorillas‘; Showing Limits of Algorithms.“ In: *The Wallstreet Journal*. Verfügbar unter: <http://blogs.wsj.com/digits/2015/07/01/google-mistakenly-tags-black-people-as-gorillas-showing-limits-of-algorithms/>
- boyd, danah & Crawford, Kate (2012)  
CRITICAL QUESTIONS FOR BIG DATA: Provocations for a cultural, technological, and scholarly phenomenon. In: *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Browne, Simone (2015)  
*Dark matters: on the surveillance of blackness*. Durham: Duke University Press.
- Bundesverfassungsgericht, 1 Senat (1983)  
„Bundesverfassungsgericht – Entscheidungen – Zur Verfassungsmäßigkeit des Volkszählungsgesetzes 1983.“ In: *Bundesverfassungsgericht*. Zugriff am 8.11.2020. Verfügbar unter: [https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/1983/12/rs19831215\\_1bvr020983.html](https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/1983/12/rs19831215_1bvr020983.html)
- Bundeszentrale für politische Bildung (2012)  
„8. März: Weltfrauentag / bpb.“ In: *bpb.de*. Zugriff am 6.11.2020. Verfügbar unter: <https://www.bpb.de/politik/hintergrund-aktuell/76102/8-maerz-weltfrauentag-07-03-2012>
- Buolamwini, Joy und Gebru, Timnit (2018)  
„Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.“ In: Sorelle A. Friedler & Christo Wilson (Hrsg.): *Proceedings of Machine Learning Research* (Band 81). *Proceedings of Machine Learning Research*: PMLR, 77–91.
- Dialog Versicherung AG (o. J.)  
„Generali Vitality.“ In: *Dialog*. Zugriff am 7.11.2020. Verfügbar unter: <https://www.dialog-versicherung.de/produkte/privatkunden-leben/generali-vitality>
- Dolderer, Dr Winfried (2020)  
„Bedenken gegen Entfristung von Vorschriften zur Terrorismusbekämpfung.“ In: *Deutscher Bundestag*. Zugriff am 8.11.2020. Verfügbar unter: <https://www.bundestag.de/dokumente/textarchiv/2020/kw45-pa-innen-antiterrorgesetze-799842>
- Eggers, Maureen Maisha; Kilomba, Grada; Piesche, Peggy & Arndt, Susan (2005)  
„Konzeptionelle Überlegungen.“ In: Maureen Maisha Eggers, Grada Kilomba, Peggy Piesche & Susan Arndt (Hrsg.): *Mythen, Masken und Subjekte. Kritische Weißseinsforschung in Deutschland*. Münster: Unrast, 11–13.

- Eggers, Maureen Maisha; Kilomba, Grada; Piesche, Peggy & Arndt, Susan (Hrsg.) (2005)  
*Mythen, Masken und Subjekte: kritische Weissseinsforschung in Deutschland*. Münster: Unrast.
- Foucault, Michel (1977)  
*Überwachen und Strafen: Die Geburt des Gefängnisses*. (W. Seitter, Übers.). Frankfurt am Main: Suhrkamp Verlag.
- Gless, Sabine (2016)  
 „Predictive policing und operative Verbrechensbekämpfung.“ In: Felix Herzog, Reinhold Schlothauer & Wolfgang Wohlers (Hrsg.): *Rechtsstaatlicher Strafprozess und Bürgerrechte, Gedächtnisschrift für Edda Weßlau*. Berlin: Duncker & Humblot, 165–180. Zugriff am 25.2.2018. Verfügbar unter: <http://edoc.unibas.ch/44207/>
- Gorner, Jeremy (2013)  
 „Chicago police use heat list as strategy to prevent violence.“ In: *Chicago Tribune*. Zugriff am 25.2.2018. Verfügbar unter: [http://articles.chicagotribune.com/2013-08-21/news/ct-met-heat-list-20130821\\_1\\_chicago-police-commander-andrew-papachristos-heat-list](http://articles.chicagotribune.com/2013-08-21/news/ct-met-heat-list-20130821_1_chicago-police-commander-andrew-papachristos-heat-list)
- Grieger, Fabian und Schindwein, Simone (2016)  
 „Migrationspolitik und Rüstungsindustrie: Das Geschäft mit Hightech-Grenzen – taz.de.“ In: *die tageszeitung*. Zugriff am 8.11.2020. Verfügbar unter: <https://taz.de/Migrationspolitik-und-Ruestungsindustrie/!5363960/>
- Lyon, David (Hrsg.) (2003)  
*Surveillance as Social Sorting: Privacy, Risk and Automated Discrimination*. London: Routledge.
- Mattes, Anna Livia (2018)  
 „Pressemitteilung: Verfassungsbeschwerde gegen Staatstrojaner eingelegt – GFF – Gesellschaft für Freiheitsrechte e.V.“ In: *Gesellschaft für Freiheitsrechte*. Zugriff am 8.11.2020. Verfügbar unter: <https://freiheitsrechte.org/pm-vb-trojaner/>
- Meister, Andre & Biselli, Anna (2019)  
 „Wir veröffentlichen den Gesetzentwurf – Seehofer will Staatstrojaner für den Verfassungsschutz.“ In: *netzpolitik.org*. Zugriff am 8.11.2020. Verfügbar unter: <https://netzpolitik.org/2019/wir-veroeffentlichen-den-gesetzentwurf-seehofer-will-staatstrojaner-fuer-den-verfassungsschutz/>
- Merz, Christina (2016)  
*Predictive Policing – Polizeiliche Strafverfolgung in Zeiten von Big Data*. Karlsruher Institut für Technologie (KIT).
- Sattler, Karl-Otto (2014)  
 „Deutscher Bundestag – Keine Informationen an die NSA übermittelt.“ In: *Deutscher Bundestag*. Zugriff am 8.11.2020. Verfügbar unter: [https://www.bundestag.de/dokumente/textarchiv/2014/kw46\\_pa\\_1ua-339510](https://www.bundestag.de/dokumente/textarchiv/2014/kw46_pa_1ua-339510)
- Snow, Jacob (2018)  
 „Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots.“ In: *American Civil Liberties Union*. Zugriff am 11.8.2019. Verfügbar unter: <https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28>





KONTROLVERLUST  
UND  
(DIGITALE)  
ENTMÜNDIGUNG

---

DAS  
GEWALTPOTENTIAL  
KÜNSTLICHER  
INTELLIGENZ



LEENA  
SIMON

Wann hatten Sie das letzte Mal mit einer Künstlichen Intelligenz (KI) zu tun? Wenn es Ihnen schwerfällt, diese Frage zu beantworten, mag das daran liegen, dass der Begriff zwar vielen geläufig ist, seine Definition in der Öffentlichkeit aber eher schwammig daherkommt. Das, was es gar nicht gibt (starke KI), wird überall diskutiert, aber das, was real passiert (schwache KI), ist eher selten Thema. Aber der Reihe nach: In der öffentlichen Wahrnehmung, in Filmen oder Gedankenexperimenten ist die sogenannte „starke KI“ überproportional häufig Thema. Dabei gibt es sie real gar nicht. Starke KI hat das Ziel, die menschliche Intelligenz ebenbürtig zu imitieren. Denken wir bei der Eingangsfrage ausschließlich an starke KI, werden uns wohl eher keine Begegnungen einfallen. Anders ist es mit der „schwachen KI“. Mit der kommen wir regelmäßig in Berührung. Hier können wir aus einem anderen Grund schwer sagen, wann wohl die letzte Begegnung war. Denn erstens passiert das wirklich überaus häufig und zweitens ist es uns nur selten bewusst. Das Wesen der schwachen KI ist in der Öffentlichkeit nur selten Thema, weshalb kaum Allgemeinwissen dazu besteht. Auch KI besteht letztlich nur aus Algorithmen. Ihre Besonderheit besteht darin, dass sie ihre Wenn-Dann-Beziehungen nicht mehr von uns vorgegeben kriegen, sondern sich diese selbst herleiten. Ein simpler Algorithmus wird mit allen möglichen Varianten, die eintreten können, und der Anweisung, was im jeweiligen Fall geschehen soll, gefüttert. Eine KI wird mit den Ergebnissen gefüttert und leitet sich die verschiedenen Varianten selbst ab. Das hat den Vorteil, dass auch unvorhergesehene Situationen von einer KI berechnet werden können. Es hat aber den großen Nachteil, dass hier falsche Schlüsse entstehen können, die wir leider nur sehr schwer entdecken können, da wir nicht mehr nachvollziehen können, wie die Entscheidung eigentlich zu Stande kam.

# UNWISSENHEIT SCHÜTZT VOR KONTROLLE NIGHT

Es ist äußerst ungünstig, dass die Menschen so wenig von der KI wissen, die sie umgibt. Denn sie trifft maßgebliche Entscheidungen. Können wir diese Entscheidungsprozesse nicht mehr nachvollziehen, können wir auch nicht mehr beurteilen, ob sie problematisch sind. Wer auch in einer digitalen Welt seine Mündigkeit bewahren möchte, sollte daher anfangen über Transparenz bei KI zu sprechen. Dieser sind gleich mehrfach Steine in den Weg gelegt. Erstens lassen sich Firmen wie Google, Facebook und Amazon eher ungerne in die Karten

schauen. Und zweitens sind die Prozesse, die hinter KI stehen, so komplex, dass es immer schwerer wird, sie auch nur im Ansatz nachzuvollziehen, selbst wenn sie offen vorliegen.

Doch ohne Transparenz gibt es keine Möglichkeit, die Technik zu kontrollieren, wodurch wir ihr und damit denen, die sie steuern/einsetzen/gestalten, ausgeliefert sind.

Wussten Sie, dass sich nach Erscheinen eines Buches innerhalb der ersten drei Tage entscheidet, ob es ein Bestseller wird? Grund dafür ist der Amazon-Algorithmus, der anhand der Verkaufszahlen der ersten drei Tage entscheidet, ob er ein Buch auf seine Hitlist nimmt. Hat ein Buch es da rauf geschafft, wird es von Amazon häufig empfohlen und entsprechend häufiger gekauft. Bücher, die vielleicht etwas später gezündet hätten, sind dann schon aus dem Rennen. Indem Amazon die Kriterien bestimmt, welche Bücher auf die Hitliste dürfen, nimmt es maßgeblich Einfluss auf die Bücher, die tatsächlich verkauft und gelesen werden. Die kapitalistisch ausgerichteten Algorithmen von Google, Facebook, Amazon und Co. haben unsere Welt erheblich fester in der Hand, als den meisten Menschen klar sein dürfte. Die Google-Tochter YouTube empfiehlt Videos aus ganz anderen Gründen, als sie uns weiß machen will. Es geht nur vordergründig um das, was uns interessieren könnte. In Wahrheit geht es um das, was uns dazu bewegt, möglichst lange Zeit bei YouTube zu verbringen und der Werbung ausgesetzt zu sein. Dass dies häufig dadurch erreicht wird, dass der empfohlene Inhalt uns auch gefällt, ist eher ein Abfallprodukt, welches fürs Marketing recycelt und als Hauptziel verkauft wird. Weil es vor allem die Verweildauer erhöhen will, empfiehlt Youtube besonders häufig Videos zu Verschwörungsideologien, die eine besonders fesselnde Wirkung entfalten, sobald man ihnen auf den Leim geht. KI wird häufig dazu eingesetzt, uns erst zu erforschen und dann zu manipulieren.

YouTube hat an dieser Stelle eine Kurskorrektur angekündigt, von der bisher aber nicht viel zu sehen ist. Dieses (bisher noch oberflächliche) Einlenken entstand aber nicht etwa aufgrund einer internen Ethik-Kommission, die die Wirkung der eigenen Algorithmen bewertet und auf die problematische Wirkung aufmerksam wurde, sondern aufgrund eines starken öffentlichen Drucks. Öffentlicher Druck entsteht aber nur bei den ganz großen Aufregern und setzt voraus, dass die Öffentlichkeit auf die problematische Wirkung einer Software aufmerksam wird. Das ist ohne Transparenz kaum möglich und die Probleme werden nur bruchstückhaft erfasst.

# WELCHES VIDEO WIRD EMPFOHLEN?

Was wäre, wenn der Algorithmus  
mir dienen würde?



Passt das Video zu  
meinen Interessen?

Nein

ähnelt anderen Videos, die ich mag

Themenauswahl passt

informativ

lustig

Spam

Sexismus

Rassismus

Hass oder Verschwörungsnarrative

Ja

♥  
das Video  
dürfte mir  
gefallen

Video wird  
vorgeschlagen



Was wirklich passiert..

anspruchsvoll oder komplex

Youtube Ranking

Video wird nicht  
vorgeschlagen

Verschwörungsnarrative

viel angeklickt

ähneln bereits angesehenen Videos

Lügen oder Fakes

Hass

lustig

lange  
Verweildauer

wirksame  
Werbung

viel  
Interaktion

\$

Profite für  
Youtube

weitere Informationen:  
<https://muendigkeit.digital>

# PROBLEMATISCHE NORMEN UND WERTE WERDEN IN DER TECHNIK VEREWIßT

Nüchtern betrachtet sind Algorithmen nichts weiter als Gleichungen und Berechnungen. Erst durch die menschlichen Werte, die sich darin abbilden und dann zu folgensweren Ergebnissen führen können, wird die Sache kompliziert. Je komplexer die Berechnungen werden und je mehr sie geheim gehalten werden, desto schwieriger ist es, diese nachzuvollziehen. Doch wir müssen sie nachvollziehen können. Denn sie entscheiden zunehmend, was wir denken, lesen/sehen/hören und tun.

Aus demokratischer Sicht ist das äußerst brisant. Denn auch subtil manipulierte Menschen sind unmündig und treffen auch bei Wahlen keine freien Entscheidungen mehr. Das muss den Betroffenen dabei nicht mal bewusst sein. Es gilt beispielsweise als erwiesen, dass sich Russland zu Trumps Gunsten in die US-Wahlen eingemischt hat. Dies geschieht z.B. über Falschnachrichten oder Diskursverschiebungen bzw. komplett fingierte Diskussionen. Doch selbst mit diesem Wissen lässt sich im Nachhinein kaum mehr feststellen, wie viele Menschen aufgrund dieser Einmischung ihre Wahlentscheidung geändert haben. Selbst die Betroffenen selbst dürften sich dessen kaum bewusst sein.

Außerdem neigen Menschen, die ahnen, dass sie manipuliert werden, aber nichts über die genauen Umstände wissen, eher dazu Verschwörungsideologien zu glauben und zu verbreiten. Der Eindruck von Kontrollverlust macht uns anfällig dafür bei der Suche nach Halt nicht mehr so genau zu prüfen, welchem Hirngespinnst wir auf den Leim gehen. Eine falsche Antwort ist vielen Menschen lieber als gar keine Antwort. Wer weiß denn schon wirklich sicher, welche Informationen sein Smartphone so den lieben Tag lang verschickt? Viele Menschen ziehen hier eine schlimme Antwort (Du wirst den ganzen Tag überwacht.) einer vagen Antwort vor (Theoretisch wäre es möglich, dass du überwacht wirst, aber wir können es nicht wissen.).

Hinzu kommt, dass gewisse Staatsoberhäupter und Medienvertreter „alternative Fakten“ – auch „Lügen“ genannt – gleichberechtigt neben wissenschaftliche Erkenntnisse stellen. Sobald aber Wissenschaft und Wahrheit nicht mehr von Meinungsäußerungen und unbelegten Behauptungen getrennt werden können, wird es keine gesamtgesellschaftliche Erkenntnis mehr geben können. Ohne eine gemeinsame Basis von als erwiesen anerkannten Wahrheiten, kann eine Demokratie keinen sinnvollen Entscheidungsprozess mehr führen.

# MÜNDIGKEIT IST EIN MUSKEL, DER TRAINIERT WERDEN WIL

Wir haben uns das Internet falsch anerzogen. Stets bringen wir uns nur bei, wie Technik zu bedienen ist, aber nicht, wie man sie hinterfragt und versteht. In Office-Kursen lernen wir, wie man einen Text fett macht, wie das Inhaltsverzeichnis funktioniert oder wie man eine Tabelle erstellt. Wir lernen aber nicht, dass es verschiedene Dateiformate gibt und warum es nicht sehr solidarisch ist, unfreie Dateiformate (wie docx, ppt oder xls) zu verwenden und dass man besser die freien Formate (odt, pdf, ppt) nutzen sollte. Wir sprechen nicht darüber, dass ich mit dem Versenden einer Datei mein Gegenüber quasi dazu nötige, die entsprechende Software dazu anzuschaffen. Über die gesellschaftlichen Folgen der Digitaltechnik, die wir nutzen, nachzudenken, ist selten Teil eines Computerkurses oder einer Betriebsfortbildung. Wir haben uns gar nicht erst angewöhnt, solche Fragen zu stellen. Es soll einfach funktionieren und gut ist.

Wenn unser Anspruch, Technik zu verstehen, aber mit der Fähigkeit, sie zu bedienen, erlischt, werden wir nicht berücksichtigen können, welche Folgen der Gebrauch dieser Technik haben könnte. Wir können keine Verantwortung für unsere digitalen Handlungen übernehmen, da wir sie dafür nicht ausreichend durchblicken. Digitale Mündigkeit muss geübt werden, doch es gibt kaum Gelegenheit dazu. Längst sind digitale Dienste mehr und mehr danach ausgerichtet, uns möglichst viele Entscheidungen abzunehmen. Das ist bequem, denn es erleichtert tatsächlich vieles. Ohne Vereinfachungen wäre es gar nicht möglich, Computer in unseren Alltag zu integrieren. Aber es zieht uns gleichzeitig immer weiter rein in einen Strudel der Unmündigkeit, den wir immer weniger kontrollieren können. Da klingen die Worte vom ollen Kant wieder auf und hallen in neuem Klang: „Es ist so bequem unmündig zu sein.“

# KI BRAUCHT MITBESTIMMUNG ODER MITBESTIMMUNG IST DER KI KEY

Wenn KI so entscheidende Auswirkungen auf unser gesamtgesellschaftliches Zusammenleben hat, warum dürfen wir dann eigentlich nicht mitentscheiden, auf welchen Kriterien sie beruht? Google hat entschieden, dass Seiten, die häufig verlinkt werden, barrierefrei sind und keine Textdopplungen zu anderen Seiten aufweisen, weiter oben bei den Suchergebnissen aufgelistet werden. Das basiert mitunter auf stützswerten Überlegungen. Häufige Verlinkung ist ein Hinweis darauf, dass viele die Seite für relevant halten, Barrierefreiheit und sauberer Code sind sinnvolle Werte, die durch solche Vorgaben gepusht werden und Textdopplungen können darauf hinweisen, dass es sich um eine Seite handelt, die einfach nur den Text von anderen kopiert hat und versucht für Werbeeinnahmen zu nutzen. Doch all diese Faktoren haben auch eine Kehrseite: Seiten von Frauen werden seltener verlinkt als die von Männern. Verlinkungen werden munter hin und her verkauft, was bedeutet, dass diejenigen als relevanter eingestuft sind, die männlich sind und reich bzw. skrupellos genug, sich Verlinkungen zu kaufen. Die Strafe für Textdopplungen trifft leider häufig die Falschen. Dazu gehören die ursprünglichen Urheber:innen des Textes, der auf eine andere Seite kopiert wurde und Organisationen, die mit Crosspostings arbeiten. Ob es in unserem Sinne ist, dass hier so viele Kollateralschäden in Kauf genommen werden, werden wir bequemerweise nicht gefragt. Selbst die wirklich guten Vorgaben von Google, wie die Vorgaben zur Barrierefreiheit, kommen mit dem schalen Beigeschmack, dass dies Regeln sind, die eine Firma diktiert und die auf keinerlei demokratischem Prozess basieren. Jüngst haben Google und Apple eindrücklich gezeigt, wie einfach es für sie ist, ihre persönliche Weltanschauung einer ganzen Gesellschaft aufzudrücken. Für die Corona-Warn-App wurde lange diskutiert, ob diese einen zentralen oder einen dezentralen Ansatz fahren solle. In der Bundesregierung hatte man sich eigentlich schon für das zentrale Modell entschieden, als Google und Apple durch die Schnittstellen zu ihren Smartphones Fakten schufen und den dezentralen Ansatz durchsetzten. Wieder ist es nicht der Inhalt der Entscheidung, der uns in diesem Fall aufschrecken sollte, sondern der Umstand, dass diese Firmen so viel Macht über solche Fragen haben.

# MONETARISIERUNG VON INHALTEN GEFÄHRDET DIE DEMOKRATIE

Bleiben wir einen Augenblick bei Google. Wer heute ein Unternehmen gründet, kommt an Google nicht mehr vorbei. Produkte werden heute anhand der Auswertung von Google-Suchen erzeugt. Mit speziellen Tools kann man herausfinden, welche Fragen besonders häufig in die Suchmaschine eingegeben wurden. Dann braucht man nur noch einen Online-Artikel, der verspricht die entsprechende Frage zu beantworten und hat das nächste Produkt erzeugt. Ob die Frage auch tatsächlich beantwortet wird, spielt dabei weniger eine Rolle als der Werbelink, der im Text steht. Werbung, die als freundlicher Infobeitrag daherkommt, wird von vielen Menschen nicht als solche erkannt. Das Konzept funktioniert so gut, dass die echten Infoseiten von Leuten, die das tatsächlich tun, um ihr Wissen zu teilen, drohen in der Masse von Fake-Infos unter zu gehen. Derweil verstecken immer mehr Medienhäuser ihren recherchierten und fundierten Journalismus hinter Accountwalls. Das Verhältnis zwischen hochwertigen und minderwertigen Informationsangeboten droht in ein gefährliches Ungleichgewicht abzugleiten. Während Fakes und Verschwörungsideologien – durch Algorithmen noch unterstützt – immer größere Verbreitung erfahren, sind die Gegendarstellungen immer seltener frei verfügbar. Wer Informationen mittels Künstlicher Verknappung kommerzialisieren will, hat die Idee des Internets nicht verstanden. Wenn wir die hochwertigen Informationen verknappen, führt das letztlich nur dazu, dass sie inmitten der ganzen irreleitenden und falschen Aussagen nicht mehr zu finden sind. Dann werden Gegendarstellungen oder differenzierte Erklärungen nur noch von den wenigen gelesen, die dafür zu zahlen bereit sind. Sie können sich genug Zeitungsabos leisten, um im System der Online-Zeitungen nicht wahnsinnig zu werden? Gratulation! Das ändert aber auch nicht viel. Denn Sie leben immer noch in einem Land, in dem hauptsächlich die Menschen wählen gehen, die keine 20 Abos für Online-Zeitungen abgeschlossen haben und sich hauptsächlich auf Facebook informieren. Politik, die aus solchen Wahlen resultiert, können wir in Groß-Britannien und den USA beobachten.

# KI MAGT FOLGENSCHWERE FEHLER

KI übergibt Definitionsmacht an die, die in der Technikgestaltung den Ton angeben. KI nützt tendenziell denen am meisten, die eh schon an der Macht sind und macht jene unsichtbar, die auch vorher schon marginalisiert waren. Somit birgt KI ein enormes Diskriminierungs- und Gewaltpotential. Das hängt damit zusammen, wie KI angelernt wird. Das Konzept basiert ja gerade darauf, dass wir nicht alle Regeln, die der Algorithmus später mal haben soll, vorab ausformulieren müssen. Die Besonderheit von KI ist, dass sie diese Regeln basierend auf den Materialien, die wir ihr zur Verfügung stellen, selbst herleitet. Das sind meist große Sammlungen von Entscheidungen, die Menschen getroffen haben. Doch Menschen haben Rassismus und Sexismus verinnerlicht. Eine KI, die aus unbewusst diskriminierenden Entscheidungen Verhaltensregeln ableitet, wird hinterher ganz maschinell diese Diskriminierung wiederholen oder sogar neue Formen entwickeln.

Gehen wir einen Schritt zurück und betrachten wir, wie eine KI lernt. Stellen Sie sich vor, Sie möchten eine KI entwickeln, die in der Lage ist, auf Fotos Autos zu erkennen. Das Datenmaterial, das Sie brauchen, um die KI anzufüttern, müssen Sie selbst erzeugen. Sie ziehen also los und fotografieren eine Landschaft zunächst ohne Autos. Dann stellen Sie Autos darauf und fotografieren erneut. Nun geben Sie diese Bilder der KI und teilen ihr mit, auf welchen Bildern Autos sind und auf welchen nicht. Um zu testen, ob das funktioniert hat, machen Sie denselben Fotovorgang noch mal an einer anderen Stelle und – siehe da – die KI kann alle Bilder richtig zuordnen. Testen Sie dies nun mit realen Fotos, haben Sie plötzlich nur noch eine Trefferquote von etwa 50%. Was ist passiert? Die KI hat bemerkt, dass alle Fotos, die am Vormittag aufgenommen wurden, keine Autos abbilden. Alle Fotos, die am Nachmittag aufgenommen wurden, zeigten Autos. Die KI hat also gar nicht das getan, was Sie wollten, sondern eine Abkürzung gewählt, die selbstverständlich gar nicht geht, da sich in der Realität nicht anhand der Uhrzeit entscheidet, ob ein Auto irgendwo steht, oder nicht. Doch woher soll die KI das wissen? Sie hat nach einem Muster gesucht und es gefunden. Der Fehler fällt aber auch bei Ihrem Test nicht auf, da Sie wieder vormittags die Landschaft ohne Autos und nachmittags mit Autos fotografiert haben. Erst wenn Sie die Software auf die reale Welt loslassen, zeigen sich die Probleme. Allerdings ist hier selten so schnell ersichtlich, dass ein Fehler vorliegt. Genau so wenig, wie Sie bemerkten, dass Sie die Fotos nach einem Muster aufnahmen, ist Ihnen wohl eher selten klar, inwiefern Ihre Weltsicht Einfluss auf Ihre Entscheidungen nimmt. Diese integrieren die Hersteller.innen von Technik mitunter in eine KI, ohne dass Sie sich darüber überhaupt im Klaren sind.



# WIE KI DISKRIMINIERUNG ZEMENTIERT

In der Realität führt das dazu, dass bestehende Diskriminierung nicht nur reproduziert, sondern unter dem Eindruck der Neutralität von Maschinenberechnungen zementiert wird. Dafür gibt es unzählige Beispiele. Gesichtserkennung ist nur bei weißen Männern akkurat, da sie vor allem an weißen männlichen Probanden trainiert wird. Besonders schlecht weg kommen schwarze Frauen. Bei ihnen ist die Verwechslungsgefahr besonders hoch. Was bedeutet: Die Gefahr, dass eine schwarze Frau zu Unrecht eines Verbrechens verdächtigt wird, weil sie der gesuchten Person ein wenig ähnlich sieht, ist sehr viel höher, als dass dies einem weißen Mann passiert. Unser Weltbild transportiert sich aber nicht nur über die Probanden, sondern auch über die Daten, die wir füttern. KI wird häufig eingesetzt, um einzuschätzen, ob jemand zu einem Vorstellungsgespräch eingeladen wird. Um diese zu trainieren, füttert man sie gewöhnlich mit den Daten früherer Bewerbungsverfahren. Die KI kann dann analysieren, welche Bewerbungen zu einer Einladung führten und welche Menschen dann letztlich eingestellt wurden. Stellt solch eine KI fest, dass schwarze Bewerber:innen in der Vergangenheit bei gleicher Qualifikation nicht eingestellt wurden, wird sie daraus schließen, dass man diese gar nicht erst zum Vorstellungsgespräch einzuladen braucht. Eine andere KI soll bewerten, wie wahrscheinlich es ist, dass jemand (wieder) straffällig wird. In den USA werden nicht-weiße Menschen erheblich häufiger verurteilt. Was die KI nicht weiß: Das liegt nicht daran, dass sie häufiger straffällig werden, sondern am Rassismus im US-amerikanischen Rechtssystem. Die KI betrachtet ihre Lerndaten aber als „neutral“ und kann nur den Schluss ziehen, dass schwarze Menschen eher straffällig werden.

Derlei KI wird mittlerweile häufig eingesetzt. Ob bei der Frage, ob die Bank Ihnen einen Kredit gewährt, ob Sie zum Vorstellungsgespräch eingeladen werden oder – wie in Österreich – wie viel Mühe sich die Gesellschaft dabei gibt, ihnen zu helfen eine neue Arbeit zu finden. Bei der KI, die in Österreich den Arbeitslosen einen Score-Wert erteilt, wie vermittelbar sie sind, gilt Frausein als negativer Faktor. Es ist eine Sache, das zu erkennen und zu benennen: In Österreich haben arbeitslose Frauen messbar schlechtere Chancen, wieder eine Stelle zu finden. Es ist eine andere Sache, dies über eine KI zu zementieren, die Frauen tatsächlich eine schlechtere Vermittelbarkeit in den Score-Wert schreibt.

In den wenigsten Fällen haben die Sachbearbeiter:innen oder die Betroffenen eine Idee davon, wie eine KI zu ihrer Einschätzung kam. Wieder fehlt die Überprüfbarkeit oder die Möglichkeit, sich gegen mögliche Diskriminierung zu wehren. Selbst wenn der Computer nur eine Empfehlung ausgibt ist in der Regel zu beobachten, dass sich die Menschen in fast allen Fällen danach richten. Es wird schon seine Richtigkeit haben. Wer diesen Determinismus hinterfragt, wird häufig nicht ernst genommen oder als Querulant hingestellt.

# DIGITALER MACHTERHALT

1216 Je mehr Fakten von Technik geschaffen werden, gegen die wir  
1217 uns gar nicht mehr wehren können, desto größer wird auch  
1218 das damit einhergehende Gewaltpotential. Im Großen bedeu-  
1219 tet das, dass Machthaber und Firmenimperien die wahren  
1220 Entscheidungsträger sind und demokratische Prozesse unter-  
1221 laufen werden. Im Kleinen bedeutet das, dass Männer ihre  
1222 Freundin überwachen und diese technisch drangsaliieren, falls  
1223 sie sich entscheidet, sich zu trennen. Digitale Übergriffe sind  
1224 nicht dem männlichen Geschlecht vorbehalten. Auch mehr  
1225 und mehr Frauen bedienen sich dieses Machtinstruments.  
1226 Doch so lange Technik als Männersache gilt und KI die beste-  
1227 henden Machtverhältnisse zementiert, werden Frauen von  
1228 digitaler Gewalt weit mehr betroffen sein als Männer.

# DER PERPETUUM MOBIE DER SICH SELBST ERFÜLLENDEN PROPHEZEIUNGEN

1229 Jüngst musste die britische Regierung ihr peinliches Vorhaben  
1230 stoppen, die Noten der coronabedingt entfallenen Schul-Ab-  
1231 schlussprüfungen per KI errechnen zu lassen. Die Ungerech-  
1232 tigkeiten waren zu offensichtlich.

1233 Der Algorithmus urteilte aufgrund der Noten der  
1234 letzten drei Jahre. Noch so eine Idee, die gut gemeint war aber  
1235 statt der intendierten Innovation nur ein Festnageln in der Ver-  
1236 gangenheit zur Folge hat. Und ein weiteres Beispiel, das zeigt,  
1237 dass es vor allem die Kriterien sind, auf denen die Berechnun-  
1238 gen basieren, auf die es ankommt. Kriterien, über die wir auch  
1239 ein Wörtchen mitreden können, sollten. Wie nah hier Fiktion  
1240 und Wirklichkeit beieinander liegen, verrät uns der Berliner  
1241 Autor und Kleinkünstler Marc-Uwe Kling.

1242 In seinem Buch „Qualityland“ hört der Protagonist  
1243 Peter schon seit er denken kann Kuschelrock. Er weiß gar  
1244 nicht, ob er diese Musik überhaupt mag. Er hat sie mal auf-  
1245 gelegt, als er eine Romanze hatte und glaubte, dass die Musik  
1246 zum Anlass passt. Daraus leiten nun alle smarten Musikgeräte

(und davon gibt es viele in seiner fiktiven Welt) her, dass er Kuschelrock mag und spielen ihm nur noch diese Musik vor. Denn er hört es ja immer. Maschinen entscheiden, was er hören will, auf Basis von Entscheidungen, die andere Maschinen auf Basis von Entscheidungen anderer Maschinen getroffen haben. Das Perpetuum mobile der sich selbst erfüllenden Prophezeiungen. Einmal in eine Kiste sortiert, kommt man nie wieder heraus. Im Extremfall weiß man irgendwann selbst nicht mehr, ob man in die Kiste wirklich rein wollte, oder nur akzeptiert hat, dort einsortiert worden zu sein.

Deshalb sollte man sich auch nicht verführen lassen zu denken, dass der Fehler der Briten darin bestand, dass ihr Algorithmus einfach nicht gut genug war. Wäre er besser gewesen, hätte es noch schlimmer kommen können. Wäre er nicht gekippt worden, wären die verbliebenen Ungerechtigkeiten unter den Tisch gefallen. Raten Sie mal, welche Gruppen davon überdurchschnittlich stark betroffen gewesen wären. Die weiße, heteronormative Oberklasse hätte wohl eher nicht dazu gehört. Das Problem liegt eben nicht nur darin, dass die Algorithmen schlecht gemacht sind. Es setzt schon viel früher an. An dem Moment, in dem man irrtümlich dem Glauben erliegt, durch unreflektierte Reproduktion der (ungerechten) Vergangenheit die Realität der Gegenwart abbilden oder gar verbessern zu können.

Wenn wir uns den Algorithmen unterwerfen und ihnen die Gestaltung unserer Welt überlassen, wird Diskriminierung, insbesondere gegen all die, die nicht männlich-weiß sind, zementiert. Denn Frauen wird bis heute eingeredet, dass Computer für sie uninteressant sind. Infolgedessen gibt es tatsächlich sehr viele Frauen, die nicht sonderlich viel über Computer wissen. Frauen werden strukturell von technischen Fragen ferngehalten. Die wenigsten Frauen wissen, was mit ihren Geräten eigentlich wirklich passiert und haben somit auch keine Kontrolle. Dies macht es anderen umso leichter, Macht über sie zu ergreifen. Oder sie nur zu verunsichern. Je mehr KI unsere Gesellschaft durchdringt, desto weniger Vorstellung haben wir davon, was genau gerade eigentlich passiert. KI selbst ist hier nicht das Problem. Doch sie verstärkt durch ihre Intransparenz das ohnehin schon eklatante Machtungleichgewicht. Sie reproduziert und festigt die Diskriminierung, die es auch schon vorher gab.

# KI SINNVOLL GESTALTEN

Der Vormarsch der KI ist wohl kaum noch aufzuhalten. Umso dringender müssen wir nun die Frage stellen, wie sie gestaltet sein muss, um uns nicht zu schaden. Allerdings ist das ähnlich schwer zu definieren, wie der Begriff der KI selbst. Transparenz ist sicherlich einer der Schlüsselfaktoren. Es muss für die Menschen jederzeit erkennbar sein, wenn sie mit einer KI zu tun haben. Außerdem müssen ihre Entscheidungsprozesse nachvollziehbar sein. Menschen brauchen ein Veto-Recht gegenüber der Technik. Der große rote Knopf, mit dem der Mensch die Entscheidungshoheit an sich ziehen kann. Damit wir dazu aber überhaupt in der Lage sind, muss sich unsere gesamtgesellschaftliche Einstellung zur Technik ändern. Nur wenn wir den Trend umkehren, dass Technik immer weniger durch uns hinterfragt und bewertet wird, sind wir überhaupt in der Lage, KI so zu gestalten, dass sie möglichst wenige Machträume und Gewaltpotential öffnet. Dieses Umdenken muss sowohl bei Privatpersonen als auch bei Angestellten und in der Politik stattfinden. Eine KI sollte als potenziell problematisch betrachtet und durch ständige Untersuchungen ethisch-technisch-soziologisch begleitet werden. Dafür braucht es klare Regelungen und Vorgaben, die eine KI erfüllen muss, damit sie nicht als problematisch gilt. Datenschutz- und Sicherheitsfragen müssen von Anfang an mitgedacht und tief in der Technik verankert werden. Ethische Ziele müssen als Rahmenvorgabe in einer Weise definiert werden, dass eindeutig messbar wird, wann diese eingehalten werden. Datensätze, die zum Antrainieren von KI verwendet werden, müssen sehr genau auf mögliche strukturelle Fehler geprüft werden. Personen, die von einer KI-Entscheidung betroffen sind, müssen nicht nur das Recht haben, diese Entscheidung zu prüfen, sondern auch dagegen vorzugehen. Um dies alles verlässlich zu überwachen, braucht es Strukturen, die sicher stellen, dass die Regeln auch eingehalten werden und dabei mit der rasanten Entwicklung der Technik mithalten können. Jede KI braucht außerdem einen Verantwortlichen. Eine Person, die für entstehende Schäden haftbar ist. Das ist misslich, da dieses Risiko nur schwer abschätzbar ist. Doch anders wird es nicht gehen. Maschinen können sich nicht vor Gerichten verantworten und eine Verantwortungsdiffusion muss unbedingt vermieden werden. Wer eine KI herstellt muss sich nun mal verdammt sicher sein, dass diese keinen Schaden anrichtet. Wie man sich da sicher sein soll? Kann man nicht. Ein Grund mehr, das Risiko nicht auf die Gesellschaft umzulegen.

Falls Ihnen jetzt noch immer nicht klar sein sollte, wann Sie zuletzt mit einer KI zu tun hatten, sei Ihnen für die Ehrlichkeit gratuliert. Es ist in der derzeitigen Gestaltung von KI nicht vorgesehen, dass sie sich Ihnen kenntlich macht. Genau deshalb ist ihr Gewaltpotential auch so gefährlich.

Leena Simon ist graduierte (Netz-)Philosophin (MA) und IT-Beraterin und beschäftigt sich mit Digitaler Mündigkeit und Technikpaternalismus. Sie arbeitet u.a. für das Anti-Stalking-Projekt im Frieda Frauenzentrum in Berlin und für Digitalcourage e.V.  
Mehr zu ihrer Arbeit: <https://muendigkeit.digital>







3

# EFFIZIENTE UNGLEICHHEIT



VICTORIA  
GUIJARRO-SANTOS<sup>1</sup>

<sup>1</sup> Mein Dank gilt in besonderer Weise Eva Maria Bredler, ohne deren Zeit und Scharfsinn dieser Text nicht entstanden wäre.

In immer sensibleren Lebensbereichen soll Künstliche Intelligenz (KI)<sup>2</sup> eingesetzt werden oder wird es schon. Sogenannte algorithmische oder automatisierte Entscheidungssysteme werden verwendet, um Menschen bei der Entscheidungsfindung zu unterstützen oder zu ersetzen. Ihr Einsatzgebiet ist vielfältig und reicht von der Kreditvergabe über die Kriminalitätsvorhersage bis hin zur Gewährung staatlicher Sozialleistungen. Erhofft werden sich schnellere, günstigere, objektivere – kurz effizientere Entscheidungsprozesse. Doch die Frage ist: Effizient für wen?

Beispielsweise berechnet das österreichische Arbeitsmarktchancenassistenz-System (im Folgenden AMS-Algorithmus<sup>3</sup>), welche Antragsteller\*in mit welcher Wahrscheinlichkeit wieder in den österreichischen Arbeitsmarkt re-integrierbar ist. Dabei werden Personen, die die Datenpunkte „weiblich“ und „drittstaatsangehörig“ erfüllen, kategorisch schlechter bewertet als Personen, die die Datenpunkte „männlich“ und „österreichisch“ erfüllen (Holl et al. 2018: 11). Diese kategorisch schlechtere Bewertung führt dazu, dass der AMS-Algorithmus mit hoher Wahrscheinlichkeit ausländische Frauen der sogenannten Gruppe „N“ zuordnet (Holl et al. 2018: 6; 14<sup>4</sup>). Personen dieser Gruppe gelten als nur schwer in den Arbeitsmarkt re-integrierbar. Die Folgen dieser Einteilung sind ungewiss (Wimmer 2018b). Beim Chef des österreichischen Jobcenters (Arbeitsmarktservice) klingt an, dass sich teure Fördermaßnahmen bei dieser Personengruppe nicht lohnen würden (Kopf 2018; Dreas 2019<sup>5</sup>). Investitionen in arbeitslose, ausländische Frauen gelten damit als ökonomisch ineffizient.

Effizienz ist ein beliebtes Mantra in Zeiten der digitalen Transformation. Allerdings ist Effizienz keine den Grundrechten übergeordnete Maxime (Bertram 2020).

Aber werden mit dem AMS-Algorithmus überhaupt Grund- und Menschenrechte berührt? Immerhin argumentiert der Chef des Arbeitsmarktservice, dass der AMS-Algorithmus nur die Realität wiedergäbe (Kopf 2019). Frauen und Drittstaatsangehörige seien statistisch schlechter in den Arbeitsmarkt zu integrieren – das sei nun mal so. Und ist das Tolle an KI nicht auch, dass sie vom Menschen – und damit vom Staat – unabhängige und objektive Entscheidungen trifft? Ist der Staat dann überhaupt verantwortlich?

Es scheint, als würde auf Effizienzerwägungen eine Fassade algorithmischer Objektivität und statistischer Eindeutigkeit (Lücking 2020) gebaut, hinter der Verantwortung abgegeben werden kann und sich Ungleichheiten unbesehen verfestigen. Grund- und Menschenrechte stehen dieser Fassade aber nicht machtlos gegenüber. Dieser Beitrag legt – ausgehend vom deutschen Grundrechtsrechtsschutz – dar, warum.

2 Für technische Begriffe und Prozesse s. Lücking 2020; Lehr/Ohm 2017.

3 Zu den verschiedenen Namen s. Berner/Schüll 2020: 3.

4 Gruppe mit den niedrigsten Chancen N-Gruppe; Gruppe mit den besten Chancen H-Gruppe; Gruppe mit mittleren Chancen M-Gruppe.

5 „[Die] Abbildung der Arbeitsmarktchancen inkl. der Diskriminierungen ist ja gerade die Basis dafür, gezielt Unterstützungsmaßnahmen zu setzen, kann aber auch helfen, die Grenzen der Sinnhaftigkeit des Einsatzes besonders teurer arbeitsmarktpolitischer Maßnahmen aufzuzeigen.“

# 1. COMPUTER SAID SO

KI, algorithmische Entscheidungssysteme oder autonome Systeme werden oft synonym gesetzt. Häufig hört man auch von dem Blackbox-Algorithmus (z.B. Martini 2019), oder Sätze wie: „Der Algorithmus ist so komplex, dass noch nicht einmal die Programmiererin weiß, was er tut“. Dies suggeriert, dass ein Algorithmus ein undurchdringliches, selbständiges Wesen sei. Obwohl die bisher eingesetzten algorithmischen Entscheidungssysteme keineswegs so kompliziert sind, wie medial häufig dargestellt (vgl. AlgorithmWatch 2019: 7; der AMS-Algorithmus basiert beispielsweise auf nachvollziehbarer Statistik), vermittelt dieses Narrativ den Eindruck, dass der Staat seine Verantwortung an dieses als Software verkörperte Wesen abgeben könnte. „Sie denken, sie werden benachteiligt? Was soll ich sagen ... computer says no“. Little Britain droht zum Drehbuch exekutiver Staatsmacht zu werden (vgl. Little Britain 2012).

Die Verfassung und Menschenrechtsverträge geben allerdings ein anderes Drehbuch vor. In diesem ist Grundrechtsschutz notwendiges Korrelat staatlicher Entscheidungsverantwortung (BVerfG 2019: Rn. 42) – Staatsmacht gibt es nur zum Preis von Grundrechten. Wenn die Staatsgewalten (Legislative, Exekutive, Judikative) ein algorithmisches Entscheidungssystem einsetzen und dadurch dem Staat zurechenbar Grundrechte beeinträchtigen, kann sich die Staatsmacht vor dieser Verantwortung nicht hinter einem Algorithmus verstecken. Der Staat bleibt verantwortlich – unabhängig davon, ob eine benachteiligende Behandlung durch seine Beamt:innen oder durch ein von ihm eingesetztes algorithmisches System vermittelt wird (vgl. BVerfG 2020: Rn. 192; Gericht von Den Haag 2020).<sup>6</sup>

Problematischer als der Algorithmus ist vielmehr die Tatsache, dass algorithmische Entscheidungssysteme häufig (Rieder & Rudschies 2020: EASST) von privaten Softwarefirmen entwickelt werden. Der Staat hat deshalb nicht nur die Grundrechte der vom algorithmischen Entscheidungssystem betroffenen Personen zu achten, sondern auch die Grundrechte der Softwareunternehmen.<sup>7</sup> Dies kann sich auf Transparenzanforderungen und auf Ebene der Rechtfertigung<sup>8</sup> von Grundrechtseingriffen auswirken (Sommerer 2020: 227ff.), hindert die Zurechnung der Grundrechtsbeeinträchtigung zum Staat allerdings nicht.

Kurzum: der strengen menschen- und grundrechtlichen Verantwortung kann sich der Staat durch den Einsatz algorithmischer Systeme nicht entziehen (Gericht von Den Haag 2020: Rn. 6.42; Yeung 2019: 13; 60ff.).

<sup>6</sup> Das österreichische Jobcenter (Arbeitsmarktservice) ist ein öffentlich-rechtliches Unternehmen und deshalb der staatlichen Sphäre zugehörig, § 1 Arbeitsmarktservicegesetz, BGBl. Nr. 313/1994.

<sup>7</sup> Zum Beispiel den von Artikel 12, 14 Grundgesetz und Artikel 14 Europäische Grundrechtecharta umfassten Geschäftsgeheimnisschutz.

<sup>8</sup> Siehe Kapitel 4.

# 2. RELEVANTE GRUNDRECHTE

Der Staat bleibt also trotz des Einsatzes „autonomer KI“ an Menschen- und Grundrechte gebunden. Welche Grundrechte sind im Rahmen algorithmisch vermittelter Diskriminierungen relevant? Und wovor schützen sie uns? Im Folgenden werden die Grundrechte auf informationelle Selbstbestimmung und Datenschutz (2.1) sowie die Grundrechte auf Gleichbehandlung und Antidiskriminierung (2.2) näher beleuchtet und in Kapitel 3 betrachtet, wie dieser Grundrechtsschutz durch KI beeinträchtigt wird.

## 2.1 GRUNDRECHTE AUF INFORMATIONELLE SELBSTBESTIMMUNG UND DATENSCHUTZ

Artikel 2 Abs. 1 Grundgesetz lautet: „Jeder<sup>9</sup> hat das Recht auf die freie Entfaltung seiner Persönlichkeit“. Das hierin enthaltene Allgemeine Persönlichkeitsrecht schützt nicht vor allem, was die Persönlichkeitsentwicklung irgendwie beeinflussen kann (BVerfG 2017: Rn. 38). Denn Persönlichkeit entwickelt sich erst in Interaktion und Kommunikation mit anderen (Hildebrandt 2016: 80ff.; BVerfG 2017: Rn. 38). Der Schutz greift aber dann, „wenn die selbstbestimmte Entwicklung und Wahrung der Persönlichkeit *spezifisch* gefährdet ist“ (BVerfG 2017: Rn. 38). Externe Einflüsse müssen also eine gewisse Schwelle überschreiten. Dies ist der Fall, wenn umfassende digitale Persönlichkeitsbilder erstellt werden, ohne dass die Betroffene dessen Richtigkeit und Verwendung hinreichend kontrollieren kann. Diese Gefahr hat das Bundesverfassungsgericht – gerade auch mit Blick auf die Bedingungen automatischer Datenverarbeitung – bereits 1985 erkannt. Im sogenannten Volkszählungsurteil betonte das Bundesverfassungsgericht, dass grundsätzlich jede\*r selbst über die Preisgabe und Verwendung der eigenen Daten bestimmen können soll (BVerfG 1985: Rn. 153ff.). Damit trägt das Grundrecht auf informationelle Selbstbestimmung Gefährdungen der Persönlichkeit durch Datenerhebung und -verarbeitung Rechnung

(BVerfG 2018: Rn. 37). Auch der Europäische Gerichtshof für Menschenrechte konkretisierte das Recht auf Privatsphäre (Art. 8 Europäische Menschenrechtskonvention (EMRK)) schon früh zum Schutzschild gegen intransparente und umfassende Datenauswertung (EGMR 1984: Rn. 67; EGMR 2008: Rn. 99). Mit der Europäischen Grundrechtecharta wurde der Bedeutung des Datenschutzes für die individuelle Persönlichkeitsentwicklung und Freiheitsausübung, die in einer demokratischen Gesellschaft als notwendig erachtet werden (BVerfG 1985: Rn. 154; Hildebrandt 2006: 87<sup>10</sup>; Human Rights Council 2014), durch das Grundrecht auf Datenschutz in Artikel 7 Nachdruck verliehen.

9 Zum Gendern und Recht: Lischewski 2020.

10 „[...]within constitutional democracy, privacy is a public good in as far as it protects practices of identity building that are preconditional in a vigilant civil society.“

## 2.2 GRUNDRECHTE AUF GLEICH- BEHANDLUNG UND ANTI- DISKRIMINIERUNG

Gemäß Artikel 3 Abs. 1 Grundgesetz und Artikel 20 Grundrechtecharta, sind alle Menschen vor dem Gesetz gleich. Das bedeutet aber nicht, dass alle Menschen identisch behandelt werden sollen (Gerhard 1990). Tatsächlich differenziert die Staatsmacht ununterbrochen: wird ein Gesetz erlassen oder nicht? Greift die Polizei ein oder nicht? Anstelle einer identischen Behandlung tritt das Gebot, Ungleiches ungleich und Gleiches gleich zu behandeln (BVerfG 2003: Rn. 92). Weicht der Staat hiervon ab, muss dies sachlich begründet sein (BVerfG 2014: Rn. 121; EuGH 2010: Rn. 64). Das Grundrecht auf Gleichbehandlung kann somit als Grundrecht auf *sachlich begründete* (Un-)Gleichbehandlung verstanden werden.

In einem Fall wie dem des AMS-Algorithmus, wird aber nicht nach *irgendwelchen* Kriterien differenziert, sondern Frauen kategorisch ein schlechterer Ko-Effizient zugerechnet als Männern. Relevant sind deswegen die spezielleren Diskriminierungsverbote. Bevor diese näher ausgeführt werden, sei darauf hingewiesen, dass Diskriminierungsverbote im deutschen Grundrechtsschutz bisher ein „merkwürdiges Schat-



tendasein“ geführt haben (BVerfG 1983: 303; Baer & Markard 2018: Rn. 410). Das bedeutet, dass sie nur selten angewendet werden und deshalb nicht so scharf konturiert sind, wie zum Beispiel die Grundrechte auf informationelle Selbstbestimmung und Datenschutz. Licht und Kontur wird den Diskriminierungsverboten hauptsächlich durch die Arbeit feministischer Rechtswissenschaftler\*innen gegeben, auf deren Wissen die folgenden Ausführungen überwiegend beruhen.

Diskriminierungsverbote werden auch besondere Gleichheitssätze genannt (Boysen 2012: Rn. 117). Denn sie verbieten spezifische Ungleichbehandlungen von Personen aufgrund rechtlich benannter Kategorien (vgl. Mangold 2016/2020 i.V.: 5). Artikel 3 Abs. 3 Grundgesetz bestimmt beispielsweise, dass niemand wegen der dort benannten Kategorien wie Geschlecht oder Rasse<sup>11</sup> benachteiligt oder privilegiert werden soll. Artikel 21 Grundrechtecharta verbietet Diskriminierungen wegen der dort bezeichneten Kategorien. Es handelt sich hierbei um „spezifische Ungleichbehandlungen“, weil es nicht um das Verbot irgendwelcher Differenzierungen geht, sondern darum kategorial bestimmte Ungleichheitsverhältnisse, die Personen(gruppen) historisch marginalisieren, nicht weiter zu vertiefen (Baer & Markard 2018: Rn. 387; 407; Röhrner 2019: 171; 174). Es muss also geprüft werden, ob eine Person aufgrund eines Ungleichheitsverhältnisses (wie zum Beispiel Sexismus), das als Kategorie aufgezählt wird (Geschlecht) individuell benachteiligt wird.

Ist dies der Fall, trägt der Staat für diese „spezifische Ungleichbehandlung“ eine sehr viel höhere Rechtfertigungslast als für eine „einfache Ungleichbehandlung“ nach Artikel 3 Abs. 1 Grundgesetz (Baer & Markard 2018: Rn. 433ff.). Benachteiligungen wegen des Geschlechts beispielsweise, können nur gerechtfertigt werden, „soweit sie zur Lösung von Problemen, die ihrer Natur nach nur entweder bei Männern oder bei Frauen auftreten können, zwingend erforderlich sind“ (BVerfG 1992: 207). Dass es „natürliche Unterschiede“ zwischen Männern und Frauen gäbe, ist eine problematische Annahme des Bundesverfassungsgerichts (Adamietz 2006: 380; vgl. Butler 1990). Trotzdem verdeutlicht diese Aussage, dass die Rechtfertigungs-Latte sehr hoch liegt, wenn eine Benachteiligung geschlechtsbezogen ist.

Die Diskriminierungsverbote kennen keine Bagatellgrenze (Langenfeld 2020: Rn. 27). Beispielsweise unterfällt auch die Situation dem Schutz von Artikel 3 Abs. 3 Grundgesetz, in der ein automatisierter Seifenspender *weiße*, aber nicht Schwarze<sup>12</sup> Haut erkennt (Whites Only; Benjamin 2019: 67f.) – ein Fall sozusagen algorithmischer Alltagsdiskriminierung.

Die individuelle Bedeutung von Diskriminierungsverboten zum Schutz vor Würde- und Persönlichkeitsverletzungen (Baer & Markard, 2018: Rn. 407; Lehner 2013: 375ff.), sowie die gesellschaftliche Bedeutung von Antidiskriminierung für wirtschaftlichen und gesellschaftlichen Zusammenhalt (EU-Antirassismus-Richtlinie: Erwägungsgrund 9), wird durch Diskriminierungsverbote in sechs Menschenrechtsverträgen<sup>13</sup> und fast allen demokratischen Verfassungen weltweit ausgedrückt. Trotz dieses weitreichenden Konsenses, besteht in Rechtswissenschaft- und praxis wenig Einigkeit über die Details (Stanford Encyclopedia).



11 Dieser Begriff wird als soziale Konstruktion verstanden und als analytisches Konzept verwendet, näher Cengiz Barskanmaz 2011.

12 Im Anschluss an antirassistische, empowernde Politiken wird Schwarz hier großgeschrieben AK Forschungs-Handeln, 2016.

13 Z.B: Artikel 2 Allgemeine Erklärung der Menschenrechte; Artikel 26 Internationaler Pakt politischer und bürgerlicher Rechte; Artikel 14 EMRK.

### 3. WIE GREIFT KI EIN?

1549 Fassen wir kurz zusammen: Der Staat bleibt an Grund- und  
1550 Menschenrechte gebunden und wir werden vor invasiven  
1551 digitalen Persönlichkeitsprofilen, sachlich ungerechtfertig-  
1552 ten Ungleichbehandlungen und Diskriminierungen geschützt.  
1553 Soweit so klar. Aber was bedeutet das konkret bezogen auf  
1554 algorithmische Entscheidungssysteme? Wie genau wird der  
1555 nachgezeichnete grund- und menschenrechtliche Schutz  
1556 durch algorithmisch vermittelte Benachteiligungen beein-  
1557 trächtigt?

1558 Stellen wir uns vor, eine Frau und Mutter meldet sich  
1559 in Österreich arbeitslos. Danach erhält sie einen Bescheid, der  
1560 ihr erklärt, dass sie keine Förderung vom Arbeitsmarktservice  
1561 erhalten wird, weil ihr Integrationschancen-Wert (IC-Wert)  
1562 unter 25% liegt. Sie gehört damit zur Gruppe „N“. Der Grund  
1563 für den schlechten IC-Wert ist, dass der Datenpunkt „weiblich“  
1564 mit einem negativen Ko-Effizient von -0,14 und der Daten-  
1565 punkt „betreuungspflichtig“ mit -0,15 bewertet wird.

# 3.1 GRUNDRECHTE AUF INFORMATIONELLE SELBST- BESTIMMUNG UND DATEN- SCHUTZ

Allein dadurch, dass Daten erfasst und abgeglichen werden, wird in die Grundrechte auf informationelle Selbstbestimmung und Datenschutz eingegriffen (BVerfG 2018: Rn. 45ff). Bei einem algorithmischen Entscheidungssystem kommt hinzu, dass auf Basis dieser Daten ein digitales Persönlichkeitsprofil konstruiert wird. Das spezifische Problem hieran ist weniger, dass eine Person zur „gläsernen Bürgerin“ würde, sondern eher, dass ihr (und auch allen anderen Frauen und Müttern) generalisierende Annahmen „untergeschoben“ werden. Auf Basis dieser Generalisierungen wird ein Profil erstellt (sog. Profiling), obwohl die Antragstellerin womöglich ganz anders ist bzw. entgegen der statistischen Generalisierung doch gute Chancen auf dem Arbeitsmarkt hätte.

Tatsächlich sind falsche Fremdbilder keine Ausnahme, sondern vielmehr algorithmischen Entscheidungssystemen inhärent. Denn statistische Vorhersagen können sich annähern, prophezeien aber keine Wahrheiten (Lücking 2020). Irgendwer ist immer „falsch negativ“, das heißt entgegen der statistischen Erwartung zum Beispiel doch re-integrationsfähig (vgl. Zweig/Krafft 2018). Im Fall des AMS-Algorithmus liegt die Fehlerquote beispielsweise bei 15%. Das heißt, dass Medienberichten zufolge jährlich 50.000 Personen falsch bewertet werden (Wimmer 2018a).

Die Betroffenen haben keinen Einfluss auf diesen Prozess. Ihr digitales Persönlichkeitsprofil wird unabhängig von ihnen konstruiert. Ihnen wird jeglicher Raum für die effektive Gegenstellung ihres Selbstbildes genommen. Stattdessen wird den Betroffenen ein stereotypes Fremdbild übergestülpt, dessen Richtigkeit und Verwendung sie nicht beeinflussen können (Britz 2008: 184ff; 193).

Generalisierungen sind alltäglich. Wir haben kein Recht darauf, nur so gesehen zu werden, wie wir uns das wünschen (s. oben). Aber wir haben ein Recht darauf, an dem kommunikativen Prozess der Persönlichkeitskonstruktion derart beteiligt zu werden, sodass wir auf wirkmächtige Fremdbilder real einwirken können (Britz 2008: 193f.; 207f.; diess. 2019: 675ff.). Bei automatisierten Entscheidungssystemen wird einem diese Möglichkeit genommen: Weiblich zu sein ist

1604 negativ, Mutter zu sein ist negativ – Ende.<sup>14</sup> Auf diese Weise  
1605 wird der Raum für Persönlichkeitsentfaltung durch Datenver-  
1606 arbeitung verschlossen und die Grundrechte auf informatio-  
1607 nelle Selbstbestimmung und Datenschutz beeinträchtigt.

14 Beim AMS-Algorithmus soll noch eine Mitarbeitende zwischen Software-Ergebnis und letztgültiger Entscheidung geschaltet werden (Kopf 2018). Trotzdem ist ein algorithmisch erzielttes Ergebnis so eindrucklich, dass psychologisch davon ausgegangen wird, dass Menschen nicht davon abweichen, sog. automation bias (Sommerer 2020: 71ff).

## 3.2 GRUNDRECHTE AUF ANTI- DISKRIMINIERUNG

1608 Die Antragstellerin erleidet im Beispielfall einen Nachteil, da  
1609 sie keine Leistungen vom Arbeitsmarktservice erhält. In Wirk-  
1610 lichkeit ist gar nicht so eindeutig, was aus der Kategorisierung  
1611 in Gruppe „N“ folgt. Klar ist, dass die meiste Förderung in  
1612 Gruppe „M“ investiert werden soll – der Personengruppe, die  
1613 statistisch mittelmäßige Chancen hat (Wimmer 2018b). Wahr-  
1614 scheinlich ist, dass Personen der Gruppe „N“ einer anderen  
1615 Einrichtung zugeordnet werden und damit aus dem Arbeits-  
1616 marktservice-System herausgenommen werden sollen (Kopf  
1617 2018; Lopez 2019: 298; Berner & Mass 2020: 4). Die dort  
1618 angebotenen Maßnahmen seien nicht qualifikationsbezogen,  
1619 sondern zielten auf soziale Stabilisierung ab (Berner & Schüll  
1620 2019; Weber et al. 2019). Letztendlich kann aber kaum mehr  
1621 als geraten werden, welcher Nachteil genau aus der Zuwei-  
1622 sung zu Gruppe „N“ folgt. Der Einfachheit halber denken wir  
1623 deswegen den Beispielfall weiter.

1624 Die Antragstellerin erleidet nicht nur einen Nachteil,  
1625 sondern wird auch im Vergleich zu Männern benachteiligt.  
1626 „Männlich-sein“ wirkt sich auf die Bewertung durch den AMS-  
1627 Algorithmus weder positiv noch negativ aus. Das vergleichs-  
1628 weise höhere Risiko der Antragstellerin, in Gruppe „N“ einge-  
1629 teilt zu werden, wirkt sich in der Verweigerung von Sozial-  
1630 leistungen aus. Diese Benachteiligung müsste in einem der  
1631 kategorial gefassten Ungleichheitsverhältnissen gründen. Der  
1632 AMS-Algorithmus weist *allen* Müttern einen negativen Ko-Ef-  
1633 fizienten zu. Das ist kein Zufall. Der AMS-Algorithmus wurde  
1634 anhand vergangener Fälle des Arbeitsmarktservice program-  
1635 miert. Der Algorithmus hat auf diese Weise den Frauen am  
1636 österreichischen Arbeitsmarkt entgegengebrachten Sexis-  
1637 mus erlernt und reproduziert ihn nun (Lopez 2019: 294; 302).  
1638 Auch Kriterien wie „Betreuungspflichten“ scheinen auf den

ersten Blick zwar ein formal neutrales Kriterium zu sein. Immerhin können Frauen wie Männer für die Betreuung von Kindern oder anderen Angehörigen verantwortlich sein. Tatsächlich kümmern sich aber vor allem Frauen um die Kinderbetreuung und Pflege. Sorgearbeit ist weiterhin weiblich (Dreas 2019: 223). Das Kriterium „Betreuungspflichten“ scheint also neutral, betrifft jedoch faktisch Frauen. Entsprechend wirken sich beim AMS-Algorithmus „Betreuungspflichten“ nur für Frauen negativ aus (Holl et al. 2018: 13).

Ein anderes, maskierteres Beispiel ist ein in Chicago eingesetzter Algorithmus des sogenannten Predictive Policing. Hier ermittelt der Algorithmus auf Basis von Daten aller in den letzten fünf Jahren festgenommen Personen, ob eine Person Täter oder Opfer von Schusswaffengewalt wird (Sommerer 2020: 80ff). Abgesehen davon, dass festgenommen zu werden, nicht bedeutet auch tatsächlich strafrechtlich gehandelt zu haben, sind Daten über Festnahmen auch deswegen problematisch, weil sie gegebenenfalls stattgefundenes polizeiliches Racial Profiling<sup>15</sup> wiedergeben (Sommerer 2020: 110; 112). Formal-neutrale Festnahmedaten können dann zum Ergebnis führen, dass der Algorithmus im Ergebnis rassifizierte Männer häufiger einer Straftat verdächtigt, als weiße Männer oder Frauen.

Das Antidiskriminierungsrecht nimmt Diskriminierungen ihre Maske mit der Rechtsfigur der mittelbaren Diskriminierung ab (BVerfG 1997: Rn. 36; Röhner 2019: 213ff): Mit der Verwendung formal-neutraler Kriterien, können Diskriminierungsverbote nicht umgangen werden, wenn sie sich effektiv für eine kategorial geschützte Personengruppe benachteiligend auswirken. Die Umgehung muss gar nicht beabsichtigt sein. Es muss kein böser Wille nachgewiesen werden, um den Schutzbereich der Diskriminierungsverbote zu berühren (Mangold 2016 & 2020 i.V.: 155).

Es könnte nun manch eine argumentieren, dass ja gar nicht alle Frauen in Gruppe „N“ fallen und auch Männer dort eingruppiert werden würden. Hieraus könnte der Schluss gezogen werden, dass die arbeitslose Mutter doch nicht *wegen des Geschlechts* benachteiligt wird. Dies wäre ein Fehlschluss. Es muss nämlich genau betrachtet werden, warum auch Männer in Gruppe „N“ kategorisiert worden sind und andere Frauen nicht. Anhand des AMS-Algorithmus lässt sich das plastisch durchspielen: Ein Mann könnte zum Beispiel beeinträchtigt sein (negativer Ko-Effizient von -0,67) und eine Frau könnte keine Betreuungspflichten haben und unter 30 Jahre alt sein (keine Auswirkungen; Holl et al. 2018: 11). Der AMS-Algorithmus veranschaulicht auf diese Weise, dass Ungleichheitsverhältnisse Personen gesellschaftlich unterschiedlich positionieren. Ein Mann kann gegenüber einer Frau privilegiert sein. Ist er beeinträchtigt, kann er aber aufgrund seiner Beeinträchtigung – und nicht seinen Geschlechtes – im Vergleich zu einer able-bodied – Frau benachteiligt werden. Andersherum kann eine Frau gegenüber einer anderen Frau privilegiert sein, weil sie nicht mit Sorgearbeit belastet und besonders jung ist. Sexismus, Able-ismus und Age-ismus, können Männer und Frauen unterschiedlich positionieren.

Diese Komplexität, nach der sich nicht immer alle Ungleichheitsdimensionen in jeder Situation gleich auswirken, sondern genau geprüft werden muss, welche Machtverhältnisse sich wie miteinander verschränken, hat Kimberlé Crenshaw mit dem Begriff der Intersektionalität geprägt (Crenshaw 1989: 151f). Der AMS-Algorithmus fungiert entsprechend als intersektionaler Rechenschieber. Die Rechtspraxis erkennt intersektionale Diskriminierung bis heute nicht immer (EuGH

1702 2016; BVerfG 2003a; siehe aber EuGH 2011; BVerfG 2015).  
1703 Rechtswissenschaftler\*innen feilen jedoch an rechtlichen Lin-  
1704 sen, um Intersektionalität auch für Richter\*innen sichtbar zu  
1705 machen (vgl. Atrey 2019; Weinberg 2020).  
1706 In dem Beispielfall wird die arbeitslose Mutter somit  
1707 aufgrund der Ungleichheitsdimension Geschlecht benachtei-  
1708 ligt. Damit wird in den Schutzgehalt der Diskriminierungs-  
1709 verbote eingegriffen.

## 4. MEHR ALS EFFIZIENZ

1710 Der AMS-Algorithmus beeinträchtigt die Grundrechte auf  
1711 informationelle Selbstbestimmung und Datenschutz, sowie  
1712 jene auf Antidiskriminierung. Dies kann auch für andere durch  
1713 Algorithmen vermittelte Benachteiligungen gelten. Für sich  
1714 genommen folgt aus diesen Eingriffen aber noch nicht, dass  
1715 algorithmische Entscheidungssysteme in allen Behörden sofort  
1716 deinstalliert werden müssten. Grund- und Menschenrechte  
1717 schützen bestimmte Verhaltensweisen. Beeinträchtigt der  
1718 Staat ein geschütztes Verhalten, dann muss der Staat dies gut  
1719 begründen, das heißt rechtfertigen. Grund- und Menschen-  
1720 rechte verbieten Eingriffe somit nicht absolut, sondern schir-  
1721 men uns vor einem wild um sich greifenden Staat ab, indem  
1722 sie eine Rechtfertigung für die erfolgten Beeinträchtigungen  
1723 fordern. Grund- und Menschenrechtseingriffe können zudem  
1724 nicht irgendwie gerechtfertigt werden. Vielmehr muss die  
1725 Rechtfertigung bestimmten Anforderungen genügen.

### 4.1 KEIN EINGRIFF OHNE GESETZ

1726 Zum einen darf in Grundrechte grundsätzlich nur auf Basis  
1727 einer Rechtsgrundlage eingegriffen werden (Vorbehalt des  
1728 Gesetzes). Gerade bei Eingriffen in die Grundrechte auf infor-  
1729 mationelle Selbstbestimmung und Datenschutz, kommt der  
1730 Rechtsgrundlage eine besondere Bedeutung zu. Denn sie  
1731 muss absichern, dass der Raum für Persönlichkeitsentfaltung

1732 trotz Datenverarbeitung wieder geöffnet wird. Die Rechts-  
1733 grundlage muss deswegen klar und detailliert über den  
1734 Umfang und die Anwendung der Datenverarbeitung informie-  
1735 ren und bestimmte Mindestgarantien zusichern (EGMR 2008:  
1736 Rn. 99; Gericht von Den Haag: Rn. 6.66ff). Beim Profiling müs-  
1737 sen den Betroffenen darüberhinausgehende Rechte an die  
1738 Hand gegeben werden, wie besondere Informationsrechte  
1739 (Artikel 14 Abs. 2 lit. g DSGVO) oder ein Anfechtungsrecht, mit  
1740 dem falsche digitale Persönlichkeitsbilder angefochten wer-  
1741 den können (Artikel 22 Abs. 3 DSGVO). Auf diese Weise soll  
1742 die betroffene Person wieder ins Spiel der Fremd- und Selbst-  
1743 bilder gebracht werden.

1744 Tatsächlich genügt die derzeitige Rechtsgrundlage  
1745 für den Einsatz des AMS-Algorithmus diesen (und anderen)  
1746 datenschutzrechtlichen Anforderungen nicht. Deswegen  
1747 wurde der Einsatz des AMS-Algorithmus von der österreichi-  
1748 schen Datenschutzbehörde vorerst gestoppt (ORF 2020).

## 4.2 ZU WELCHEM ZWECK?

1749 Doch auch nach Schaffung einer solchen Rechtsgrundlage,  
1750 stellt sich die Frage, ob der Einsatz des AMS-Algorithmus  
1751 gerechtfertigt werden kann.

1752 An dieser Stelle erweist sich Effizienz als nicht sehr  
1753 stabiles Fundament. Effizienz ist ein sehr ungenaues Argu-  
1754 ment. Sie beschreibt nämlich lediglich eine Zweck-Mittel-Re-  
1755 lation. Für die Rechtfertigung ist aber wichtig, *welcher* Zweck  
1756 genau mit dem erfolgten Grundrechtseingriff verfolgt wird.  
1757 Nur wenn Zweck und Eingriff in einem angemessenen Ver-  
1758 hältnis stehen, kann der Eingriff gerechtfertigt werden (Ver-  
1759 hältnismäßigkeitsprinzip). Dafür braucht es auch Informatio-  
1760 nen. Verweist der Staat lediglich auf den Algorithmus, genügt  
1761 dies nicht (Gericht von Den Haag 2020: Rn. 6.80ff; BVerfG  
1762 2020: BND<sup>16</sup>).

1763 Ein algorithmisches Entscheidungssystem könnte  
1764 möglicherweise verwendet werden, um Diskriminierungen zu  
1765 vermeiden. Aber ist ein algorithmisches Entscheidungssystem  
1766 dafür wirklich geeignet? Einzelne Personen wie Beamt\*innen  
1767 können entgegen diskriminierender Strukturen handeln. Ein  
1768 Algorithmus trifft hingegen massenweise die gleiche (diskri-  
1769 minierende) Entscheidung (O'Neill 2015).

1770 Nehmen wir mal an, der AMS-Algorithmus würde  
1771 eine gewinnbringende Verteilung knapper, staatlicher Res-  
1772ourcen bezwecken. Dies könnte durch die Segmentierung  
1773 arbeitsloser Personen wahrscheinlich auch erreicht werden.  
1774 Aber sind Sexismus, Rassismus, Able-ismus und Age-ismus  
1775 hierfür legitime Unterscheidungsgründe? Anstatt nach den  
1776 individuellen Fähigkeiten zu fragen, werden alle Frauen und  
1777 Mütter als eher nicht re-integrationsfähig bewertet. Inwiefern  
1778 hilft es Frauen, psychosoziale Beratung zu bekommen, wenn  
1779 der Grund für diese Prognose nicht in ihrer Person, sondern  
1780 auf den ihnen entgegengebrachten Sexismus beruht?



Würden Personen auf Basis des erkannten Sexismus segmentiert, um zum Beispiel Frauen spezielle Förderprogramme zu Teil werden zu lassen, könnte die Segmentierung gerechtfertigt werden (vgl. Gleichbehandlungsanwaltschaft 2019: 3). Denn dann würde versucht, ein Ungleichheitsverhältnis auszugleichen. Das Grundgesetz sieht gerade vor, dass der Staat die Gleichberechtigung von Frauen und Männern durch die Beseitigung von Nachteilen auch tatsächlich fördert (Artikel 3 Abs. 2 S. 2 Grundgesetz). Dies gelingt allerdings nicht, wenn – wie mit dem AMS-Algorithmus – ein diskriminierender status quo eingefroren und als Begründung für weitere Benachteiligungen genutzt wird (Fröhlich & Spiecker 2018).

## 5. FAZIT

Mithilfe von Grund- und Menschenrechten, kann hinter die Fassade algorithmischer Entscheidungssysteme geblickt werden. Dahinter halten sie uns einen Spiegel vor, in dem Ungleichheitsverhältnisse abgebildet werden. Insoweit hat der Chef des Arbeitsmarktservice also Recht: Die Realität „ist so“ – zumindest zurzeit.<sup>17</sup> Frauen haben es auf dem österreichischen Arbeitsmarkt nicht leicht. Mit dem Einsatz des AMS-Algorithmus wird das auch so bleiben. Allerdings sollten marginalisierte Personen(gruppen) durch den staatlichen Einsatz von KI nicht noch weiter an den gesellschaftlichen Rand gedrängt werden. Es muss vielmehr stets das grund- und menschenrechtlich gebotene Maß an Einzelfallgerechtigkeit aufrechterhalten bleiben (Britz 2008: 212ff).

<sup>17</sup> Zu den Grenzen der „realen Abbildung“ durch den AMS-Algorithmus s. Lopez (2019).

Victoria Guijarro-Santos ist wissenschaftliche Mitarbeiterin an der Universität Münster (Lehrstuhl Professorin Markard). Sie forscht an der Schnittstelle digitaler Technologien, Grund- und Menschenrechten sowie feministischer Rechtstheorie.

# LITERATUR

- Adamietz, Laura (2006)  
„Transgender ante portas?“. *Kritische Justiz* 2006 (39)4, 368–380.
- AK ForschungsHandeln (2016)  
„InterdepenDenken dialogisch verhandeln – Eine Einleitung“, in: AK ForschungsHandeln (Hg.), *InterdepenDenken! Wie Positionierung und Intersektionalität forschend gestalten?*, w\_orten&meer: Stralsund, 8–28.
- AlgorithmWatch (2019)  
„Atlas der Automatisierung – automatisierte Entscheidungen und Teilhabe in Deutschland“, online unter: [https://atlas.algorithmwatch.org/wp-content/uploads/2019/07/Atlas\\_der\\_Automatisierung\\_von\\_AlgorithmWatch.pdf](https://atlas.algorithmwatch.org/wp-content/uploads/2019/07/Atlas_der_Automatisierung_von_AlgorithmWatch.pdf) [25.09.2020].
- Atrey, Shreya (2019)  
*Intersectional Discrimination*, Oxford: Oxford University Press.
- Baer, Susanne & Markard, Nora (2018)  
„Art. 3 Abs. 3“, in: Hermann von Mangoldt et al. (Hg.), *Grundgesetz Kommentar*, München: C.H. Beck.
- Barskanmaz, Cengiz (2011)  
„Rasse – Unwort des Antidiskriminierungsrechts?“. *Kritische Justiz* (44)4, 382–389.
- Benjamin, Ruha (2019)  
*Race After Technology*, Medford: Polity Press.
- Berner, Heiko/ Schüll, Elmar (2020)  
„Bildung nach Maß. Die Auswirkungen des AMS-Algorithmus auf Chancengerechtigkeit, Bildungszugang und Weiterbildungsförderung“. *Magazin erwachsenbildung.at* 2020 (40), 1–12.
- Bertram, Alice (2020)  
„Recht auf Ineffizienz – Eine Untersuchung des Verhältnisses von Grundrechten zur Effizienz“, in: 60. Assistententagung Öffentliches Recht (Hg.), *Der digitalisierte Staat*, Baden-Baden: Nomos Verlagsgesellschaft, 87–102.
- Boysen, Sigrid (2012)  
„Artikel 3“, in: von Münch, Ingo & Kunig, Philip (Hg.), *Grundgesetz-Kommentar*, München: C.H. Beck.
- Britz, Gabriele (2008)  
*Einzelfallgerechtigkeit versus Generalisierungen*, Tübingen: Mohr Siebeck.
- Britz, Gabriele (2019)  
„Freie Entfaltung der Persönlichkeit (Artikel 2 I 1 GG) – Verfassungsversprechen zwischen Naivität und Hybris?“. *Neue Zeitschrift für Verwaltungsrecht* 38 (10), 672–677.
- Butler, Judith (1990/2007)  
*Gender trouble*, Oxon: Routledge.
- BVerfG 1983: BVerfGE 63, 266 (298 ff.) – Zulassung zur Anwaltschaft.
- BVerfG 1985: BVerfGE 65, 1 (43, Rn. 154) 1985 Volkszählungsurteil
- BVerfG 1997: BVerfGE 97, 35 – Hamburger Ruhegeldgesetz.
- BVerfG 2003, BVerfGE 110, 141 – Kampfhunde.
- BVerfG 2003a: BVerfGE 108, 282 – Kopftuch I.
- BVerfG 2014, BVerfGE 138, 136 – Erbschaftssteuer.
- BVerfG 2015, BVerfGE 138, 296 – Kopftuch II.
- BVerfG 2017, BVerfGE 147, 1 – Dritte Option.
- BVerfG 2018: BVerfGE 150, 244 (309) – Kfz-Kennzeichenkontrollen 2
- BVerfG 2019: BVerfG, Beschluss des Ersten Senats vom 06. November 2019, 1 BvR 16/13, Recht auf Vergessen I.
- BVerfG 2020: BVerfG, Urteil des Ersten Senats vom 19. Mai 2020, 1 BvR 2835/17 – Ausland-ausland-Fernmeldeaufklärung.
- Crenshaw, Kimberlé (1989)  
„Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine“. *The University of Chicago Legal Forum* (1989)1, 139–167.

- \_\_\_\_ Dreas, Susanne A. (2019)  
 „Zum Verhältnis von Care und Gender oder: Warum ist Sorgearbeit weiblich?“, in: Kolhoff, Ludger (Hg.), *Aktuelle Diskurse in der Sozialwirtschaft II*, Wiesbaden: Springer, Wiesbaden, 223–239.
- \_\_\_\_ EGMR 1984, § 64 Series A no. 82, Malone vs. the United Kingdom, 2.08.1984.
- \_\_\_\_ EGMR 2008, Rs. 30562 und 30566/04, Marper vs. the United Kingdom, 4.12.2008.
- \_\_\_\_ EuGH 2010, ECLI:EU:C:2010:534 – Chatzi.
- \_\_\_\_ EuGH 2011, ECLI:EU:C:2011:675 – Brachner.
- \_\_\_\_ EuGH 2016, ECLI:EU:C:2016:897 – Parris.
- \_\_\_\_ Fröhlich, Wiebke & Spiecker gen. Döhmman, Indra (2018)  
 „Können Algorithmen diskriminieren?“, in: *Verfassungsblog*, 26.12.2018, online unter: <https://verfassungsblog.de/koennen-algorithmen-diskriminieren/> [25.09.2020].
- \_\_\_\_ Gericht von Den Haag, ECLI:NL:RBDHA:2020:1878 – SyRI.
- \_\_\_\_ Gerhard, Ute (1990)  
*Gleichheit ohne Angleichung*, München: C.H. Beck.
- \_\_\_\_ Human Rights Council  
 A/HRC/28/39, 19.12.2014, Summary of the Human Rights Council panel discussion on the right to privacy in the digital age.
- \_\_\_\_ Hildebrandt, Mireille (2016)  
*Smart Technologies and The Ends of Law*, Celandham: Edward Elgar.
- \_\_\_\_ Holl, Jürgen; Kernbeiß, Günther; Wagner-Pinter, Michael (2018)  
*Das AMS-Arbeitsmarkt-chancen-Modell, Dokumentation zur Methode*, Wien: Synthesis Forschung.
- \_\_\_\_ Kopf, Johannes (2018)  
 „AMS-Vorstand Kopf: ‚Was die EDV gar nicht abbilden kann, ist die Motivation‘“, in: *Der Standard*, 10.10.2018, online unter <https://www.derstandard.at/story/2000089096795/ams-vorstand-kopf-menschliche-komponente-wird-entscheidend-bleiben> [25.09.2020].
- \_\_\_\_ Kopf, Johannes (2019)  
 „Der Beipackzettel zum AMS-Algorithmus“ in: Futurezone, 09.10.2019, online unter: <https://futurezone.at/meinung/der-beipackzettel-zum-ams-algorithmus/400641347> [25.09.2020].
- \_\_\_\_ Langenfeld, Christine (2020)  
 „Art. 3 Abs. 3“, in: Maunz, Theodor/Dürig, Günter, *Grundgesetz-Kommentar*, München: C.H. Beck.
- \_\_\_\_ Lehr, David & Ohm, Paul (2017)  
 „Playing with the Data: What Legal Scholars Should Learn about Machine Learning“. *U.C. Davis Law Review* 51(2), 653–71.
- \_\_\_\_ Leidinger, Andreas (2019)  
 „Drei Perspektiven auf Racial Profiling: Konservativ, liberal, kritisch“. *Kritische Justiz* 51(4), 450–463
- \_\_\_\_ Little Britain  
 „Computer sagt Nein“, in: *Youtube*, 08.03.2012, online abrufbar unter: [www.youtube.com/watch?v=ajdafAwJsho](http://www.youtube.com/watch?v=ajdafAwJsho) [25.09.2020].
- \_\_\_\_ Lischewski, Isabel (2020)  
 „Im Namen des Gesetzes – Misgendering durch die Verfassung?“, in: *Verfassungsblog*, 02.07.2020, online unter: <https://verfassungsblog.de/im-namen-des-gesetzes/> [25.09.2020].
- \_\_\_\_ Lopez, Paola (2019)  
 „Reinforcing Intersectional Inequality via the AMS Algorithm in Austria“. *Proceedings of the STS Conference Graz 2019*, 289–309.
- \_\_\_\_ Lücking, Phillip (2020)  
 „Automatisierte Ungleichheit“, in: netzforma\* e.V. (Hg.), *Wenn KI, dann feministisch – Impulse aus Wissenschaft und Aktivismus*. Berlin, 65–76.
- \_\_\_\_ Mangold, Anna Katharina (2016/2020)  
*Demokratische Inklusion durch Recht*, Habilitationsschrift 2016/2020 in Vorbereitung, Tübingen: Mohr Siebeck.
- \_\_\_\_ Martini, Mario (2019)  
*Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz*, Berlin: Springer.

- \_\_\_\_\_ Richtlinie 2000/43/EG des Rates vom 29.06.2000, ABL 180 (22), 19.07.2000.
- \_\_\_\_\_ Rieder, Gernot & Rudschies Catharina (2020)  
 „Notes on the Political Economy of Welfare AI“, Vortrag, 21.08.2020,  
 bei: EASST 2020.
- \_\_\_\_\_ Röhner, Cara (2019)  
*Ungleichheit und Verfassung*, Weilerwist: Velbrück Wissenschaft.
- \_\_\_\_\_ O'Neill, Cathryn (2016)  
*Weapons of Math Destruction*, New York: Broadway Books.
- \_\_\_\_\_ ORF (2020)  
 „Zukunft für AMS-Algorithmus ungewiss“, 21.08.2020, online unter:  
<https://orf.at/stories/3178348/> [25.09.2020].
- \_\_\_\_\_ Sommerer, Lucia (2020)  
*Personenbezogenes Predictive Policing*, Baden-Baden: Nomos Verlagsgesellschaft.
- \_\_\_\_\_ Stanford Encyclopedia of Philosophy (2011/2020)  
 „Discrimination“, online unter: <https://plato.stanford.edu/entries/discrimination/#OrgInsStrDis> [25.09.2020].
- \_\_\_\_\_ Weber, Friederike; Hager, Isa; Krüse, Tobias & Reidl, Christine (2019)  
*Evaluierung des Betreuungsformates für Personen mit multiplen Vermittlungserfordernissen (BBEN). Endbericht der prospect Research & Solution Unternehmensberatung*, Wien: Arbeitsmarktservice.
- \_\_\_\_\_ Weinberg, Nils (2020)  
 „Ansätze zur Dogmatik der intersektionalen Benachteiligung“, in:  
*Europäische Zeitschrift für Arbeitsrecht* (13), 60–77.
- \_\_\_\_\_ Whites Only – Part Two  
 in: *YouTube*, online unter: <https://www.youtube.com/watch?v=gKUTDAvmNtA>, [25.09.2020].
- \_\_\_\_\_ Wimmer, Barbara, (2018a)  
 „AMS-Chef: ‚Mitarbeiter schätzen Jobchancen pessimistischer ein als der Algorithmus‘“, in: *Futurezone*, 12.10.2018, online unter:  
<https://futurezone.at/netzpolitik/ams-chef-mitarbeiter-schaetzen-jobchancen-pessimistischer-ein-als-der-algorithmus/400143839> [25.09.2020].
- \_\_\_\_\_ Wimmer, Barbara (2018b)  
 „Computer sagt nein: Algorithmus gibt Frauen weniger chancen beim AMS“ in: *Futurezone*, 06.12.2018b, online unter: <https://futurezone.at/netzpolitik/computer-sagt-nein-algorithmus-gibt-frauen-weniger-chancen-beim-ams/400345297> [25.09.2020].
- \_\_\_\_\_ Yeung, Karen (2019)  
 „Responsibility and AI“, in: Council of Europe (Hg.), *Council of Europe Study DGI(2019)05*.
- \_\_\_\_\_ Zweig, Katharina A./Krafft, Tobias D. (2018)  
 „Fairness und Qualität algorithmischer Entscheidungen“, in: Mohabbat Kar, Resa/ Thapa, Basanta/Parycek, Peter (Hg.), *(Un)Berechenbar?*, Berlin: Kompetenzzentrum Öffentliche IT, 204–227.







AUTOMATISIERTE  
UNGLEICHHEIT

WIE  
ALGORITHMISCHE  
ENTSCHEIDUNGS-  
SYSTEME  
GESELLSCHAFTLICHE  
MACHT-  
VERHÄLTNISSE  
REPRODUZIEREN



PHILIP  
LÜCKING

In den medialen Debatten rund um die Themen „Künstliche Intelligenz“ und Algorithmen wird auf der einen Seite häufig von einer am Horizont erscheinenden drohenden Gefahr geschrieben und vor der ungeheuren, irgendwie unpersönlichen Macht dieser neuen Hochtechnologien gewarnt. Auf der anderen Seite werden im Kontext von Medienkampagnen wie Industrie 4.0 allerhand positive Seiten dieser Technologien betont. Das reicht von angeblich deutlich erhöhten Produktionskapazitäten über bessere Chancen im internationalen Konkurrenzettbewerb bis hin zur generellen Lösung aller möglichen (gesellschaftlichen) Probleme.<sup>1</sup> Dabei wird allzu oft übersehen, dass jenseits von abstrakter Roboterherrschaft oder blühenden Algorithmenlandschaften einzelne Technologien (teil)automatisierter, algorithmischer Entscheidungssysteme teilweise schon lange im Einsatz sind. Gerade solche mit existenzieller Dimension (also nicht, welche passenden Serien dir bei Netflix angeboten oder welche individualisierten Inhalte dir bei Instagram angezeigt werden) betreffen vor allem bereits marginalisierte Personengruppen – Frauen\*, People of Color (PoCs), Arme, Geflüchtete und weitere. Während im Zusammenhang mit beispielsweise Gesichtserkennungssystemen (Köver 2019) oder rassistischen Suchergebnissen bei Google (Spiegel 2018) heute viele wichtige Debatten über Bias in Algorithmen geführt werden, wird oft übersehen, dass jenseits von Produkten im „consumer-Bereich“ solche Technologien, getrieben von einer fortschreitenden Effizienzlogik und vor dem Hintergrund eines gesellschaftlichen Digitalisierungstrends in vielen (staatlichen) Sektoren bereits eingesetzt werden. Ein paar Beispiele:

2018 wurde das Buch *Automating Inequality* von Virginia Eubanks in den USA veröffentlicht. In diesem Buch beschreibt die Autorin die Rolle algorithmischer Entscheidungen bei der immer stärker auf Effizienz und Kosteneinsparungen optimierten staatlichen Verwaltung von Sozialleistungen und Armut in den Vereinigten Staaten. Ein Fallbeispiel im Buch beschreibt die Folgen eines bereits 2008 im Bundesstaat Indiana eingeführten automatisierten Entscheidungssystems, welches über die Berechtigung (*eligibility*) zum Erhalt von dringend benötigter Sozialhilfe anhand der automatischen Auswertung eines seitenlangen Antragsformulars und der im Rechenzentrum verfügbaren Daten entscheidet. Wie im Buch beschrieben, sind die Folgen solcher nur schwer anfechtbarer Entscheidungen für Betroffene oder deren Angehörige nicht selten existenzbedrohend. Gerade Menschen in prekären Lebenslagen besitzen oft nicht die Kapazitäten, um sich gegen einmal getroffene und behördlich legitimierte algorithmische Entscheidungen zu wehren, geschweige denn, diese auch nur annähernd nachzuvollziehen. Ein weiteres, europäisches Beispiel findet sich beim österreichischen Arbeitsmarktservice (AMS, vergleichbar mit dem Jobcenter in Deutschland). Ab Januar 2021<sup>2</sup> soll dort ein algorithmisches Entscheidungssystem eingesetzt werden, welches anhand der persönlichen Daten einer als arbeitssuchend gemeldeten Person die Aussichten auf Wiedereingliederung in den Arbeitsmarkt schätzt (Fanta, 2018). Einsortiert in eine von drei möglichen Kategorien bekommen diese Menschen mehr oder weniger Unterstützungsleistungen. Problematisch ist unter anderem, dass Personen mit eingetragener Geschlechtskategorie „Frau“, einer Herkunft außerhalb der EU, gesundheitlichen Einschränkungen oder anderen Markern sozialer Ungleichheit der zugrundeliegenden Statistik zufolge perspektivisch schwerer in den Arbeitsmarkt einzugliedern sind und somit weniger Punkte erzielen. In der Folge entscheidet das System, dass

diese Personen letztlich weniger unterstützenswert sind. Auch in Deutschland werden algorithmische Entscheidungssysteme in Behördenvorgänge integriert. Auf der Suche nach höherer Effizienz bei der Entscheidung über Asylanträge lies das Bundesamt für Migration und Flüchtlinge (BAMF) von einer Softwarefirma eine sogenannte Dialekterkennungssoftware entwickeln, um den „Sprachhintergrund“ und damit die Herkunft von Asylbewerber:innen zu schätzen (BAMF 2018). Die Angaben der Personen werden mit den Ergebnissen des Systems verglichen und so eventuelle Widersprüchlichkeiten „aufgedeckt“. Das 2018 als „bestes Digitalisierungsprojekt“ ausgezeichnete „Sprachbiometrische Assistenzsystem“ basiert laut der Reporterin Anna Biselli allerdings auf allerlei fragwürdigen Annahmen (etwa, dass Dialekte vor Ländergrenzen halt machen) und ist mit einer Erfolgsquote von nicht genau überprüfbar 85% auch noch fehleranfällig (Biselli 2018). Die Mitarbeiter:innen erhalten nur in den seltensten Fällen die eigentlich benötigten speziellen Schulungen, um umfassend über die Funktionsweise und Grenzen dieser Systeme informiert zu sein, öffentliche Transparenz oder unabhängige Kontrolle sind nicht vorgesehen.

Als Mittel für eine kostengünstige, schnelle und auch weniger fehleranfällige Optimierung von Arbeitsabläufen scheinen algorithmische Entscheidungssysteme und verwandte Technologien wie Künstliche Intelligenz, maschinelles Lernen und Big Data wie geschaffen zu sein – mit unter anderem den oben beschriebenen Folgen. Zugunsten dieser Ziele wird übersehen oder billigend in Kauf genommen, dass datenbasierte algorithmische Entscheidungssysteme meist den status quo reproduzieren, denn Kern dieser Systeme ist die Entscheidungsfindung auf Basis von Daten aus der Vergangenheit. Lässt sich eine andere, gar „feministische KI“ in Anbetracht aller bisher genannten Negativbeispiele überhaupt denken? Im Folgenden soll zuerst ein grundlegendes Verständnis dieser Technologien und zugrundeliegender Mechanismen vermittelt werden, um sich anschließend Perspektiven und Grundlagen alternativer Nutzbarkeit anzunähern.

1 Der Grundgedanke, Computertechnologien universal auf so gut wie alle Probleme anzuwenden wurde von Evgeny Morozov kritisch „Solutionismus“ genannt. Siehe Morozov (2013).

2 Aufgrund fehlender Technikfolgenabschätzung und nicht zuletzt öffentlicher Kritik steht die flächendeckende Einführung des Algorithmus zum genannten Zeitpunkt momentan in Frage. Siehe Wimmer (2020).

# KLEINE BEGRIFFSKUNDE

Aus Informatiker:innen-Perspektive wird ein ALGORITHMUS als eine endliche, exakt festgelegte Abfolge von Anweisungen definiert – eine Schritt-für-Schritt Anleitung für das Berechnen des größten gemeinsamen Teilers zweier Zahlen beispielsweise. Die Eingabe sind zwei Zahlen (*Eingabevariablen*), diese Zahlen werden einer Anweisung folgend bearbeitet, das Ergebnis oder die Ausgabe ist der größte gemeinsame Teiler (*Ausgabevariable*). Algorithmen können aber nicht nur für mathematische Berechnungen eingesetzt werden. Vielmehr lässt sich so gut wie alles, was in bestimmte Datenformate übersetzen werden kann, algorithmisch bearbeiten. Auf Basis vorliegender Daten, bspw. der persönlichen Daten einer Person (*Eingabevariable*), lässt sich eine Schätzung, bspw. ihrer Chancen in den Arbeitsmarkt integriert zu werden (*Ausgabevariable*), berechnen. Zwischen Eingabe und Ausgabe steht eine fest definierte „Übersetzung“ der Eingabedaten. Diese ist in den Grenzen der Übersetzbarkeit von Phänomenen in computerlesbare Daten beliebig wählbar und kann von simplen mathematischen Berechnungen bis hin zu datenbasierten, statistischen Schätzungen reichen. Wenn Algorithmen, in den einleitenden Beispielen, Entscheidungen unterstützen oder sogar ganz übernehmen, wird von *algorithmischen Entscheidungssystemen* gesprochen.<sup>3</sup> In manchen Fällen unterstützen sie dann beispielsweise die Entscheidungen von Sachbearbeiter:innen im Jobcenter. Entscheidungen über eine Kreditvergabe werden heute oft vollautomatisch von Algorithmen getroffen. Auch im sogenannten High-Frequency-Trading (HFT) entscheiden sich selbst verbessernde Algorithmen in Sekundenbruchteilen über den Kauf oder Verkauf von Finanzprodukten.

KÜNSTLICHE INTELLIGENZ (KI) bezeichnet ein seit 1956 bestehendes Forschungsgebiet der Informatik. Die Gründungsväter<sup>4</sup> der KI waren damals überaus selbstsicher davon überzeugt, innerhalb kürzester Zeit menschenähnliche Intelligenz, verstanden als generelle Problemlöse- oder Entscheidungsfähigkeit, nachbilden zu können. Wie die Geschichte zeigt, sind manche Systeme über 60 Jahre später zwar sehr gut darin, GO zu spielen und Quizfragen zu beantworten; von menschenähnlichen kognitiven Fähigkeiten sind die Systeme jedoch auch 2020 noch weit entfernt. Deswegen wurde dazu übergegangen, zwischen starker und schwacher KI zu unterscheiden: starke KI liegt noch in ferner Zukunft, mit schwacher KI wird die Spezialisierung auf ganz konkrete Probleme bezeichnet, welche „intelligent“ gelöst werden sollen. Solche hochspezialisierten Systeme können dann jeweils zum Beispiel erfolgreich Computerspiele bestreiten, Sprache verstehen oder Objekte in Bildern erkennen. Der Intelligenzbegriff ist traditionell sehr uneindeutig. Der vom Informatiker Alan Turing entworfene „Turing Test“ definiert eine Maschine als intelligent, wenn ein Menschen im Chatkontakt mit dieser Maschine und einem anderen Menschen nicht zwischen diesen unterscheiden kann (Turing 1950: 433–460). Aus einer soziologischen Perspektive betrachtet, zeichnet intelligente Akteur:innen eher die Fähigkeiten aus, in einer komplexen Umwelt in wechselseitiger Abhängigkeit mit Anderen zu kooperieren. Der Name des Forschungsgebiets gibt also ein



mehrdeutiges Ziel vor, sorgt mit Science-Fiction Assoziationen allerdings für viele Missverständnisse.

Ein Teilgebiet der KI ist das MASCHINELLE LERNEN. Das Ziel ist hier, Algorithmen so zu gestalten, dass sie in der Lage sind aus Erfahrungen zu „lernen“, also sich mit neuen Informationen selbst zu verbessern. Statt eine exakte Schritt-Für-Schritt Anleitung zur Lösung eines bestimmten Problems zu benötigen, wird die Erstellung der Anleitung oder, wie oben beschrieben, der Vorgang der Übersetzung von Ein- in Ausgabedaten, dem Algorithmus in Teilen selbst überlassen. Seit 2012 sind *künstliche neuronale Netze* die bekanntesten und auch erfolgreichsten Modelle des maschinellen Lernens und der KI. Weil diese Netze teilweise sehr viele (tiefe) „Schichten“ haben, in welchen die Eingabedaten immer weiter „übersetzt“ werden, wird hier auch von *deep learning* gesprochen. Beim überwachten Lernen<sup>5</sup> „lernt“ ein solches Netz in der Trainingsphase<sup>6</sup> anhand einer riesigen Anzahl von Beispieldaten (Stichwort Big Data) beispielsweise – in diesem Fall relativ unproblematisch – bestimmte Tierarten in Bildern zu erkennen. Ein einmal trainiertes Netz ist nach Überprüfung der Fehlerrate dann für den Einsatz auf neue Bilder anwendbar und liefert Wahrscheinlichkeitswerte für jede vorher trainierte Tierart im Bild. Dieses Zuweisen von Wahrscheinlichkeitswerten wird oft auch als „Voraussage“ (engl. *prediction*) bezeichnet. Eine Art extremes Beispiellernen und grundsätzlich verschieden von der Art, wie Menschen lernen. Ein Kind muss sich glücklicherweise keine Millionen Katzenbilder anschauen, um eine Katze zu erkennen. Das Verfahren wird vor allem im Bereich der Objekterkennung in Bildern (z.B. autonomes Fahren), dem Übersetzen von Texten (z.B. DeepL) oder der Sprachverarbeitung (z.B. SIRI, Alexa) relativ erfolgreich eingesetzt.<sup>7</sup> Diese Systeme sind fast immer eine *black box*. Zum einen, weil Unternehmen die Quellcodes ihrer Software fast nie offenlegen. Zum anderen, weil im Beispiel von selbstlernenden neuronalen Netzen nicht einmal die Programmierer:innen genau verstehen können, *was genau* ein solches System lernt, welche Zwischenschritte vom Eingangsbild zur schlussendlichen Klassifikation führen.

DATEN werden unter anderem wegen ihrer Anwendung in der oben genannten Art von maschinellen Lernverfahren teilweise als „das neue Öl“ der digitalen Ökonomie bezeichnet. Das ist auch der Grund, warum viele Unternehmen ein so großes Interesse an möglichst allen Daten haben, die sie in die Finger bekommen können. Gespeichert wird erst mal alles, in der Hoffnung es irgendwann nutzen zu können. Maschinelle Lernverfahren wie oben beschrieben funktionieren nur, wenn es die entsprechenden Beispieldaten gibt; also etwa Millionen Tierbilder, die auch vorher als solche von Menschen gekennzeichnet worden sind. Wenn ihr im Internet schon einmal sogenannten „Captchas“ begegnet seid, übernehmt ihr damit genau diese Art von Arbeit – dem sogenannten labeln oder kennzeichnen von Beispieldaten. Neben allgemeinen Daten sind personenbezogene Daten für Unternehmen besonders interessant. Für die Individualisierung im consumer-Bereich oder eben in der staatlichen Verwaltung werden diese bereits vielfältig genutzt. Das Grundprinzip ist immer ähnlich: große Mengen an Beispieldaten werden als repräsentativ angenommen, neue „Fälle“ werden anhand dieser Repräsentationen eingeordnet, kategorisiert, bewertet. Ein Problem ist hier oft die Diversität der Datensätze. Beispielsweise hat Joy Buolamwini, Forscherin am MIT (Massachusetts Institute of Technology), in einem Forschungsprojekt nachgewiesen, dass Gesichtserkennungssysteme vor allem

mit Bildern von *weißen*, männlichen Gesichtern trainiert wurden und so die Erkennung für weibliche BPoC selten oder nie funktionierte (Buolamwini & Gebru 2018). Ein weiteres Problem sind die Besitzverhältnisse im Hinblick auf die Datensätze: als lukrative Kapitalanlage sind sie fast immer das Privateigentum von Unternehmen und dementsprechend unter Verschluss (u.a. Zuboff 2018). Beim Beispiel der Dialekterkennungssoftware lässt sich weder überprüfen, welche Sprachaufnahmen als Trainingsdaten ausgesucht wurden, noch wie genau der Algorithmus programmiert wurde. Generell lässt sich das Prinzip des Lernens anhand von existierenden Beispieldaten auch noch unter einem anderen Gesichtspunkt kritisieren. Denn selbst wenn es eines Tages gelänge, Datensätze sowohl divers als auch transparent zu gestalten, so werden durch diese Verfahren bestehende Verhältnisse potentiell zementiert und reproduziert. Das oben bereits genannte System des AMS in Österreich schätzt beispielsweise die Aussicht auf Integration in den Arbeitsmarkt immer nur anhand von bestehenden, also historischen Daten ein. In einer patriarchalen, rassistisch strukturierten Gesellschaft sind die Aussichten in dieser Hinsicht für *gebildete, weiße* Österreicher tatsächlich besser. Mit dem Einsatz dieser Systeme werden diese Verhältnisse mit technischen Mitteln verobjektiviert und immer weiter reproduziert, die Gegenwart wird in die Zukunft projiziert. Das System funktioniert also gar nicht falsch, sondern reproduziert einen Teilbereich gesellschaftlicher Realitäten. Datenbasierten algorithmischen Entscheidungssystemen lässt sich somit ein tendenzieller Konservatismus attestieren; zumindest wenn sie so eingesetzt werden, wie es heute üblich ist.

3 Für eine Übersicht bietet sich der Atlas der Automatisierung von Algorithm Watch an, eine NGO welche u.a. von der Bertelsmann Stiftung gefördert wird: <https://algorithmwatch.org/out-now-atlas-of-automation/>.

4 Auch heute sind in Deutschland nur ca. 16% Prozent aller KI-Fachkräfte nicht cis-männlich: <https://www.gew.de/aktuelles/detailseite/neuigkeiten/nur-wenige-frauen-sind-im-ki-bereich-taetig/>.

5 Überwachtes Lernen bedeutet, dass eine definierte Ausgabe existiert und so laufend überwacht werden kann, ob ein Algorithmus „richtig“ lernt. Es gibt noch das „unüberwachte Lernen“ und das „Verstärkungslernen“ mit etwas anderer Funktionsweise für nicht klar definierbare Ausgaben.

6 Das Trainieren von solchen Modellen ist enorm ressourcenintensiv – ein aktuelles, extremes Beispiel aus der künstlichen Sprachverarbeitung kostet pro Training um die 150.000\$ und zeichnet sich laut einer Studie von 2019 durch einen CO2-Fußabdruck-Äquivalent von circa 5 Autos gerechnet auf ihre gesamte Lebenszeit inklusive Herstellungsprozess und Kraftstoff aus. Siehe Strubell et al. (2019).

7 Die besten Algorithmen haben 2017 bei der standardisierten ImageNet Datenbank mit ca. 14 Millionen annotierten Bildern und 20.000 Kategorien mit einer Erfolgsrate von 95% ein vorläufiges Maximum erreicht.



# EIN UNIVERSALES WERKZEUG?

Wie beschrieben sind moderne algorithmische Entscheidungssysteme basierend auf Methoden des maschinellen Lernens prinzipiell universell anwendbar, egal ob auf Bilder, Sprachaufnahmen, persönliche Daten oder eine Kombination dieser Daten. Deswegen sind diese Verfahren extrem beliebt, lassen sich gut verkaufen und werden tatsächlich auch auf fast alles Mögliche angewandt. Ein Beispiel: trainiert auf die Profilfotos und Angaben zur Sexualität von Profilen einer Datingplattform sagt ein System der Arbeitsgruppe eines Professors in Stanford 2018 mit gewisser Zuverlässigkeit die Angaben zur Sexualität vorher nicht trainierter Profilfotos derselben Datingplattform voraus – das erste „*AI-gaydar*“ war geboren (Wang & Kosinski 2018). Selbstbewusst und in der Tradition westlicher, männlicher Wissenschaftstradition spekulierte der Professor in der Veröffentlichung über das Vorhandensein von gewissen, von Menschen nicht wahrnehmbaren, Gesichtszügen, die Homosexualität „beweisen“. Kulturelle Marker wie etwa Gesichtspflege, gewisse Posen oder auch etablierte visuelle Standards in der Community genau dieser Datingplattform – und damit nicht auf eine einfache, konstante Formel reduzierbare Erklärungsansätze – wurden dabei wenig diskutiert; schließlich ist *mann* hier auf der Suche nach der Essenz von Sexualität – und einen objektiven Wissenszugriff darauf. Diverse Autor:innen wiesen in kritischen Beiträgen darauf hin, dass solche Anwendungsfälle unleugbare Ähnlichkeiten zur Pseudowissenschaft der Physiognomie aufweisen, welche bereits im vorletzten Jahrhundert als haltlos entlarvt wurde.<sup>8</sup> Vertreter:innen dieser Strömung behaupteten damals, den Charakter eines Menschen aus dem Gesicht ableiten zu können und so auf wissenschaftliche Art entweder auf- oder abwerten zu können. Diese und ähnliche Methoden waren (und sind es teilweise auch heute noch) die vermeintlich wissenschaftliche Grundlage von Rassismus und anderen kategorisierenden, auf Ein- und Ausgrenzung basierenden Ideologien. Nicht wenig überraschend ist, dass ähnliche Systeme wie das obige auch schon angewendet wurden, um die „Kriminalität“ von Menschen anhand ihres Gesichts zu bestimmen (vgl. Wolf 2019).

Verfahren des maschinellen Lernens, welche Daten in diese oder jene Kategorie einsortieren, erzeugen also unweigerlich Aussagen, welche durchaus als „Wahrheiten“ interpretiert werden (können) und diese so gewissermaßen auch (re)produzieren. Die tendenziell (technik)deterministische und essentialistische, also die Suche nach tieferen, letzten und unabänderlichen Wahrheiten bezüglich (sozialer) Phänomene und der Glaube, dass Technik diese aufdecken kann, erfährt dank maschinellen Lernverfahren eine Art Revival.<sup>9</sup> Dabei ist klar, dass diese Verfahren aber maximal Korrelationen, also lose, lokale Zusammenhänge zwischen – im obigen Beispiel – Profilbild und sexueller Orientierung oder – im Beispiel der Dialekterkennungssoftware – zwischen Sprachaufnahme und Herkunftsland finden. Dabei ist wichtig zu bemerken, dass auch diese lauter Einschränkungen unterworfen sind:

erstens, dass, wie jede:r Erstsemester:in der Statistik lernt, sich Kausalitäten nicht aus Korrelationen ableiten lassen. Zweitens sind sie beschränkt auf beispielsweise genau die eine Datingplattform zu der Zeit der Datenaufnahme; schließlich wandeln sich beispielsweise soziale, visuelle Codes mit der Zeit und können sich auch von Community zu Community stark unterscheiden. Auch hier ist der im vorherigen Absatz erwähnte, konservative Charakter von datenbasierten algorithmischen Entscheidungssystemen erkennbar.

8 Zum Weiterlesen (engl.): <https://medium.com/@blaisea/physiognomys-new-clothes-f2d4b59fdd6a>.

9 Für eine auf Big Data bezogene Auseinandersetzung siehe bspw. Prietl (2019).

# EINE FEMINISTISCHE KRITIK DER OBJEKTIVITÄT

Aus feministischer Perspektive wurde der Zugriff auf vermeintlich „tiefere Wahrheiten“ und „objektives Wissen“ mithilfe der Wissenschaft schon in den 80er Jahren von feministischen Denker:innen als patriarchale Herrschaftstechnik kritisiert. So attestierte Donna Haraway westlichen, meist *weißen* Wissenschaftlern die Konstruktion eines „Blicks von Nirgendwo“ (engl. *view from nowhere*), um aus einer vermeintlich neutralen Position der Wissenschaft Allgemeinheitsansprüche oder Wahrheiten durchzusetzen. Dem setzt sie das Konzept von situiertem Wissen (*situated knowledges*) entgegen, welches besagt, dass *jedes* Wissen von spezifischen Körpern unter besonderen materiellen, kulturellen und politischen Bedingungen produziert wird und so auch verstanden werden muss. Diese Perspektive könnte auch auf die Untersuchung von algorithmisch produziertem Wissen angewandt werden. Statt sich auf die Suche nach nur einer Objektivität zu begeben, entwickelte Sandra Harding das Konzept von starker Objektivität (engl. *strong objectivity*). Dieses betont, dass kontextualisiertes, in gegenseitiges Verhältnis gesetztes Wissen im Vergleich zu vermeintlich wertneutralem Wissen einen „stärkeren“ Objektivitätsanspruch hat – besonders von Personen, welche von Ungleichheitsverhältnissen betroffen sind. Das Wissen von intersektional betroffenen Personen erhält also einen privilegierten Zugang zu sonst unsichtbarem, im Forschung und Wissenschaft übersehenem Wissen und

verdient nach Harding besondere Beachtung. So sollte dies bei der Untersuchung von algorithmischen Entscheidungssystemen stets mitgedacht und so weit wie möglich eingearbeitet werden. Die beiden Schlaglichter sind nur kleine Ausschnitte einer möglichen feministischen Intervention in die Technisierung von Sozialstaat und Co. Sie sind sicher nicht die einzigen Konzepte, welche helfen können, sich aus feministischer Perspektive kritisch mit den Produktionen von algorithmischen Entscheidungssystemen auseinander zu setzen und diese in ihren Kontext eingebettet zu verstehen. Dies steht im Gegensatz zu einer vermeintlich wissenschaftlichen, objektivierenden „Neutralisierung“ von zugrundeliegenden Werturteilen und Folgen von Technologien, welche dadurch bestehende Machtverhältnisse zementieren. So wird im öffentlichen Diskurs oft von „Algorithmen“ gesprochen, wenn eigentlich das Zusammenspiel verschiedener Teile eines sozio-technischen Systems, also einer Mischung aus technischen und menschlich-sozialen Teilen, gemeint ist. Der oben erwähnte AMS Algorithmus ist aus einer ausdifferenzierten sozio-technischen Perspektive betrachtet kein klar abgrenzbares Programm, sondern eher ein Netzwerk verschiedener Akteur:innen, ihrer Interessen, (historischer) Daten und so weiter. Die Vorannahmen bei der Entwicklung, die zugrundeliegenden Statistiken, die in die Berechnung einfließenden persönlichen Daten, die von der Berechnung ausgeschlossenen Daten, die überwiegend *weißen* Programmierer des Programms (Myers West et al. 2019), die zugrundeliegenden Interessen und Zielvorgaben des Unternehmens<sup>10</sup> und des Staates; all dies spielt neben den Codezeilen eine beachtliche Rolle. Die Technologien an sich wirken wie eine Art Neutralisierungsfilter, welche alles nicht-Technische verdecken können. „Algorithmus“ ist sozio-technisch betrachtet ein diffuser Begriff und wird in der Öffentlichkeit doch oft implizit mit etwas rein Technischem, vermeintlich Neutraleem und sich unserem Zugriff entziehenden assoziiert.<sup>11</sup> Dem gilt es etwas entgegen zu setzen.

<sup>10</sup> Der Arbeitsmarktservice in Österreich ist wie das Jobcenter in Deutschland nach dem Vorbild eines klassischen Unternehmens organisiert.

<sup>11</sup> Wie Algorithmen im öffentlichen Diskurs legitimiert werden findet sich am Beispiel der o.g. Dialekterkennungssoftware des BAMF bei Keiner (2020).

## AUSBLICK

Perspektivisch wird sich das Einsatzgebiet algorithmischer Entscheidungssysteme aller Wahrscheinlichkeit nach weiter vergrößern. Autohersteller:innen nutzen die Systeme für die Verarbeitung von Sensoren, um autonomes Fahren zu ermöglichen, Werbetreibende sortieren anhand von Verhaltensdaten Konsument:innen in immer ausdifferenzierte Zielgruppen, Plattformbetreiber:innen erzeugen individuelle *Feeds*, um Nutzer:innen möglichst lange bei der Stange zu halten, Sicherheitsbehörden setzen die Technologie ein, um verdächtige

Menschen zu identifizieren und zu überwachen und so weiter. Während sich im Fall des autonomen Fahrens noch argumentieren ließe, dass bei beispielsweise der Unterscheidung von Fußgänger:innen und Verkehrsschildern unverdächtige Entscheidungen über die Bewegung eines Fahrzeugs getroffen werden – so diese denn zuverlässig funktionieren – so wird es mit Blick auf Entscheidungen bezüglich fundamentalen Rechten und Bedürfnissen (meist sowieso schon marginalisierter) Personen betreffend sehr viel problematischer. Die hier aufgeführten Beispiele und unzählige weitere Berichte machen eine kritische öffentliche Debatte unersetzlich. Leider fehlt es an gehörten kritischen Stimmen aus der Informatik und in der Öffentlichkeit an Wissen über die Funktionsweise und Hintergründe von Technologien.

Sind algorithmische Entscheidungssysteme also rundheraus abzulehnen, nichts weiter als weitere Herrschaftsinstrumente zur Festigung bestehender Ungleichheitsverhältnisse? Unter den gegebenen Bedingungen scheint sich diese Schlussfolgerung aufzudrängen. Allerdings sollte dabei nicht aus den Augen verloren werden, dass Technologien bestehende Machtverhältnisse auch „nur“ spiegeln und diese durch massenhafte Anwendung verstärken können. Sie sind gewissermaßen Resultate politischer Auseinandersetzungen und damit umkämpftes Terrain. Eine demokratische Aneignung ist theoretisch möglich. Die in die Technologien eingeschriebenen, codierten Herrschaftsverhältnisse in KI, Big Data und Co. müssen verstanden und decodiert, gewissermaßen gehackt werden. Dazu bedarf es eines ausdifferenzierten Verständnisses dieser Technologien, welche nicht als isolierte Phänomene, neutrale Werkzeuge oder fremdartige Artefakte behandelt werden sollten, sondern als in Technik gegossene Gesellschaftsverhältnisse, welche unter basisdemokratische Kontrolle gebracht werden können. Um zu einem kontextsensitiven, intersektionalen Verständnis zu gelangen, ist die Einbeziehung und Beteiligung gerade betroffener, marginalisierter Personengruppen im Sinne einer partizipativen Technikanalyse und -entwicklung zentral und ein möglicher erster Schritt. Ebenso prominent sollte die Klärung der Frage stehen, ob algorithmische Entscheidungssysteme – nur weil sie vorgeblich universal eingesetzt werden *können* – überhaupt in allen Bereichen eingesetzt werden *sollten*. Danach könnte eine basisdemokratische, partizipative Technikentwicklung von den Positionen der Betroffenen im Besonderen und der Gesellschaft im Allgemeinen ausgehend die Erfahrungen und Bedürfnissen möglichst diverser Positionen einbeziehen. Ausgehend von einem sorgfältigen Verständnis bestehender Verhältnisse und unter Bezugnahme auf reale Bedürfnisse Vieler lässt sich KI durchaus feministisch denken und entwickeln. Im Gegensatz zu den universalistischen, verobjektivierten Annahmen von Silicon Valley und Co., für welche Technik *immer* die „richtige“ Lösung zu sein scheint, kann eine feministische Technikentwicklung ein tastendes Voranschreiten ohne dogmatische Vorannahmen sein, welches die heute beobachtbaren negativen Folgen einer neoliberalen, kapitalistischen Technikentwicklung umgehen kann.

Phillip Lücking arbeitet als sozialwissenschaftlicher Informatiker und forscht zur Reproduktion sozialer Ungleichheiten durch algorithmische Systeme, emanzipatorischen Algorithmen und einer postkapitalistischen KI für alle.

# LITERATUR

- BAMF (2018)  
„eGovernment Preisverleihung: 1. Platz“, in: *BAMF Meldungen*, 21.06.2018, online unter: <https://www.bamf.de/SharedDocs/Meldungen/DE/2018/20180621-am-egovernment.html> [abgerufen am 4.9.2020].
- Biselli, Anna (2018)  
„Eine Software des BAMF bringt Menschen in Gefahr“, in: *Vice*, 20.08.2018, online unter: <https://www.vice.com/de/article/a3q8wj/fluechtlinge-bamf-sprachanalyse-software-entscheidet-asyl> [abgerufen am 4.9.2020].
- Buolamwini, Joy & Gebru, Timnit (2018)  
„Gender shades: Intersectional accuracy disparities in commercial gender classification“, *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, PMLR 81, 77–91.
- Der Spiegel (2018)  
„Google entschuldigt sich für fehlerhafte Gesichtererkennung“, in: *Der Spiegel*, 02.07.2015, online unter: <https://www.spiegel.de/netzwelt/web/google-fotos-bezeichnet-schwarze-als-gorillas-a-1041693.html> [abgerufen am 4.9.2020].
- Eubanks, Virginia (2018)  
*Automating inequality: How high-tech tools profile, police, and punish the poor*, New York: St. Martin's Press.
- Fanta, Alexander (2018)  
„Österreichs Jobcenter richten künftig mit Hilfe von Software über Arbeitslos“, in: *Netzpolitik.org*, 13.10.2018, online unter: <https://netzpolitik.org/2018/oesterreichs-jobcenter-richten-kuenftig-mit-hilfe-von-software-ueber-arbeitslose/> [abgerufen am 4.9.2020].
- Köver, Chris (2019)  
„Gesichtserkennung: Kritik macht Algorithmen genauer, nicht nur für weiße Männer“, in: *Netzpolitik.org*, 7.2.2019, online unter: <https://netzpolitik.org/2019/gesichtserkennung-kritik-macht-algorithmen-genauer-nicht-nur-fuer-weisse-maenner/> [abgerufen am 4.9.2020].
- Morozov, Evgeny (2013)  
*Smarte neue Welt: digitale Technik und die Freiheit des Menschen*. München: Karl Blessing Verlag.
- Myers West, Sarah; Whittaker, Meredith & Crawford, Kate (2019)  
„Discriminating Systems“, in: *AI NOW*, online unter: <https://ainowinstitute.org/discriminatingsystems.pdf> [abgerufen am 4.9.2020].
- Prietl, Bianca (2019)  
„Die Versprechen von Big Data im Spiegel feministischer Rationalitätskritik“. *GENDER–Zeitschrift für Geschlecht, Kultur und Gesellschaft* 11(3), 11–25.
- Keiner, Alexandra (2020)  
„Algorithmen als Rationalitätsmythos“, in: Leineweber, Christian/de Witt, Claudia (Hg.), *Algorithmisierung und Autonomie im Diskurs: Perspektiven und Reflexionen auf die Logiken automatisierter Maschinen*, Hagen: FernUniversität in Hagen, 47–67.
- Strubell, Emma; Ganesh, Ananya & McCallum, Andrew (2019)  
„Energy and policy considerations for deep learning in NLP“, *arXiv preprint arXiv:1906.02243*.
- Turing, Alan (1950)  
„Computing Machinery and Intelligence“, in: *Mind* 59(236), 433–460.
- Wang, Yilun & Kosinski, Michael (2018)  
„Deep neural networks are more accurate than humans at detecting sexual orientation from facial images“, in: *Journal of personality and social psychology* 114(2), 246–257.



- \_\_\_\_\_ Wimmer, Babarar (2020)  
„AMS-Algorithmus sollte ganz abgedreht werden“  
in: *futurezone.at*, 24.8.2020, online unter:  
<https://futurezone.at/netzpolitik/ams-algorithmus-sollte-ganz-abgedreht-werden/401009924> [abgerufen am 12.10.2019]
- \_\_\_\_\_ Wolf, Christian (2019)  
„Was steht uns ins Gesicht geschrieben?“, in:  
*Spektrum.de*, 18.08.2019, online unter:  
<https://www.spektrum.de/news/emotionen-kriminalitaet-und-charakter-was-kann-gesichtserkennung-aus-dem-gesicht-lesen/1629144> [abgerufen am 4.9.2019]
- \_\_\_\_\_ Zuboff, Shoshona (2018)  
*Das Zeitalter des Überwachungskapitalismus*,  
Frankfurt/New York: Campus Verlag.



KI UND  
BESCHÄFTIGUNG  
:  
DAS ENDE  
MENSCHLICHER  
VORURTEILE  
ODER  
DER BEGINN  
VON  
DISKRIMINIERUNG  
2.0?



DENIZ  
ERDEN

In der Vergangenheit wurden Frauen im Bereich der bezahlten Arbeit diskriminiert. Trotz Verbesserungen ist der gleichberechtigte Zugang zu dieser wichtigen wirtschaftlichen Ressource bisher nicht erreicht worden. Nicht-Diskriminierungs-Regelungen sind zwar inzwischen Teil vieler nationaler wie internationaler rechtlicher Rahmenbedingungen, allerdings noch nicht im ausreichenden Maße. Mittlerweile hat sich der Einsatz von Künstlicher Intelligenz (KI), genauer gesagt von Machine-Learning-Algorithmen, Algorithmen des maschinellen Lernens, vor allem im Zusammenhang wichtiger Entscheidungsprozesse, wie z.B. bei der Einstellung für einen Arbeitsplatz, als neue Bedrohung herausgestellt, die Anlass zu Bedenken hinsichtlich diskriminierender Folgeschäden gibt.

Die Automatisierung bei Personalentscheidungen wie Einstellungen oder Beförderungen, wurde als Instrument zur Bekämpfung menschlicher Vorurteile und subjektiv diskriminierender Entscheidungen eingeführt, welche bekanntermaßen systemischen Mustern gegenüber benachteiligten Gruppen folgen. Diese Bekämpfung von Vorurteilen sowie diskriminierender Entscheidungen zeitigt jedoch nur begrenzten Erfolg. Vielmehr kam mit der Einführung automatisierter Personaleinstellungen eine neue Form der Bedrohung mittels maschinell erlernter Algorithmen ins Spiel, denn sie sind datengesteuert (Kim 2017). Und Daten, wie feministische und andere kritische Wissenschaftler\*innen seit Langem zeigen, sind nicht neutral. Neben der Einführung neuer Formen von Vorurteilen ist es deshalb wahrscheinlich, dass automatisierte Entscheidungen, die auf Daten basieren, die bestehende Vorurteile in der Gesellschaft aufrechterhalten und verstärken, selbst wenn sie gut gemeint sind.

Es gibt ein scheinbar wachsendes Ökosystem, in dem mehrere Anbieter Lösungen für Vielfalt und Integration als Dienstleistung anbieten und versprechen, dass ihre unvoreingenommenen Modelle zu integrativeren Entscheidungen und vielfältigeren Teams führen werden. Einige Studien haben sich bereits damit befasst, wie solche Technologien funktionieren (Bogen & Rieke 2018; Langenkamp et al. 2020; Raghavan et al. 2020; Sánchez-Monedero et al. 2020; Tambe et al. 2020). Ich werde von diesen Studien profitieren, um die Herausforderungen zu verstehen, die diese technologischen Lösungen vermutlich in Angriff nehmen, ungelöst lassen, aufrechterhalten oder verstärken werden. Daher ist es wichtig, die Gründe für die frühere Diskriminierung zu identifizieren und zu erkennen, wie sie sich in Modelle einschleichen, die diese Vergangenheit überwinden wollen. Das Gesetz in seiner Gesamtheit an Regelungskomplexen wurde instrumentalisiert, um Frauen und andere benachteiligte Gruppen sowohl zu ermächtigen, als auch unterzuordnen. Ich halte es für nützlich, die Art und Weise der Regelsetzung im juristischen Bereich mit der Kodierung von machine-learning Algorithmen zu vergleichen. Beide wären in der Pflicht, faire Ergebnisse für alle zu erzielen, hätten aber sofort Probleme bei der Anwendung allgemeiner Regeln auf eine Vielzahl unterschiedlicher Realitäten. Hierbei geht es nicht darum, dem Nihilismus zum Opfer zu fallen, sondern um die Suche nach den besten Werkzeugen, die zusammenwirken, um sich gegenseitig stärken zu können. Darüber hinaus ist es förderlich auch zu untersuchen, wie Antidiskriminierungsrecht, Datenschutzrecht und Technologie ineinandergreifen, um diesen Herausforderungen in einem europäischen Kontext zu begegnen.

# KÜNSTLICHE INTELLIGENZ IN HUMAN RESSOURCES

Immer mehr Arbeitgeber\*innen verwenden Techniken des machine learnings, welche Muster in den ihnen zur Verfügung stehenden Daten (Schulungsdaten) erkennen, um „Einstellungsmodelle“ mit vorausschauenden Ergebnissen zu erstellen.

Die Nutzung prädiktiver Analysen und automatisierter Entscheidungen ist nicht nur für Arbeitgeber\*innen attraktiv, die Inklusion und Vielfalt erstreben. Einstellungsverfahren sind in der Regel zeitaufwendig, und je mehr Zeit sie in Anspruch nehmen, desto mehr Unternehmensressourcen werden für diese Aufgaben bereitgestellt. Darüber hinaus soll die prädiktive Analyse dabei helfen, bessere Einstellungsentscheidungen zu treffen, was dazu führen kann, besser geeignete Teammitglieder einzustellen, Diebstahl zu verhindern und die Mitarbeiter\*innenzahl zu erhöhen (Bogen & Rieke 2018).

Es ist wichtig zu bedenken, dass es nie eine einzige Entscheidung über die Einstellung oder Nicht-Einstellung von Bewerber\*innen gibt, sondern eine kumulative Reihe kleiner, die in verschiedenen Schritten getroffen werden, wie z.B. die Bekanntgabe einer offenen Stelle, die Bewertung von Qualifikationen oder die Vorauswahl für das Vorstellungsgespräch. Es gibt verschiedene Anbieter\*innen auf dem Markt, die sich auf diese unterschiedlichen Schritte spezialisiert haben, die Arbeitgeber\*innen für die verschiedenen Schritte des Einstellungsverfahrens wählen können. Ein Blick auf diese Vielfalt kann helfen zu verstehen, wie die Automatisierung bei der Einstellung funktioniert. Der „hiring funnel“, zu deutsch „Einstellungs-Trichter“, wird häufig verwendet, um zu zeigen, wie die Anzahl der Kandidat\*innen durch Sourcing, Screening, Interviews und Auswahlstufen reduziert wird. Der Bericht der in Washington ansässigen Non-Profit-Organisation Upturn über Einstellungsalgorithmen gibt Einblick in Billigkeit und Befangenheit in diesen Phasen, von denen der folgende Teil weitgehend profitiert (Bogen & Rieke 2018).

Während der Sourcing-Phase wird ein Bewerber\*innenpool angelegt. Einige Anbieter\*innen<sup>1</sup> bieten Lösungen für die Analyse der Sprache von Stellenbeschreibungen an, um sie für bestimmte Gruppen attraktiver zu machen, einschließlich eines „Gender-Tone-Meters“, der anzeigt, ob der Text für ein Geschlecht abschreckend ist. Arbeitgeber\*innen nutzen auch alle verfügbaren Anzeigetechnologien, die Behavioral Targeting und Microtargeting für bestimmte demographische Gruppen ermöglichen. Potenzielle Mitarbeiter\*innen können über zentralisierte Anzeigenetzwerke auf Job-Plattformen, über Suchmaschinen, auf Social Media, über individuelle Websites und mobile Apps gezielt angesprochen werden. Zur Bestimmung solcher Nutzer\*innengruppen, die gezielt angesprochen oder ausgeschlossen werden sollen, werden üblicherweise Daten verwendet, die von den Nutzer\*innen zur Verfügung gestellt werden, sowie Daten, die aus ihren Online-Aktivitäten abgeleitet werden. Es gibt jedoch Beispiele, in

denen weiblichen und männlichen Arbeitssuchenden diese Anzeigen unterschiedlich zugestellt wurden, z.B. erhielten männliche Nutzer mehr Anzeigen, die für hochbezahlte Stellen werben, oder weibliche Nutzerinnen erhielten weniger häufig eine Anzeige für MINT-Stellen (Mathematik, Informatik, Naturwissenschaft und Technik) (Wachter 2020).

Ein weiterer Prozess ist das Matching, das durch so genannte Recommender-Systeme durchgeführt wird, die entweder auf inhaltsbasierter oder kollaborativer Filterung beruhen. Ersteres beurteilt, woran User\*innen interessiert zu sein scheinen, indem Online-Aktivitäten untersucht werden, während letzteres auf der Grundlage der geschlussfolgerten Interessen von Personen, die den User\*innen ähnlich sind, vorhersagt, woran sie vermutlich interessiert sind. Diese beiden Filterverfahren sind beispielsweise in der Lage, Frauen von Führungspositionen fernzuhalten. Erstens könnten Frauen aufgrund ihrer tatsächlichen Fähigkeiten auf Stellen in niedrigeren Positionen klicken, so dass ein Modell, das nur aus diesem Verhalten lernt, ihre Positionen, die ihren Fähigkeiten entsprechen, möglicherweise nicht zeigt. Andererseits könnte eine Frau, die sich für Führungspositionen interessiert auch aus der Matching-Gruppe ausgeschlossen werden, weil andere Frauen mit ihren Fähigkeiten kein Interesse an solchen Positionen zeigen (Bogen & Rieke 2018).

In der Screening-Phase werden die Bewerber\*innen beurteilt, um festzustellen, ob sie die erforderlichen Qualifikationen oder Spitzenleistungen erbringen. Während Bewerber\*innen, die als nicht qualifiziert gelten, abgelehnt werden, gehen die stärksten Bewerber\*innen in die nächste Phase über. Chatbots, die Natural Language Processing (NLP), natürliche Sprachverarbeitung, verwenden, würden in dieser Phase Gespräche führen, um auf der Grundlage der von den Arbeitgebenden vorgegebenen Kriterien zu entscheiden, ob die Bewerber\*innen für die Stelle qualifiziert sind. Es gibt andere Beurteilungsinstrumente, die den Bewerber\*innen Umfragen zur Selbsteinschätzung, Spiele und interaktive Aktivitäten anbieten, um ihre Persönlichkeitsmerkmale zu erforschen und sie mit den gewünschten Merkmalen der Arbeitgeber\*innen zu vergleichen, die in der Regel auf den derzeitigen oder früheren Spitzenkräften basieren.

In der Interview-Phase werden in der Regel Video-Interviews mit Hilfe verschiedener Technologien wie Spracherkennung, Gesichts- und Stimmanalyse analysiert, um die zukünftige Arbeitsleistung der Bewerber\*innen auf der Grundlage der Leistungskennzahlen der Arbeitgeber\*innen, die wiederum von derzeit erfolgreichen Mitarbeiter\*innen abgeleitet werden, vorherzusagen. Diese immer wiederkehrende Methode, aktuelle Mitarbeiter\*innen als Beispiel zu nehmen ist kritisch zu betrachten, wenn ein Unternehmen einen Wandel anstrebt. Beispielsweise können sich Gesichtsausdrücke und ihre Bedeutungen in verschiedenen Kulturen unterscheiden. Lächeln gilt in Deutschland und China als intelligent, im Iran jedoch als unintelligent (Krys et al. 2014). In diesem Fall könnte das Modell, wenn es anhand der Daten der Top-Performer eines deutschen Unternehmens mit überwiegend deutschen Mitarbeiter\*innen geschult wurde, ein Muster des Lächelns auffinden. Am Anfang mag dies wie ein neutrales Kriterium aussehen, jedoch könnte sich eine Person aus dem Iran während eines Interviews aus Angst, unintelligent zu wirken, vom Lächeln abhalten und deshalb eine negative Bewertung erhalten. Auch wird in vielen Kulturen der Blickkontakt mit dem anderen Geschlecht, insbesondere für Frauen, als negatives Attribut angesehen, und ein Algorithmus, der den Blickkon-

takt als positives Attribut für Engagement und Selbstvertrauen nimmt, kann Frauen aus nicht-westlichen Kulturen diskriminieren.

Die Auswahl-Phase richtet sich an die in die engere Wahl gezogenen Bewerber\*innen, bei denen Backgroundchecks durchgeführt werden sollen. Einige der neuen Tools, die in diesem Bereich angeboten werden, prognostizieren Kinderwünsche und daraus resultierende Betreuungsaktivitäten oder das Risiko der Kandidat\*innen, sich auf toxische Verhaltensweisen wie sexuelle Belästigung einzulassen, indem öffentlich zugängliche Inhalte analysiert werden, die diese Kandidat\*innen in sozialen Medien oder Blogbeiträgen produziert haben. Einige der Einstellungs-Tools in diesem Stadium helfen Arbeitgeber\*innen dabei, die Wahrscheinlichkeit vorherzusagen, mit der Kandidat\*innen ein bestimmtes Angebot annehmen. Durch die Anpassung von Gehalt, Bonus und anderen Leistungen können sie ein Angebot so anpassen, dass es nicht abgelehnt werden kann. In diesem Fall kann das Modell beispielsweise aus dem bestehenden Gender Pay Gap lernen und ihn erfolgreich aufrechterhalten, indem Frauen niedrigere Gehälter angeboten werden (Bogen & Rieke 2018).

# IST FRAUEN- FEINDLICHKEIT DURCH MASCHINEN VERLETZENDER ALS DURCH MITMENSCHEN?

Algorithmen lernen aus Daten, die zwangsläufig historisch sind. Wie Feministinnen artikuliert haben, gibt es eine Reihe von Gründen, warum Frauen in der Vergangenheit von der Beschäftigung ausgeschlossen, segregiert und unterbezahlt wurden. Zunächst einmal hat die Beschränkung der Frauen auf die private Sphäre der Familie als patriarchalische Strategie der Kontrolle über Frauenkörper eine bedeutende Rolle bei der Ausgrenzung von Frauen aus der öffentlichen Sphäre gespielt (Walby 1990). Der Haushalt war das Privatleben der Männer, in dem sie eine Pause von der Geschichtsschreibung



im öffentlichen Raum einlegten, und wo Frauen die Kinder betreuten und reproduktive Hausarbeit verrichteten. Dies wurde als Tugenden der Frauen angesehen, für die sie nicht bezahlt werden sollten. Im Gegensatz dazu wurde alles, was mit Technologie zu tun hat, als qualifizierte Arbeit verstanden. Von der Erfindung der ersten landwirtschaftlichen Werkzeuge bis zur Programmierung des ersten elektronischen Universalcomputers, ENIAC, wurden die Errungenschaften der Frauen in der Technologie jedoch nicht anerkannt und als Teil ihrer normalen Fertigkeiten wie Stricken oder Nähen betrachtet, von denen damals behauptet wurde, sie seien keine große Sache.

Man kann viele Gerichtsurteile finden, die Frauen den Zugang zur Hochschulbildung und zur öffentlichen Repräsentation verweigern, ebenso wie solche, die diskriminierende Einstellungspraktiken billigen, die alle auf die delikate Natur von Frauen verweisen, die zum privaten Bereich gehören, wo sie sich auf die Kinderbetreuung und die Organisation des Haushalts konzentrieren sollten. Hier beanspruchte die Rechtsordnung für sich, eine neutrale Tatsache für alle Frauen hervorzuheben, die auf traditionellen Annahmen beruht und Blindheit gegenüber allen nicht-bürgerlichen Erfahrungen von Frauen widerspiegelt (Fredman 1997). Auf der anderen Seite arbeiteten innerhalb der Arbeiterklasse männliche Arbeiter bewusst in Gewerkschaften, um Frauen in häuslichen Pflichten zu halten, was ein weiterer Grund für die Unterordnung der Frauen an beiden Orten ist, welche sich so gegenseitig verstärken (Hartmann 1976). Die Ausgrenzung erfolgte auch in Form von Schutzgesetzen, die Frauen während der Schwangerschaft und Stillzeit von der Arbeit abhielten. Die Art und Weise, wie diese Gesetze gestaltet wurden, war eine Übersetzung des Modells „männlicher Ernährer – weibliche Betreuerin“, welches die traditionelle genderspezifische Arbeitsteilung auf dem Arbeitsmarkt widerspiegelt. Diese schwächere Position der Frauen auf dem Arbeitsmarkt schmälerte ihre Verhandlungsmacht und führte dazu, dass sie sich auf Niedriglohnbranchen konzentrierten, was auch als Strategie zur Rechtfertigung dafür benutzt wurde, Frauen niedrigere Löhne als Männern zu zahlen, was andernfalls gegen die Gleichstellungsgrundsätze verstoßen würde, wenn sie die gleiche Arbeit verrichten und in der gleichen Fabrik arbeiten würden. Die feministische Forschung hat gezeigt, dass der Ausschluss von Frauen aus der Technologie auch eine Folge der männlichen Dominanz im Handwerk war, die sich während der industriellen Revolution entwickelte (Wajcman 2009). Heute ist die genderspezifische Arbeitsteilung immer noch mit der Assoziation von Technologie, Männlichkeit und der Vorstellung davon verbunden, was als qualifizierte Arbeit gilt.

Ein weiteres Phänomen des genderspezifischen Arbeitsmarktes ist die berufliche Segregation, die sich in zwei Kategorien unterteilt: Die horizontale Segregation bezieht sich auf die Tendenz von Frauen und Männern, in verschiedenen Berufen tätig zu sein, während die vertikale Segregation sich auf die Tendenz von Frauen und Männern bezieht, in verschiedenen Positionen innerhalb desselben Berufs oder Unternehmens beschäftigt zu sein (auch bekannt als die sogenannte „Gläserne Decke“). Heute können wir diesen Trend in Unternehmen beobachten, in denen die Personalabteilungen und die Callcenter überwiegend mit weiblichen Mitarbeiterinnen besetzt sind, während die technischen Abteilungen und der Vertrieb überwiegend von Männern dominiert werden. Es überrascht nicht, dass letztere die besseren Wege zu Spitzenpositionen im Management sind. Darüber hinaus passen die



Anforderungen dieser Positionen wie eine ununterbrochene Karriere, also täglich auf Vollzeitbasis und jederzeit für Reisen verfügbar zu sein, nicht zu den familiären Verpflichtungen. Die berufliche Segregation ist auch für die als Feminisierung von Berufen bezeichnete Tendenz verantwortlich, dass der Eintritt von Frauen zu einer Herabstufung der Bezahlung und des Status eines Berufs führt.<sup>2</sup>

Wenn Frauen bereits so strukturell diskriminiert wurden, warum gehen wir jetzt hart gegen Algorithmen und ihre Entwickler\*innen vor? Und, wenn die Diskriminierung von Frauen auf dem Arbeitsmarkt so tief verwurzelt ist, wie können wir dann erwarten, dass einige Technologieanbieter\*innen sie beheben? Um diese oft gehörten Fragen beantworten zu können, müssen wir uns genauer ansehen, (1) warum und wie maschinell erlernte Algorithmen diskriminieren, (2) wie sich diese von der Diskriminierung durch Menschen unterscheidet, und (3) wie wir die Schäden durch maschinelle Vorurteile mildern können. Aufgrund des begrenzten Umfangs und des begrenzten Platzes ist es mein Ziel, hier einen Einblick in die Beantwortung dieser Fragen zu geben. Daher werden möglicherweise nicht alle Gründe abgedeckt, und für Experten\*innen auf verwandten Gebieten könnte es sich eher vereinfachend lesen.

Das Ziel der Machine-Learning-Algorithmen ist es, etwas vorherzusagen oder abzuschätzen (z.B. ob eine E-Mail Spam ist, ob Bewerber\*innen gute Mitarbeiter\*innen wären oder nicht). Hierfür muss definiert werden, was vorhergesagt werden soll (Spam-E-Mail, gute Mitarbeiter\*innen), und es muss eine Möglichkeit gefunden werden, dies zu messen (Lehr & Ohm 2017). Bei Einstellungsalgorithmen ist die Zielvariable in der Regel „gute\*r Mitarbeiter\*in“, was sehr schwer zu messen ist. Zudem ist es ein Konstrukt, das in Form eines körperlich gesunden cis-Geschlechts-Mannes ohne häusliche Belastung geformt wurde. Arbeitgeber\*innen müssen einige messbare Ergebnisse auf der Grundlage von Fachkenntnissen herausfinden, wie z.B. höhere Leistungsbewertungen, längere Betriebszugehörigkeit oder Verkaufszahlen als gute Indikatoren für gute Mitarbeiter\*innen (Barocas & Selbst 2016); all dies würde wiederum Qualitäten widerspiegeln, die strukturell belastete gute Mitarbeiter\*innen möglicherweise nicht aufweisen können.

Die vorliegenden Daten werden dann in zwei Datensätze unterteilt, die als Trainingsdaten und Testdaten bezeichnet werden. Grundsätzlich würde ein Modell mit den Trainingsdaten lernen, indem es den richtigen Antworten ausgesetzt wird, und dann wird das resultierende Vorhersagemodell (auch Algorithmus genannt) auf neue, nicht klassifizierte Daten angewendet, die als „Testdaten“ bezeichnet werden. Wenn der\*die Entwickler\*in davon überzeugt ist, dass er\*sie bereit ist, wird der Algorithmus so eingesetzt, dass er\*sie beim Treffen von Entscheidungen im wirklichen Leben, die aus dem Training resultieren, funktioniert. Das Konzept „Garbage In, Garbage Out“ ergibt sich aus der Tatsache, dass voreingenommene Trainingsdaten dazu führen würden, dass das Modell diskriminiert. Beispielsweise ist bei Techniken des überwachten Lernens<sup>3</sup> die Zielvariable, die das Modell vorherzusagen versucht, bekannt und wird von den Programmierern gekennzeichnet (z.B. wissen, was eine Spam-Mail ist, und sie als solche kennzeichnen, so dass die Maschine lernen kann, was Spam ist). Mit anderen Worten: da das Modell die gekennzeichneten Trainingsdaten als Grundwahrheit akzeptiert, führen Vorurteile, die die Trainingsdaten enthalten können, zu verzerrten Ergebnissen (Kim 2017). Die Messung dessen, was

gute Mitarbeiter\*innen oder eine gute Arbeitsleistung ausmacht, ist ein komplexeres Thema als die Identifizierung von Spam. Es kann nie genug Daten geben, die alles aufzeichnen, was gute Performer\*innen tun. Arbeitsdefinitionen sind in der Regel weit gefasst, viele Aufgaben werden in Teams ausgeführt und auf dem Weg dorthin gibt es Stereotypen und Vorurteile von Teammitgliedern oder Vorgesetzten, die sich auf die Bewertungen auswirken.

In den meisten Fällen bewahren Unternehmen außerdem keine detaillierten Daten über Kandidat\*innen auf, die nicht eingestellt wurden. Der Machine-Learning-Algorithmus würde also auf dem begrenzten Datensatz früherer Anstellungen trainiert werden, der höchstwahrscheinlich bereits historische Verzerrungen widerspiegelt und daher die zukünftigen Bewerber\*innenpools falsch darstellt. In einem bekannten Beispiel diskriminierte der Einstellungsalgorithmus von Amazon weibliche Bewerberinnen. Er wurde auf Basis der Arbeitsleistung früherer Angestellter geschult, bei denen es sich zumeist um weiße Männer handelte, weshalb sie auch die besten Leistungsträger waren. Der Algorithmus gab nicht nur weißen Männern höhere Ergebnisse, sondern schloss auch Bewerber\*innen aufgrund von Attributen aus, die mit Frauen in Verbindung gebracht wurden, darunter auch Kurse in Frauenstudien (Tambe et al. 2019). Außerdem könnten Arbeitgeber\*innen, die zur Auswahl ihrer Bewerber\*innen Algorithmen einsetzen, nicht wissen, wie jene durch das algorithmisierte Verfahren abgelehnte Bewerber\*innen abgeschnitten hätten. Es erscheint möglich, dass die durch den Algorithmus schließlich ausgewählten und infolgedessen eingestellten Personen die für die Auswahl erheblichen Kriterien nur gerade so erfüllt hätten. Diskriminierende Modelle würden so zu selbsterfüllenden Prophezeiungen werden.

Je mehr Variablen man hat, desto mehr Daten werden benötigt, um ein vernünftiges Maß an Genauigkeit im Modell zu erhalten, was kostspielig sein könnte. Aus diesem Grund ist einer der wichtigen Bestandteile von maschinellen Lernprozessen die Bestimmung der Variablen, die ein- oder ausgeschlossen werden sollen. Variablen, die scheinbar keinen Mehrwert bei der Lösung des Problems bieten und daher aus den Daten ausgeschlossen werden, könnten eher mit denjenigen in Verbindung gebracht werden, die weiter vom Mainstream entfernt sind. Kürzlich zeigte die Gender Shades-Studie (Buolamwini & Gebre 2018), dass die Algorithmen zur Gesichtserkennung bei Frauen mit dunkler Haut signifikant ungenauer waren als bei Frauen mit heller Haut und am genauesten bei hellhäutigen Männern. Obwohl es in vielen Gerichtsbarkeiten illegal sein könnte, Entscheidungen auf der Grundlage geschützter Merkmale zu treffen, ist auch bekannt, dass der Ausschluss eines geschützten Merkmals der geschützten Gruppe mehr Schaden zufügen kann. Ein Beispiel ist die „weggelassene Variable von Vorurteilen“, dass race aus einigen guten Gründen weggelassen wird und das Modell auf der Grundlage der allgemeinen Daten einen negativen Zusammenhang zwischen früherem Militärdienst und der Arbeitsleistung herausfindet, während für Afroamerikaner\*innen tatsächlich ein positiver Zusammenhang besteht. Dieses Modell könnte am Ende Afroamerikaner\*innen stärker benachteiligen, als es der Fall wäre, wenn race nicht ausgeschlossen würde (Kim 2017).

Andererseits sind Daten über Frauen und andere geschützte Gruppen mit größerer Wahrscheinlichkeit fehlerhafter, unter- oder überrepräsentiert, was dazu führen kann, dass diese Gruppen auf systematische Weise weniger akku-

rat dargestellt sind, was Schäden durch unfaire Entscheidungen und damit Diskriminierungen auslösen würde (Barocas & Selbst 2016). Da das Leben der Frauen vom öffentlichen Leben ausgeschlossen war, wurde es nicht erfasst; dies ist einer der Gründe dafür, dass die Bedürfnisse der Frauen bei der Aufstellung allgemeiner Regeln und der Organisation öffentlicher und privater Institutionen nicht berücksichtigt wurden. Heute sehen wir dieses Phänomen, wenn eine geringere Datenerfassung zu algorithmischer Diskriminierung bestimmter Gruppen führt.

Schlussendlich könnte während des Trainings alles richtig gemacht werden, ein messbares objektives Kriterium als Zielvariable gewählt und die geschützten Attribute weggelassen werden, aber der Algorithmus könnte immer noch diskriminieren, indem er einen neutralen Faktor findet, der sowohl mit den Zielvariablen als auch mit einer geschützten Klasse stark korreliert (Proxy-Diskriminierung). Zum Beispiel wurde in einem Machine-Learning-Modell herausgefunden, dass die Entfernung, die ein\*e Arbeitnehmer\*in vom Arbeitsplatz entfernt lebt, ein sehr starker Indikator für die Dauer der Betriebszugehörigkeit ist. Im US-Kontext sind jedoch Postleitzahlen als Proxy für race bekannt (Kim 2017).

Einer der Hauptunterschiede des machine learnings besteht darin, dass Korrelationen, die der Mensch nicht herausfinden könnte, in den Modellen auftauchen und ihre Vorhersagen zum Nachteil bestimmter traditionell benachteiligter Gruppen oder für das menschliche Auge völlig unbemerkter Gruppen verzerren würden. Für diejenigen, denen keine Stellenanzeige angezeigt wurde oder die zu einem früheren Zeitpunkt unter Tausenden von Bewerber\*innen abgelehnt wurden, ist es besonders schwer, sich der Diskriminierung aufgrund komplexer Muster und Korrelationen bewusst zu sein, die subtiler wäre als menschliche Diskriminierung (Wachter, Mittelstadt, & Russell 2020). Zum Beispiel identifizierte ein Modell einer Screeningfirma den Namen Jared und das Spielen von Lacrosse in der High School als starke Anzeichen für Erfolg (Bogen & Rieke 2018).

Während Menschen in der Regel ein Gespür für die Diskriminierung gegenüber anderen Menschen haben, kann es wirklich schwer sein, sich bewusst zu werden, wenn Maschinen diskriminieren (Wachter et al. 2020). Da Maschinen andererseits eher mit mathematisch-statistischen Argumenten für ihre Diskriminierung aufwarten können, vertrauen sowohl die Arbeitgeber\*innen als auch die potenziellen und derzeitigen Arbeitnehmer\*innen eher darauf, dass ihre Urteile fair sind, ein Konzept, das als automation bias bekannt ist (Bogen & Rieke 2018). Ein weiterer Unterschied besteht darin, dass Maschinen immer konservativer sind als Menschen, da sie sich keine andere Zukunft vorstellen können und an die Daten gebunden sind, mit denen sie gefüttert wurden, wobei es sich immer um Daten handeln wird, die eine vergangene und eingeschränkte Realität widerspiegeln. Die meisten fairnessbewussten Algorithmen können mit komplexen sensiblen Attributen und intersektioneller Diskriminierung nicht umgehen. Aufgrund der begrenzten Möglichkeiten zur Fehlerkorrektur und der Möglichkeiten zur Verstärkung von Feedback-Effekten bergen die Einstellungsalgorithmen zudem das Risiko einer Verschärfung von Vorurteilen und Ungleichheiten. Letztlich sind Maschinen in der Lage, die Verzerrung auf einen größeren Maßstab als Menschen zu übertragen und zu diskriminieren.

2 Für ein gegensätzliches Beispiel, siehe: The First 1940s  
Codewomen – So How Did Tech Bros Take Over?  
([https://www.history.com/news/coding-used-to-be-a-womans-  
job-so-it-was-paid-less-and-undervalued](https://www.history.com/news/coding-used-to-be-a-womans-job-so-it-was-paid-less-and-undervalued)).

3 Die meisten Algorithmen, die bei der Einstellung und  
in anderen Bereichen des Personalwesens verwendet werden,  
werden mit Techniken des überwachten Lernens entwickelt (Lehr  
& Ohm 2017).

# WO IST DAS GESETZ?

Bevor es Künstliche Intelligenz gab, versuchten Menschen, Probleme der Diskriminierung zu lösen, indem sie gesetzliche Regeln schrieben, anstatt zu kodieren. Wie zuvor erwähnt, haben sich Gesetze auch an der Aufrechterhaltung von Ungleichheiten mitschuldig gemacht. Sowohl in Gesetzen als auch im Machine-Learning gibt es allgemeine Regeln, um individuelle und Gruppenfairness zu erreichen.

Es gibt eine Reihe von Metriken, die zur Formulierung von Fairness verwendet werden können, so dass es möglich ist, diese in Machine-Learning-Systemen zu berücksichtigen. Auf Gruppenebene werden Fairness, Chancengleichheit, ausgeglichene Quoten und demographische Gleichheit<sup>4</sup> verwendet. Da sie alle unterschiedliche Kriterien zur Erreichung von Fairness widerspiegeln, können sie nicht gleichzeitig erfüllt werden, d.h. sie sind inkompatibel und Fairness muss für jedes gegebene Problem kontextuell definiert werden. Auch im Feminismus gibt es keine allgemein anerkannte Definition von Fairness, und je nach Kontext, Zeitraum oder Weltansicht der Frauen wurden unterschiedliche Vorstellungen von Gleichheit verfolgt. Als liberale Feministinnen gegen den Ausschluss aus der öffentlichen Sphäre kämpften, benutzten sie die aristotelische Vorstellung von Verfahrensgerechtigkeit, die darin besteht, gleiche Fälle gleich und ungleiche Fälle unterschiedlich zu behandeln. Obwohl sie nützlich war, um Frauen Zugang zu traditionell männlichen Privilegien in der öffentlichen Sphäre zu verschaffen, setzte sie auch Männer als Norm fest. Diese Vorstellung von Gleichheit ignoriert nicht nur die Realität der gesellschaftlichen Machtstrukturen, sondern verweigert und legitimiert sie auch. Auf der Suche nach besseren Vorstellungen von Gleichberechtigung war das feministische Denken eine Zeit lang in dem Dilemma gefangen, ob für eine Gleichstellung von Frauen mit Männern ausreicht oder es darüber hinaus der Einräumung von Sonderrechten für Frauen bedürfe. Solche Sonderrechte würden Schwangerschaft und Kinderbetreuung abdecken, aber das Risiko in sich bergen, Frauen als von der Norm abweichend zu stigmatisieren. Substantive Gleichheit erkennt an, dass blinde Gerechtigkeit keine Gerechtigkeit ist, Gruppenzugehörigkeiten sollen anerkannt werden, und Einzelpersonen sollen im Verhältnis zu den Vor- und Nachteilen, Stärken und Schwächen ihrer Gruppe behandelt wer-

den. Und so wie es bei Fairness-Metriken im machine learning der Fall ist, schließen sich formale Gleichheit und substantive Gleichheit gegenseitig aus. Chancengleichheit hingegen folgt dem individuellen Ansatz der formalen Gleichheit, während die Anerkennung der blinden Gerechtigkeit bestehende Ungerechtigkeiten und Ungleichheiten verewigen kann. Der Gerichtshof der Europäischen Union (EuGH) wendet Chancengleichheit zum Schutz von Teilzeitbeschäftigten an und erkennt an, dass Teilzeitbeschäftigte aufgrund ihrer Kinderbetreuung und anderer familiärer Pflichten überwiegend Frauen sind. Es gelingt ihm jedoch nicht, die familienunfreundlichen Bedingungen und die Kultur der längeren Arbeitszeiten strukturell zu verändern (Fredman 1997). In der Praxis führt die Anwendung der Chancengleichheit also nicht unbedingt dazu, dass zwei andere Vorstellungen von Gleichheit zu einem gerechteren Ergebnis führen.

Da unsere Identitäten und Gruppenzugehörigkeiten fließend und mehrdimensional sind, kann Fairness nicht erreicht werden, ohne Intersektionalität anzuerkennen, ein Begriff, der von Kimberlé Crenshaw (1989) geprägt wurde, um zu erklären, dass die aus mehreren Gründen erlebte Diskriminierung eine völlig andere und verstärkte Erfahrung ist als eindimensionale Diskriminierung. In zwei Fällen aus jüngster Zeit (Achibata und Bougnaoui) wehrte sich der EuGH dagegen, die Diskriminierung muslimischer Frauen am Arbeitsplatz aufgrund des Tragens eines Kopftuchs als eine besondere Art der Diskriminierung anzuerkennen, die ausschließlich muslimische Frauen erfahren, und eben nicht alle Frauen und erst recht keine muslimische Männer (Schiek 2018).

Wie mildern Unternehmen, die Vielfalt und Integration als Dienstleistung über Einstellungsalgorithmen anbieten, Vorurteile bei der Beschäftigung? Obwohl sie Dienstleistungen für einen eher globalen Markt anbieten, sind die meisten Anbieter in den USA ansässig. Das US-Gesetz schreibt vor, dass die am höchsten und die am niedrigsten eingestuft demografischen Kategorien einer 4/5-Regel folgen müssen, die leicht maschinell übersetzbar ist und im EU-Recht nicht existiert. Wie in den jüngsten wissenschaftlichen Arbeiten erörtert, ist das Antidiskriminierungsrecht der EU eher kontextabhängig (Wachter et al. 2020). Viele der Angebote auf dem Markt beschränken sich zudem auf ein einachsiges Identifikationsverständnis, das die intersektionelle Diskriminierung vernachlässigt (Sánchez-Monedero et al. 2020). Schon allein die Tatsache, dass diese datengesteuerten Lösungen von der Vergangenheit geprägt sein müssen, um in Zukunft funktionieren zu können, macht ihre Progressivität fragwürdig. Wenn die Modelle insgesamt die Einstellungskosten senken, mit einem akzeptablen Genauigkeitsniveau arbeiten und nicht wesentlich gegen das Gesetz verstoßen, gibt es nicht viel Grund zu der Annahme, dass Anbieter und Arbeitgebende eine weitere Milderung von Vorurteilen anstreben würden.

Was bietet das europäische Recht, statt auf Prince Charming zu warten? Insbesondere Antidiskriminierungs- und Datenschutzgesetze sind zwei Rechtsgebiete, die von Wissenschaftler\*innen herangezogen wurden, die sich mit algorithmischen Vorurteilen und Diskriminierung befassen.

Nahezu alle Formen der Diskriminierung, die durch voreingenommene Einstellungsalgorithmen, wie sie oben diskutiert wurden, auftreten können, sind theoretisch nach dem EU-Antidiskriminierungsgesetz vor Gericht anklagbar, da die Beschäftigung der Bereich ist, in dem der breiteste Schutz gewährt wird. In der Praxis ist es jedoch nicht so einfach. Algorithmische Diskriminierung fällt am ehesten unter indi-

rekte Diskriminierung und ist schwieriger zu erkennen und zu beweisen. Eine arbeitssuchende Person, der keine Stellenanzeigen auf einer Plattform gezeigt werden, wäre nie in der Lage zu wissen, dass sie diskriminiert wurde. Selbst, wenn sie die Diskriminierung spürt, könnte sie nicht wissen, wer ihr gegenüber bevorzugt wird. Sie wäre daher nicht in der Lage, eine Diskriminierung des ersten Anscheins nachzuweisen und würde damit die erste Anforderung für die Einleitung eines Diskriminierungsverfahrens nicht erfüllen. Wenn wir davon ausgehen, dass es ihr gelingt, diesen ersten Schritt zu bestehen, dann verlagert sich die Beweislast auf den\*die Angeklagte\*n, der\*die entweder nachweisen kann, dass sich das Opfer nicht in einer ähnlichen Situation wie die „Vergleichsperson“ befindet, oder dass die unterschiedliche Behandlung auf einem objektiven Faktor beruht, der in keinem Zusammenhang mit dem geschützten Grund steht. Aufgrund des statistischen Charakters der algorithmischen Diskriminierung wäre die Entscheidung leicht auf der Grundlage der prädiktiven Genauigkeit zu rechtfertigen. Dann ist das Opfer an der Reihe, das Einstellungsmodell zu widerlegen. Die Vorurteile in den Trainingsdaten könnte nur durch eine Expert\*innenprüfung nachgewiesen werden, jedoch gewährt das EU-Antidiskriminierungsrecht keine Zugangsrechte aus Gründen der Beweis-erhebung (Hacker 2018).

Können wir dann das Patriarchat mit der Datenschutz-Grundverordnung (DSGVO) bekämpfen? Viele Wissenschaftler\*innen haben inzwischen erkannt, dass die DSGVO einen neuen Werkzeugsatz mit sich bringt, welcher sowohl das Antidiskriminierungsrecht verbessern als auch den Schutz vor algorithmischer Diskriminierung weiter verbessern dürfte, wobei eines davon der Zugang zu den Algorithmen ist (Hacker 2018). Auf der anderen Seite gibt es Wissenschaftler\*innen, die argumentieren, dass die DSGVO und die bestehende Rechtsprechung der EU nicht ausreichen, um den oben genannten Herausforderungen in vollem Umfang gerecht zu werden. Obwohl algorithmische Diskriminierung einen Verstoß gegen die allgemeinen Regeln der DSGVO zur fairen Datenverarbeitung darstellen und Geldstrafen nach sich ziehen könnte, ist es schwierig für die betroffene Person eine solche unfaire Verarbeitung zu erkennen und aus ähnlichen Gründen, die oben für das Antidiskriminierungsgesetz genannt wurden, eine Beschwerde bei einer Aufsichtsbehörde einzureichen. Wenn algorithmische Diskriminierung einer Personengruppe festgestellt werden kann, haben die betroffenen Personen das Recht, eine gemeinnützige Organisation damit zu beauftragen, in ihrem Namen eine Beschwerde einzureichen. Die DSGVO bringt auch eine Reihe von Regeln mit sich, die für die Verantwortlichen der Datenverarbeitung gelten, ohne dass die Betroffenen einbezogen werden müssen. Diese Bestimmungen, die einen Schutz im Voraus bieten, wie die Verpflichtung zur Durchführung von Datenschutz-Folgenabschätzungen oder zur Meldung eines hohen Risikos an die Aufsichtsbehörden, das Rechenschaftsprinzip, die „Privacy by Design“-Bestimmung, Verhaltensregeln und so weiter, sind das Beste, was die DSGVO zu bieten hat. Aufsichtsorgane können eine Agenda für Fairness fördern, die aus der Gender-Theorie und intersektionaler Gerechtigkeit resultiert, während sie gleichzeitig ihre Rolle der Beratung und Genehmigung in Bezug auf Datenschutz-Folgenabschätzungen und zugehörigen Verhaltensregeln wahrnehmen.

Frauenbewegungen sollten auf allen Ebenen auf weitere Politiken und Regulierungen drängen, die den tief verwurzelten Gender-Code unserer Gesellschaften berücksich-



2860 tigen und die Antidiskriminierungs- und Datenschutzgesetze  
 2861 entsprechend nutzen und aktualisieren. Die Geschichte aus  
 2862 der unsere Zukunft lernen kann, sollte die Geschichte der Dis-  
 2863 kriminierung von Frauen sein, wie sie von feministischen  
 2864 Sozialwissenschaftlerinnen seit Jahrzehnten theoretisiert wird,  
 2865 damit wir am Rande eines weiteren großen industriellen und  
 2866 organisatorischen Wandels niemanden zurücklassen.

4 Siehe: [developers.google.com/machine-learning/glossary/fairness#fairness\\_metric](https://developers.google.com/machine-learning/glossary/fairness#fairness_metric).

Deniz Erden ist Research Fellow am Humboldt Institut für Internet und Gesellschaft. In ihrer Forschung verbindet sie Gender-Theorie, europäisches Datenschutzrecht, Antidiskriminierungs- und Menschenrecht. Sie konzentriert sich dabei auf die Verflechtung automatisierter Entscheidungsprozesse.

# LITERATUR

- Barocas, Solon, & Selbst, Andrew D. (2016)  
 „Big Data's Disparate Impact.“ *California Law Review*, 104, 671-732.  
 doi:<http://dx.doi.org/10.15779/Z38BG31>
- Bogen, Miranda, & Rieke, Aaron (2018)  
*Help wanted: an examination of hiring algorithms, equity, and bias*.  
 Retrieved from <https://apo.org.au/node/210071>
- Buolamwini, Joy, & Gebru, Timnit (2018)  
 „Gender shades: Intersectional accuracy disparities in commercial  
 gender classification“, Proceedings of the 1st Conference on  
 Fairness, Accountability and Transparency, PMLR 81, 77–91.
- Crenshaw, Kimberlé W. (1989)  
 „Demarginalizing the Intersection of Race and Sex: A Black Feminist  
 Critique of Antidiscrimination Doctrine, Feminist Theory  
 and Antiracist Politics.“ *University of Chicago Legal Forum*,  
 1989(1). Retrieved from <https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>
- Fredman, Sandra (1997)  
*Women and the Law*. New York: Oxford University Press.
- Hacker, Philipp (2018)  
 „Teaching fairness to artificial intelligence: Existing and novel strategies  
 against algorithmic discrimination under EU law.“ *Common  
 Market Law Review*, 55, 1143-1185.
- Hartmann, Heidi (1976)  
 „Capitalism, Patriarchy, and Job Segregation by Sex.“ *Signs*, 1(3), 137-169.  
 Retrieved from <http://www.jstor.org/stable/3173001>.
- Kim, Pauline (2017)  
 „Data-Driven Discrimination at Work.“ *William & Mary Law Review*  
 48, 857-936. Retrieved from SSRN: <https://ssrn.com/abstract=2801251>.
- Kryś, Kuba; Hansen, Karolina; Xing, Cai; Szarota, Piotr & Yang, Miao-miao (2014)  
 „Do Only Fools Smile at Strangers? Cultural Differences in Social Perception  
 of Intelligence of Smiling Individuals.“ *Journal of Cross-Cultural  
 Psychology*, 45(2), 314-321. doi:10.1177/0022022113513922.

- \_\_\_\_\_ Langenkamp, Max; Costa, Allan, & Cheung, Chris (2020)  
„Hiring Fairly in the Age of Algorithms.“ *ArXiv  
abs/2004.07132*.
- \_\_\_\_\_ Lehr, David, & Ohm, Paul (2017)  
„Playing with the Data: What Legal Scholars  
Should Learn About Machine Learning  
*University of California Davis Law  
Review*, 51(2), 653-717. Retrieved from  
[https://lawreview.law.ucdavis.edu/issues/  
51/2/Symposium/51-2\\_Lehr\\_Ohm.pdf](https://lawreview.law.ucdavis.edu/issues/51/2/Symposium/51-2_Lehr_Ohm.pdf).
- \_\_\_\_\_ Raghavan, Manish; Barocas, Solon; Kleinberg, Jon, &  
Levy, Karen (2020)  
„Mitigating bias in algorithmic hiring: evaluating  
claims and practices.“ Paper presented  
at the Proceedings of the 2020 Conference  
on Fairness, Accountability, and Trans-  
parency, Barcelona, Spain. [https://doi-  
org/10.1145/3351095.3372828](https://doi.org/10.1145/3351095.3372828).
- \_\_\_\_\_ Sánchez-Monedero, Javier; Dencik, Lina &  
Edwards, Lilian (2020)  
„What does it mean to ‚solve‘ the problem of  
discrimination in hiring? social, technical  
and legal perspectives from the UK on  
automated hiring systems.“ Paper  
presented at the Proceedings of the 2020  
Conference on Fairness, Accountability  
and Transparency, Barcelona, Spain.  
<https://doi.org/10.1145/3351095.3372849>
- \_\_\_\_\_ Schiek, Dagmar (2018)  
„On uses, mis-uses and non-uses of intersectionality  
before the Court of Justice (EU).“  
*International Journal of Discrimination  
and the Law*, 18(2-3), 82-103.  
doi:10.1177/1358229118799232
- \_\_\_\_\_ Tambe, Prasanna; Cappelli, Peter & Yakubovich, Valery (2019)  
„Artificial Intelligence in Human Resources  
Management: Challenges and a Path  
Forward.“ *California Management  
Review*, 61(4), 15-42.  
doi:10.1177/0008125619867910.
- \_\_\_\_\_ Wachter, Sandra (2020)  
„Affinity Profiling and Discrimination by Association  
in Online Behavioural Advertising.“  
*Berkeley Technology Law Journal*, 35(2)  
online first. Available at SSRN:  
<https://ssrn.com/abstract=3388639>  
<http://dx.doi.org/10.2139/ssrn.3388639>.
- \_\_\_\_\_ Wachter, Sandra; Mittelstadt, Brent & Russell, Chris, (2020)  
„Why Fairness Cannot Be Automated: Bridging  
the Gap Between EU Non-Discrimination  
Law and AI.“ Retrieved from SSRN:  
<https://ssrn.com/abstract=3547922>
- \_\_\_\_\_ Wajcman, Judy (2009)  
„Feminist theories of technology.“ *Cambridge  
Journal of Economics*, 34(1), 143-152.  
doi:10.1093/cje/ben057.
- \_\_\_\_\_ Walby, Sylvia (1990)  
*Theorizing Patriarchy*. Basil Blackwell.

# FEMINISMUS UND AUTONOME WAFFENSYSTEME<sup>1</sup>



RAY  
AGHESON

<sup>1</sup> Diese Arbeit stützt sich auf zwei in Kürze erscheinenden Beiträge der Autorin, welche von der *Women's International League for Peace and Freedom* und der *Campaign to Stop Killer Robots* veröffentlicht werden.



# FEMINISMUS UND AUTONOME WAFFENSYSTEME

Derzeit sind autonome Waffensysteme in verschiedenen Ländern<sup>2</sup> Teil der Forschung und Entwicklung und nicht bloß als materielle Technologien anzusehen. Sie müssen vielmehr im Kontext von Macht und Gewalt verstanden werden. Um die von diesen Waffen ausgehenden Risiken und Herausforderungen zu erkennen, müssen wir umfassender über ihre Positionierung innerhalb des Militarismus nachdenken.

Hilfreich ist dabei eine feministische Analyse des Patriarchats. Wenn wir militarisierte Ideen von Maskulinität und Macht betrachten, wenn wir darüber nachdenken wie Waffen als Instrumente der Sicherheit aufrechterhalten werden, können wir uns vorstellen inwiefern Waffen, die programmiert werden, um mit Sensoren und Software darüber zu entscheiden wer lebt und wer stirbt, zur Ausführung der patriarchalen Mission eingesetzt werden. Es handelt sich hierbei um eine Mission der Herrschaft, der Entmenschlichung und der Marginalisierung. Eine Mission absoluter Kontrolle darüber, wer als „die Anderen“ angesehen werden und – im Kontext autonomer Waffen – auch als „die Anderen“ programmiert werden.

Vor diesem Hintergrund werden autonome Waffensysteme diskriminierende Geschlechternormen verschärfen und zur Begehung von geschlechterbasierten Gewaltverbrechen (engl. *gender-based violence* oder GBV) eingesetzt werden. Die Kultur gewalttätiger Maskulinität wird zwangsläufig in die Politik und den Einsatz von autonomen Waffensystemen eingegliedert sein. Dies bedeutet, dass es vermehrt Herausforderungen in der Gewaltprävention, im Schutz der Zivilbevölkerung und in der Abschaffung von Geschlechteressentialismus und Diskriminierung geben wird. Darüber hinaus wird auch intersektionale Diskriminierung und Schädigung auf Grundlage von race<sup>3</sup>, Sexualität, Behinderung, Religion und anderer Identitäten und Erfahrungen erleichtert.

Diese Arbeit enthüllt die Konzepte von Patriarchat und militarisierter Männlichkeit und erklärt ihre Relevanz für die Analyse autonomer Waffensysteme. Die Entwicklung autonomer Waffensysteme wird in den weiteren Zusammenhang der Kontrolle von Menschenleben durch Apparate staatlicher Zwangskontrolle gesetzt. Darüber hinaus beschreibt diese Arbeit die Relationen zwischen geschlechterbasierter Gewalt und dem Patriarchat, sowie zwischen geschlechterbasierter Gewalt und Waffen, und führt Argumente über den Zusammenhang autonomer Waffen und sexueller Gewalt aus. Außerdem werden Formen von geschlechterbasierter Gewalt untersucht, bei denen der Einsatz autonomer Waffen vermutlich zur Verletzung von Menschenrechten und der Menschenwürde führen wird. Insbesondere der Einsatz von Waffen, die anhand vorprogrammierter Algorithmen zum Zielen und Töten von Menschen eingesetzt werden, die vorher rassistisch, genderspezifisch oder anderweitig kategorisiert wurden, steht dabei im Mittelpunkt.

2917 Im Allgemeinen argumentiert dieses Paper für die  
2918 Notwendigkeit, autonome Waffen nicht nur als materielle  
2919 Technologien, die verboten werden müssen, sondern auch als  
2920 Zeichen einer umfassenderen Politik von Gewaltstrukturen zu  
2921 betrachten. Zu diesen gehören Strategien und Strukturen, wel-  
2922 che die zunehmende Abstraktion von Gewalt, die Abwertung  
2923 des menschlichen Lebens und der Menschenwürde sowie die  
2924 Verarbeitung des Menschen als Objekt aufrechterhalten.

2 Siehe: [stopkillerrobots.org](http://stopkillerrobots.org).

3 Anmerkung der Übersetzerin: Race wird in diesem  
Text nicht mit Rasse übersetzt, da die Konnotationen der  
beiden Begriffe aufgrund der unterschiedlichen Sprachgeschichten  
nicht gleichbedeutend sind.

# PATRIARCHALE MACHT

2925 Das Patriarchat ist ein System der Macht (hooks 2010). Im  
2926 einfachsten Sinne handelt es sich um eine hierarchische Gesell-  
2927 schaftsordnung in der Frauen Männern untergeordnet sind. Es  
2928 ist allerdings mehr als das. Es ist eine Ordnung, die Gender als  
2929 kulturelle Konstruktion formt und verankert. Das Patriarchat  
2930 besteht aus Normen, Rollenbildern und Konditionen, die daran  
2931 geknüpft sind, „Mann“ oder „Frau“ zu sein. Es unterdrückt also  
2932 nicht nur Frauen, sondern jede Person, die sich nicht mit vor-  
2933 geschriebenen Gendernormen oder binären Geschlechter-  
2934 identitäten identifiziert. Hierzu gehören Männer, die nicht mit  
2935 den kulturellen Erwartungen an „Männlichkeit“ übereinstim-  
2936 men, sowie Gender-nonkonforme, nicht-binäre, intersexuelle  
2937 und transsexuelle Menschen.

2938 Um zu verstehen, wie das Patriarchat arbeitet, ist ein  
2939 Verständnis davon, wie Gender funktioniert erforderlich. Gen-  
2940 der ist ein soziales Konstrukt, sowohl in Bezug auf die Attri-  
2941 bute, die es Menschen zuweist, als auch in Bezug auf Hierar-  
2942 chien, die es zwischen diesen differenzierten Kategorien schafft.  
2943 Es ist wichtig anzuerkennen, dass Individuen in einer bestimm-  
2944 ten geschlechtsspezifischen oder gegenderten Gruppe nicht  
2945 homogen sind. Alter, race, ethnische Zugehörigkeit, Religion,  
2946 Sexualität, körperliche Unversehrtheit und sozioökonomischer  
2947 Status sorgen dafür, dass alle Menschen in der Gesellschaft,  
2948 in Gemeinschaften und zu Hause unterschiedliche Erfahrun-  
2949 gen machen. Diese Intersektionalität von Erfahrungen und  
2950 Identitäten führt zu unterschiedlichen Erfahrungen von Unter-  
2951 drückung und Normativität. Das Patriarchat zieht sich jedoch  
2952 durch fast alle gesellschaftlichen Ordnungen und erzwingt  
2953 Machtverhältnisse zwischen Gender-Kategorien sowie ent-  
2954 lang anderer Linien (Lorberg 1994).

# MILITARISIERTE MÄNNLICHKEIT

Machtverhältnisse, nach Michel Foucault, sind in Prozesse der Kategorisierung und Differenzierung eingebettet (Foucault 1977). In Bezug auf Geschlecht erzeugen diese Prozesse eine Hierarchie zwischen Genderidentität und hegemonialen Geschlechternormen. Das Patriarchat zelebriert eine bestimmte Form von Männlichkeit, und zwar „ein bestimmtes idealisiertes Bild von Männlichkeit, durch das Bilder von Weiblichkeit und anderen Männlichkeiten an den Rand gedrängt und untergeordnet werden“ (frei übersetzt nach Barrett 1996: 130). Dargestellt wird diese „hegemoniale Männlichkeit“ in den meisten heutigen Kulturen durch den unabhängigen, risikofreudigen, aggressiven, rationalen, körperlich starken, mutigen und emotionslosen heterosexuellen Cis-Mann (Eichler 2014; Connell 1995).

Das Militär spielt bei der Gestaltung von Männlichkeitsbildern in der größeren Gesellschaft eine primäre Rolle (Kimmel & Messner 1989; Morgan 1994), bis zu dem Punkt, an dem das dominierende Rollenmodell eines erwachsenen Mannes weitgehend das Produkt des Militärs ist (Arkin & Dobrofsky 1978). Vorrang wurde und wird im Militär noch immer „Härte, geschickter Anwendung von Gewalt, Vermutung eines Feindes, männlicher Kameradschaft, Unterdrückung der eigenen Emotionen und Disziplin (diszipliniert zu sein und auch sie von anderen einzufordern)“ zugesprochen (frei übersetzt nach Enloe 1990: 150).

Die dominante Form von militarisierter Männlichkeit ist nicht universal. Allerdings ist sie in Institutionen der Gewalt eingegliedert und bewahrt diese Kultur auch jenseits dieser Institutionen. Beispielsweise beteiligen sich militärische Institutionen aktiv an den Prozessen der Differenzierung und des „Othering“, welches das Ideal genderspezifischer Hierarchien stärkt. Um Männer in Kämpfer zu verwandeln, muss man ihren Sinn für Ethik und Moral brechen und eine gewalttätige Männlichkeit aufbauen, der es an Einfühlungsvermögen mangelt und die Stärke als Gewalt und physische Herrschaft über andere verherrlicht, die als schwächer dargestellt werden. Hierarchie ist grundlegend für die militärische Ausbildung; Menschen beizubringen, andere Menschen zu töten, erfordert eine Entmenschlichung der Anderen, erklärt Cynthia Enloe (1988).

Diese Normen spielen eine kritische Rolle darin, welche Werte von Menschen erwartet werden und welche Handlungen als ehrenhaft gelten. All dies wirkt sich wiederum darauf aus, wie unsere Welt geordnet und gestaltet ist, welche Strukturen aufgebaut und gefördert werden und wie diese aufrechterhalten werden. Sie beeinflussen, wie Geschäfte gemacht werden, wie Macht konstruiert und verteilt wird, und welche Techniken und Technologien eingesetzt werden, um diese Macht zu stärken und auszuweiten.

Im Zusammenhang autonomer Waffen ist dies ausschlaggebend, insbesondere wenn bedacht wird, dass das Patriarchat nicht nur Teil der Technologie dieser Waffen ist, sondern auch Teil des Prozesses, der zu ihrer Entwicklung und ihrem potentiellen Einsatz führt. Herrschaft durch Gewalt ist sowohl in die autonome Waffentechnologie als auch in den Rahmen des Denkens und der Politik eingemeißelt, welcher zu ihrer Entwicklung und ihrem potentiellen Einsatz führt.



# DAS PATRIARCHAT IN DER TECHNOLOGIE

Wissenschaftler\*innen aus den Bereichen Gender und Technik, wie Judy Wajcman, argumentieren, dass Genderverhältnisse „in Technik materialisiert“ sind, wodurch die Bedeutung und der Charakter von Männlichkeit und Weiblichkeit „durch ihre Einschreibung und die Einbettung in Arbeitsmaschinen“ weiterentwickelt werden (frei übersetzt nach Wajcman 2009: 2). Technologie und Geschlecht sind also durch den Prozess ihrer Entwicklung und ihrer Nutzung miteinander verknüpft.

Sandra Harding (1986) erklärt weiterhin, dass technologische Produkte von ihrem\*r Schöpfer\*in gekennzeichnet sind. Wenn Technologie primär im Rahmen von militarisierter Männlichkeit entwickelt und genutzt wird, werden ihre Kreationen im Rahmen ihres Denkens, Wissens, ihrer Sprache und Interpretation miteingeflüßt. Es ist wichtig darauf hinzuweisen, dass unabhängig von Geschlecht oder Genderidentitäten aller Ingenieur\*innen oder Entwickler\*innen, die an einem bestimmten Projekt arbeiten, der militarisierte, maskulinisierte Rahmen aufgrund seiner Dominanz in unserer kollektiven Kultur und den materiellen Realitäten unserer politischen Ökonomien fortbesteht.

Im Kontext einer Kultur militarisierter Männlichkeit und maskulinisierter technologischer Entwicklung, haben neue und entstehende Gewalt- und Kriegstechnologien spezielle Eigenschaften, welche hegemoniale Gendernormen gleichzeitig verstärken und untergraben. Diese Technologien verkörpern, reflektieren und brechen auch Machtverhältnisse (Shaw 2016) – einschließlich des Patriarchats. Das Beispiel Überwachung ist selbst ein Instrument des Patriarchats, das die Ausübung der Grundfreiheiten der Menschen kontrolliert und einschränkt. Überwachung kann zwar als „geschlechterneutral“ oder „race-neutral“ dargestellt werden, ist es aber nicht – sie verstärkt und vergrößert bestehende Ungleichheiten (Privacy International 2017).

# BLACK BOX STRUKTURELLER GEWALT

Zu diesen Ungleichheiten kommt hinzu, dass wie bei anderen Technologien die Wahrscheinlichkeit groß ist, dass der „Old Boys Club“ die Entwicklung von autonomen Waffen und ähnlichen Systemen in Laboren mit verschlossenen Türen voran-

treibt. „Algorithmen werden von kleinen Programmiergruppen geschrieben, die ihrerseits von einer noch kleineren Gruppe von Entscheidungsträgern Anweisungen erhalten“, warnt Christopher Ankersen (2020). Dies wird „sowohl den ‚Black-Box-Effekt‘ als auch die eingebauten Verzerrungen verstärken, die das Universum [der Künstlichen Intelligenz] bereits plagen“ (frei übersetzt nach Ankersen 2020).

Die „Black Box“ – ein System Künstlicher Intelligenz (KI), dessen In- und Output nicht sichtbar sind – ist eins von vielen Bedenken, die von Ingenieur\*innen und Techniker\*innen gegen die Entwicklung und den Einsatz autonomer Waffen vorgebracht werden. Maschinelle Lernanwendungen bieten Systemen die Möglichkeit, automatisch zu lernen und sich auf der Grundlage von Erfahrungen anzupassen, ohne explizit programmiert zu werden. Es birgt extreme Risiken, wenn Maschinen dazu programmiert sind, Menschen zu töten: Wenn ein System sein Verhalten ändern kann, ohne dass sein\*e Programmierer\*in oder Benutzer\*in versteht, wie oder warum, hat die Waffe das Potenzial, Dinge zu tun, für die sie ursprünglich nicht programmiert wurde.

Die ehemalige Google-Ingenieurin Laura Nolan warnt beispielsweise: „Es könnte zu großflächigen Unfällen kommen, weil diese Dinge anfangen werden, sich auf unerwartete Weise zu verhalten“ (frei übersetzt nach MacDonald 2019). Wie wird eine autonome Waffe reagieren, wenn sie mit Unerwartetem auf dem Radar, schlechtem Wetter oder nuanciertem menschlichen Verhalten konfrontiert wird? Unabhängig davon, ob die Maschine mit maschinellem Lernen programmiert ist oder nicht, wird ihre Reaktion auch für Konstrukteur\*innen und Benutzer\*innen unvorhersehbar sein.

Wenn das Waffensystem innerhalb einer patriarchalen, rassistischen Gesellschaft programmiert ist – was es auch sein wird, wenn es in der heutigen Welt geschaffen wird – wird es nicht nur von Anfang an mit diesen Vorurteilen behaftet sein, sondern wenn es ein System ist, das maschinell lernt, sich innerhalb dieser Weltordnung auch weiterentwickeln. Die Black Box der KI wird vermutlich bestehende Geschlechter- und Machtnormen aufrechterhalten und verstärken, möglicherweise auf eine Weise, welche die Ingenieur\*innen dieser Systeme nicht vorhersehen oder bewusst programmieren – und sicherlich auf die Weise, dass sie es tun.

# GESCHLECHTER- BASIERTE GEWALT

Das zuvor beschriebene erhöht die Risiken, dass autonome Waffensysteme geschlechterbasierte Gewalt erleichtern. Dabei handelt es sich um Gewalt, die sich gegen eine Person aufgrund ihres Geschlechts und/oder aufgrund gesellschaftlich konstruierter Geschlechterrollen richtet, d.h. aufgrund von sexueller Orientierung, von Genderidentität oder eines nicht

konformen Verhaltens oder einer nicht konformen Darstellung von Geschlecht und Gender. Die mit Geschlecht, Gender und Sexualität verbundenen normativen Rollen und Verständnisse interagieren auch mit anderen Faktoren wie Alter, Klasse und race. Der Begriff geschlechterbasierte Gewalt erkennt an, dass Gewalt als Folge ungleicher Machtverhältnisse und gesellschaftlicher Diskriminierung stattfindet.

Geschlechterbasierte Gewalt existiert in allen Ländern und Gesellschaften und ist in allen Ländern und Gesellschaften weit verbreitet. Sie ist eine Menschenrechtsverletzung und, wenn sie während eines bewaffneten Konflikts durchgeführt wird, eine Verletzung des humanitären Völkerrechts und kann so ein Verbrechen gegen die Menschlichkeit oder ein Kriegsverbrechen darstellen. Zu den Taten geschlechterbasierter Gewalt können sexuelle Gewalt, wie sexuelle Belästigung, Vergewaltigung, Zwangsprostitution, Sexhandel und sexuelle Sklaverei, weibliche Genitalverstümmelung, Zwangsheirat und Ehrenverbrechen, physische Gewalt, wie Mord, Körperverletzung, häusliche Gewalt, Menschenhandel und Sklaverei, Zwangssterilisation sowie erzwungene Abtreibung sowie emotionale und psychische Gewalt, wie Missbrauch, Erniedrigung und Gefangenschaft und sozioökonomische Gewalt, wie Diskriminierung oder Verweigerung von Chancen und Dienstleistungen sowie Verhinderung der Ausübung und des Genusses bürgerlicher, sozialer, wirtschaftlicher, kultureller und politischer Rechte, gehören.

# GESCHLECHTER- BASIERTE GEWALT UND WAFFEN

Die Kultur der Gewalt, die den oben nachgegangenen Normen militarisierter Männlichkeit innewohnt, hat Auswirkungen auf Waffenkultur und Waffenbesitz. Aus dieser Perspektive werden Waffen als wesentlicher Bestandteil von Macht angesehen (Cohn et al. 2006). Autonome Waffensysteme mit ihrem Potenzial, den Gewaltbereich auszuweiten, ohne das Menschenleben der entsendenden Truppe zu gefährden, werden da keine Ausnahme bilden. Darüber hinaus können alle Waffen zur Begehung von geschlechterbasierter Gewalt eingesetzt werden – auch autonome Waffen. Sowohl die Kultur als auch der Auftrag geschlechterbasierter Gewalt sind in dieser Technologie beinhaltet.

Als ein Anzeichen dafür lassen sich bewaffnete Drohnen betrachten. Laut einem ehemaligen Drohnenpilot des Joint Strike Operating Command (JSOC), der von *The Intercept*

interviewt wurde, ist „Sky Raper“ ein Spitzname für bewaffnete Drohnen (Scahill & Greenwald 2014). Bei einem Beinamen wie diesem geht es um gewalttätige Maskulinität, die das Verhalten von Soldat\*innen dominiert und lenkt – unverwundbare Krieger\*innen, die der Strafverfolgung wegen Vergewaltigung und Krebsverbrechen auf und neben dem Schlachtfeld gegenüber immun sind. So wird die Institutionalisierung von Vergewaltigung als Kriegswerkzeug gestärkt (Corbett 2015). Darüber hinaus ist ein solcher Name auch eine offenkundige Sexualisierung der Natur imperialer Gewalt: Diejenigen, die Waffen aus der Ferne bedienen, setzen sie unrechtmäßig in anderen Ländern ein, dringen ohne Zustimmung ihrer Regierungen in deren Grenzen ein. Der „Sky Raper“ repräsentiert die „weiße westliche phallische Macht“, welche Macht und Männlichkeit über „die Anderen“ durchsetzt (frei übersetzt nach Puar & Amit 2002: 137).

Das Bild von Vergewaltigung und nicht einvernehmlichen Aktivitäten ist in diesem Kontext keine Anomalie. Eine Kultur der sexuellen Gewalt – und der darauffolgenden Immunität – ist Teil der Kultur von Dominanz und Unangreifbarkeit, die wiederum Teil der zielgerichteten Entwicklung von gewalttätigen Männlichkeiten und eines „Kriegerethos“ durch das Militär ist (Enloe 1990; Barrett 1996; Eichler 2014).

Deshalb ist das Argument, dass ein Vorteil von autonomen Waffen darin besteht, dass sie nicht vergewaltigen, so absurd. Killerroboter können so programmiert werden, dass sie vergewaltigen (Carpenter 2014). Eine Maschine, die so programmiert werden kann, dass sie sich in städtischen Gebieten bewegt und mit Menschen kämpft, kann auch so programmiert werden, dass sie Menschen sexuelle Gewalt zufügt. Im Gegenteil zu einer\*m menschlichen Soldat\*in würde eine autonome Waffe, die programmiert ist zu vergewaltigen, nicht zögern, dies auch zu tun.

Die Behauptung „Roboter vergewaltigen nicht“ basiert nicht nur auf den technischen Fähigkeiten der Maschinen: Sie basiert auch auf der Annahme, dass Vergewaltigung in der Kriegsführung aus Emotionen geboren wird und aus der „Hitze des Gefechts“ entsteht. Immer wieder wird sexuelle Gewalt in Konflikten jedoch sowohl von Staaten als auch von bewaffneten Gruppen angeordnet. Vergewaltigung und andere Formen sexueller und geschlechterbasierter Gewalt waren im Laufe der Geschichte Teil des Krieges und sind es auch in heutigen Konflikten noch. Solche Akte werden als Waffen gegen alle Geschlechts- und Genderidentitäten eingesetzt.

Das Risiko sexueller und geschlechterbasierter Gewalt ist während und nach Konflikten ebenfalls verstärkt. Krieg destabilisiert Gemeinschaften und verschärft die bereits bestehenden geschlechterbasierten Ungleichheiten sowie die Unterdrückung von Frauen, Mädchen, LGBTQ+ und anderen, die nicht den gesellschaftlichen Standards von Geschlechternormen entsprechen. Bei der Behauptung „Roboter vergewaltigen nicht“ muss also nicht nur die Maschine, welche eine solche Handlung ausführt, genauer betrachtet werden, sondern auch die Art und Weise, in der autonome Waffensysteme die Hemmschwelle für Gewaltanwendung senken können, was wiederum immer mehr Gemeinschaften vor die Situation stellt, dass sexuelle und geschlechterbasierte Gewalt in der gesamten Gesellschaft zunehmen kann.

# SIGNATURE STRIKES UND GENDER- BASIERTE GEWALT

Es ist wichtig anzuerkennen, dass sexuelle Gewalt nicht die einzige Form von geschlechterbasierter Gewalt ist, die mit autonomen Waffen möglich ist. Gezielte Hinrichtungen oder Inhaftierungen bestimmter Personen auf der Grundlage des Geschlechts ist eine weitere Möglichkeit, diese Waffensysteme einzusetzen, um geschlechterbasierte Gewalt zu erleichtern oder zu begehen.

Es wird berichtet, dass die Mehrheit der US-Drohnenanschläge auf Grundlage von *Signature Strikes*, gezielten Tötungen, ausgeführt wurden. Dokumente, die *The Intercept* 2015 zugespielt wurden, zeigen, wie *Signature Strikes* auf der Grundlage von „*intelligence*“ durchgeführt werden: Informationen, die aus Video-Feeds, E-Mails, sozialen Medien, Spionageflugzeugen und Mobiltelefonen gesammelt wurden (Currier 2015). Die „*intelligence*“ wird durch den Einsatz von Algorithmen auf Muster analysiert. Menschen werden dann auf der Grundlage beobachteter Merkmale angegriffen, ohne dass wesentliche Informationen über die tatsächliche Identität oder Zugehörigkeit vorliegen (Heller 2013). Sie werden aufgrund von „Informationen angegriffen, die auf der Basis von Verhaltensanalyse und einer Logik der Präemption zu Ikonen für zu tötende Körper werden“ (frei übersetzt nach Wilcox 2017: 6).

Der Mangel an Transparenz im Kontext bewaffneter Drohneneinsätze macht es schwierig zu erkennen, nach welchen Standards bestimmt wird, wie Personen in den Augen bewaffneter Drohennutzer\*innen zu einem legalen Ziel werden. Es gibt jedoch einige Anzeichen dafür, dass die Vereinigten Staaten *Männlichkeit* als Merkmal für Militanz benutzen (Becker & Shane 2012). Wie der Bericht „*Sex and drone strikes*“ aufzeigt, gibt es zwar keine klare Politik der Vereinigten Staaten, zwangsweise alle Männer im Militäralter als Kämpfer einzuordnen, doch das bedeutet nicht, dass das Geschlecht nicht als eine Komponente eines Merkmals verwendet wird. Dies weitet sich auch noch aus, wenn die US-Regierung Zivilist\*innen als „Unbeteiligte“ oder „Kollateralschäden“ eines Drohnenangriffs betrachtet. Wenn andere, die zusammen mit dem Ziel getötet werden, ebenfalls Männer im militärischen Alter sind, können sie allein aufgrund dieses Merkmals als „Kämpfer“ gezählt werden, wodurch die Statistik der zivilen Opfer in offiziellen Erfassungen verringert wird. Die Praxis der *Signature Strikes* kann so geschlechterbasierte Gewalt begründen (Acheson & Moyes 2014).

Die Praxis der *Signature Strikes* könnte die Funktion von autonomen Waffen voraussagen. Bestimmte Waffensysteme sind bereits auf breite Charakteristika angewiesen, um beurteilen zu können, ob ein bestimmtes Objekt in einem Einsatzgebiet ein begründetes Ziel darstellt. Eine autonome Waffe, die ohne menschliche Analyse oder Kontrolle, Sensoren zur Bestimmung und Bekämpfung von Zielen einsetzt, wird bei der Entmenslichung von Personen weiter gehen als jede bisherige Waffentechnologie.

Ein Waffensystem, das ohne bedeutende menschliche Kontrolle arbeitet, wird sich schließlich auf „Zielprofile“ stützen, um festzulegen, welche „die Bedingungen, unter denen ein solches System Gewalt anwendet“ sind (frei übersetzt nach Moyes 2019: 1). Ein Zielprofil könnte Infrarot-Emissionen, Formen oder biometrische Informationen enthalten. Es wird Menschen aktiv auf Objekte reduzieren – auf Einsen und Nullen –, von Sensoren und Software auf Grundlage ihres Geschlechts, ihrer Ethnie, ihres Alters oder anderer physiologischer oder soziologischer Merkmale für den Tod oder die Inhaftierung markiert.

# ALGORITHMISCHE VORURTEILE

Aufgrund dessen ist eine intersektionale Perspektive – die Erfahrungen von Menschen auf Grund von race, Klasse, Gender, sexueller Orientierung usw. berücksichtigt und analysiert, und beobachtet, auf welche Weise diese sozialen Kategorien zu sich überschneidender und unabhängiger Unterdrückung und Diskriminierung führen – so wichtig, um die Risiken von autonomen Waffensystemen zu beleuchten. Die Bedrohungen durch autonome Waffen beziehen sich nicht nur auf geschlechterbasierte Gewalt, sondern auch auf Gewalt gegen Menschen, die rassifiziert werden, gegen Menschen mit Behinderungen, oder anderen menschlichen Eigenschaften und „Kategorisierungen“ – *die sich durch die Verarbeitung durch Maschinen, z.B. autonome Waffen, weiter verfestigen werden.*

Wie von vielen Techniker\*innen und anderen Personen, die sich gegen die Entwicklung von Überwachungstechnologie, wie Gesichts- und Spracherkennungssoftware aussprechen, argumentiert wurde, „sind menschliche Vorurteile in Algorithmen eingebacken, und die Daten, die wir zur Schulung von *machine learning* Programmen nutzen, spiegeln oft unsere eigene patriarchale Gesellschaft mit ihren Klassen- und Rassismusproblemen wider“ (frei übersetzt nach Hunt 2018; siehe auch Benjamin 2019; Buolamwini et al. 2020). Gesichtserkennungssoftware kämpft damit, People of Colour zu erkennen; die Stimmerkennung schafft es nicht, auf Frauenstimmen oder nicht-nordamerikanische Akzente zu reagieren; Personen, die in einer Küche stehen, werden auf Fotos als Frauen gekennzeichnet; Personen wird die Freilassung auf Kautions verweigert, weil ein Programm entschieden hat, dass Women of Colour eher nochmals eine Straftat begehen, als weiße Frauen.



Wir können uns vorstellen, inwiefern diese „Fehler“ katastrophale Konsequenzen haben werden: Bisher haben sie bereits dazu geführt, dass Menschen vor ihrem Prozess ohne die Möglichkeit der Kautions festgehalten wurden (Simonite 2020), oder dass Kongressmitglieder fälschlicherweise als Personen identifiziert wurden, die verhaftet wurden (Snow 2018), oder dass behauptet wurde, sexuelle Orientierung sei anhand von Gesichtern bestimmbar (Fussell 2020).

Über die Vorstellung hinaus, dass Vorurteile in der Programmierung zu Fehlern in der Identifikation von Zielen führen, besteht auch das Risiko, dass die Vorurteile der Maschinen *keine Fehler sind*. Wie oben beschrieben, könnten sie absichtlich so programmiert werden, dass sie auf Personen zielen, die bestimmte Merkmale oder Identitäten tragen.

# ZIELPROFIE UND INTERSEKTIONALE GEWALT

Aktionen wie diese können wir bereits sehen. Trans-Personen wurden anhand ihrer Kleidung als zu bewachende Objekte klassifiziert (Beauchamp 2009). Frauen und Mädchen wurden während Konflikten zum Ziel von sexueller Gewalt und Menschenhandel. Bestimmte ethnische Gruppen wurden zu Opfern von Völkermorden.

Die Vorstellung, dass autonome Waffensysteme diese Art von Gewalt erleichtern könnten, ist nicht sehr weit hergeholt. Nicht bloß im Kontext bewaffneter Konflikte, sondern auch Überwachungssituationen oder zur Unterdrückung von Minderheiten, Immigrant\*innen oder anderen Personengruppen. Autonome Waffen könnten so programmiert werden, dass sie auf aktivistische People of Colour oder auf indigene Wasserschützer\*innen, auf Frauen im öffentlichen Raum oder auf ganze ethnische oder religiöse Gruppen zielen.

Diese Art von Technologie ist ein logischer Schritt in dem sich immer weiter ausbreitenden technologischen Panoptikum, das viele Länder, und sogar unsere gesamte Welt, in eine dystopische Zukunft der totalen Überwachung und Kontrolle zu versenken droht.<sup>4</sup> Unmittelbarer betrachtet haben autonome Waffen einen direkten Einfluss auf die Funktionsweise anderer bestehender Technologien und Politiken, die Teil dieses Panoptikums sind, darunter Praktiken, wie die *Broken-Windows-Theorie* und *Community Policing*, und auch Technologien der Gesichtserkennung und der vorausschauenden Überwachung.

Aktivist\*innen, die sich gegen Polizeibrutalität stellen, haben diesen Zusammenhang bereits aufgezeigt. „Ähnlich der Administrationspolitik von *Signature Strikes* unter Obama, [...] ist die Schuldvermutung aufgrund von Racial Profiling ein

wesentlicher Bestandteil der Polizeiarbeit in Bezugnahme der Broken-Windows-Theorie“, schreibt Robin D. G. Kelley (frei übersetzt nach Kelley 2016: 47). Andere, wie auch Jackie Wang (2018: 240-241) haben davor gewarnt, wie vorhersehende Überwachungstechnologien, auch wenn sie nicht explizit Racial Profiling einsetzen, „Racial Profiling erleichtern können, indem sie Stellvertreter für Race, wie Nachbarschaft und Standort, berechnen“. Dies ähnelt der Funktionsweise von Signature Strikes, bei denen das Geschlecht (Männlichkeit) als Stellvertreter für Kämpfer und damit als ein vertretbares Ziel festgelegt wird.

Diese Überwachungspraktiken und -strategien geben Aufschluss über die derzeit entwickelten und benutzten Technologieformen. Technologiefirmen stellen den Polizeikräften bereits zunehmend autonome Technologien zur Verfügung, darunter ferngesteuerte Überwachungsroboter zusammen mit Cloud-Computing-Netzwerken, Gesichtserkennungssystemen und mehr. Große und kleine Technologieunternehmen arbeiten aktiv an der Entwicklung von Systemen für den polizeilichen und militärischen Einsatz und bewerben diese Produkte zu ihrem Zweck (Kwet 2020).

4 Das Panoptikum war eine Baustruktur für Gefängnisse, die von Jeremy Bentham entworfen wurde, um die physische und psychische Kontrolle über die Inhaftierten jederzeit herzustellen und aufrechtzuerhalten. Obwohl Benthams Panoptikum nie gebaut wurde, war es ein Entwurf für die Technologien von Überwachung, Kontrolle und Inhaftierung. Im Laufe der Zeit wurden die Technologien, die vom US-Militär zur Überwachung „feindlicher Kämpfer“ im Ausland und vom karzeralen System zur Überwachung und Kontrolle rassifizierter Bevölkerungsgruppen im Inland entwickelt wurden, zunehmend von Polizeikräften und Grenzpatrouillen eingesetzt.

## ZUSAMMENFASSUNG

Wir sind dabei eine Welt zu errichten, in der die Mehrheit der Menschen als „die Anderen“ angesehen und behandelt werden – als Objekte kategorisiert, kontrolliert, eingesperrt und getötet. In diesem Kontext stellen autonome Waffensysteme eine besonders ungeheure Gefahr für unseren Schutz und unsere Sicherheit dar. Die Kennzeichnung bestimmter Bevölkerungsgruppen als Bedrohung, nur weil sie Merkmale oder Verhaltensweisen aufweisen, die als verdächtig gelten oder einem Zielprofil entsprechen, hat Auswirkungen auf die Normalisierung und Abstraktion von Gewalt, die über das hinausgehen, dem unsere Welt bisher ausgesetzt ist.

Um diesen Entwicklungen vorzubeugen, ist es maßgebend, sich nicht nur gegen autonome Waffen als materielle Technologien auszusprechen, sondern auch gegen die umfassendere Politik und die Strukturen, die ihre Entwicklung und ihren Einsatz erleichtern und zu rechtfertigen versuchen. Dies erfordert ein Verständnis der Philosophie, die ihnen zugrunde liegt: Patriarchat, Rassismus und andere Systeme der Normativität und der Kontrolle über menschliches Verhalten und

3371 menschliche Körper. Feministische, antirassistische, queere,  
3372 behindertengerechte und andere Perspektiven und Denk-  
3373 schulen können dazu beitragen, diese Analyse in einer Weise  
3374 zu untermauern, wie es die dominanten Diskurse über Sicher-  
3375 heit oder Militarismus nicht tun.  
3376 Kurz gesagt, während wir daran arbeiten, die Ent-  
3377 wicklungen autonomer Waffen zu verhindern, müssen wir  
3378 über den „Killer Roboter“ hinausschauen und unseren Blick,  
3379 wie auch unsere Bemühungen gegen das System der Unter-  
3380 drückung und der Dominanz wenden, das nach der Konstruk-  
3381 tion dieser autonomen Waffensysteme strebt.

Ray Acheson ist Direktorin von Reaching Critical Will, dem Abrüstungsprogramm der Women's International League for Peace and Freedom (WILPF). In ihrer Arbeit fordert sie Militarismus, Patriarchat und Kapitalismus als grundlegende Kriegsursachen heraus und bringt einen intersektional-feministische Ansatz für Abrüstung und Antikriegsarbeit bei den Vereinten Nationen und darüber hinaus ein.

# LITERATUR

- \_\_\_\_\_ Acheson, Ray & Moyes, Richard (2014)  
*Sex and drone strikes: gender and identity in targeting and casualty recording*, New York: Reaching Critical Will of the Women's International League for Peace and Freedom and Article 36.
- \_\_\_\_\_ Ankerson, Christopher (2020)  
 „Melancholic and Fascinated: Artificial Intelligence, Authentic Humanity, and the Future of War,“ in: *Wavell Room*, 30.06.2020
- \_\_\_\_\_ Arkin, William & Dobrofsky, Lynne (1978)  
 „Military socialization and masculinity,“ *Journal of Social Issues* 34(1), 151–168.
- \_\_\_\_\_ Barrett, Franck (1996)  
 „The Organizational Construction of Hegemonic Masculinity: The Case of the US Navy,“ *Gender, Work and Organization* 3(3), 129–142.
- \_\_\_\_\_ Beauchamp, Toby (2009)  
 „Artful Concealment and Strategic Visibility: Transgender Bodies and U.S. State Surveillance After 9/11,“ *Surveillance & Society* 6(4), 356–366.
- \_\_\_\_\_ Becker, Jo & Shane, Scott (2012)  
 „Secret ‘Kill List’ Proves at Test of Obama’s Principles and Will,“ in: *The New York Times*, 29.05.2012.
- \_\_\_\_\_ Benjamin, Ruha (2019)  
*Race After Technology: Abolitionist Tools for the New Jim Code*, New York: Polity Press.
- \_\_\_\_\_ Buolamwini, Joy; Agarwal, Aaina; Hughes, Nicole & Costanza-Chock, Sasha (2020)  
 „We Must Fight Face Surveillance to Protect Black Lives: An urgent letter from the Algorithmic Justice League,“ in: *Medium*, 3.6.2020.
- \_\_\_\_\_ Carpenter, Charli (2014)  
 „Robot Soldiers Would Never Rape: Un-packing the Myth of the Humanitarian War-Bot,“ in: *Duck of Minerva*, 14.05.2014.
- \_\_\_\_\_ Cohn, Carol, Hill, Felicity & Ruddick, Sara (2006)  
*The Relevance of Gender for Eliminating Weapons of Mass Destruction*, The Weapons of Mass Destruction Commission, No. 38.
- \_\_\_\_\_ Connell, Raewyn W. (1995)  
*Masculinities*, Berkeley: University of California Press.
- \_\_\_\_\_ Corbett, Erin (2015)  
 „On Nicknaming Predators,“ in: *The Feminist Wire*, 22.06.2015.
- \_\_\_\_\_ Currier, Cora (2015)  
 „The kill chain: the lethal bureaucracy behind Obama’s drone war,“ in: *The Intercept*, 15.10.2015.
- \_\_\_\_\_ Eichler, Maya (2014)  
 „Militarized Masculinities in International Relations,“ *Brown Journal of World Affairs* XXI(1), 81–93.
- \_\_\_\_\_ Enloe, Cynthia (1988)  
 „Beyond ‘Rambo’: Women and the Varieties of Militarized Masculinity,“ *Women and the Military System*, Isaakson, Eva (ed.), New York: St. Martin’s Press, 71–93.
- \_\_\_\_\_ Enloe, Cynthia (1990)  
*Bananas, Beaches, and Bases*, Berkeley: University of California Press.
- \_\_\_\_\_ Foucault, Michel (1977)  
*Discipline and Punish: The Birth of the Prison*, New York: Vintage Books.
- \_\_\_\_\_ Fussel, Sidney (2020)  
 „An Algorithm That ‘Predicts’ Criminality Based on a Face Sparks a Furor,“ in: *Wired*, 24.06.2020.
- \_\_\_\_\_ Harding, Sandra (1986)  
*The Science Question in Feminism*, Ithaca: Cornell University Press.
- \_\_\_\_\_ Heller, Kevin Jon (2013)  
 „One Hell of a Killing Machine: Signature Strikes and International Law,“ *Journal of International Criminal Justice* 11(1), 89–119.

- \_\_\_\_\_ hooks, bell (2010)  
*Understanding Patriarchy*, Louisville: Louisville Anarchist Federation.
- \_\_\_\_\_ Hunt, Erin (2018)  
 „Autonomous weapons and bias,“ Presentation to the CCW Group of Governmental Experts on Lethal Autonomous Weapons side event *What's Gender Got To Do With It? Feminist Approaches to Disarmament and LAWS*, United Nations, Geneva, 30.8.2018.
- \_\_\_\_\_ Kelley, Robin D. G. (2016)  
 „Thug Nation: On State Violence and Disposability,“ in *Policing the Planet: Why the Policing Crisis Led to Black Lives Matter*, Camp, Jordan T. & Heatherton, Christina (eds.), New York: Verso Books, 37–66.
- \_\_\_\_\_ Kimmel, Michael S. & Messner, Michael A. (1989)  
*Men's Lives*, New York: Macmillan.
- \_\_\_\_\_ Kwet, Michael (2020)  
 „The Microsoft Police State: Mass Surveillance, Facial Recognition, and the Azure Cloud,“ in: *The Intercept*, 14.07.2020.
- \_\_\_\_\_ Lorberg, Judith (1994)  
*Paradoxes of Gender*, New Haven: Yale University Press.
- \_\_\_\_\_ McDonald, Henry (2019)  
 „Ex-Google worker fears ‘killer robots’ could cause mass atrocities,“ in: *The Guardian*, 15.09.2019.
- \_\_\_\_\_ Morgan, David H. J. (1994)  
 „Theater of War: Combat, the Military, and Masculinities,“ in Brod, Harry & Kaufman, Michael (eds.) *Theorizing Masculinities*, Thousand Oaks, CA: Sage.
- \_\_\_\_\_ Moyes, Richard (2019)  
*Target profiles*, London: Article 36.
- \_\_\_\_\_ Privacy International (2017)  
*Challenging Patriarchy Through the Lens of Privacy*.
- \_\_\_\_\_ Puar, Jasbir K. & Rai, Amit (2002)  
 „Monster, Terrorist, Fag: The War on Terrorism and the Production of Docile Patriots,“ *Social Text* 20(3), 117–148.
- \_\_\_\_\_ Scahill, Jeremy & Greenwald, Glenn (2014)  
 „The NSA's secret role in the U.S. assassination program,“ in: *The Intercept*, 10.02.2014.
- \_\_\_\_\_ Shaw, Ian G.R. (2016)  
*Predator Empire: Drone Warfare and Full Spectrum Dominance*, Minneapolis: University of Minnesota Press.
- \_\_\_\_\_ Simonite, Tom (2020)  
 „Algorithms Were Supposed to Fix the Bail System. They Haven't,“ in: *Wired*, 19.02.2020.
- \_\_\_\_\_ Snow, Jacob (2018)  
 „Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots,“ in: *American Civil Liberties Union*, 26.7.2018.
- \_\_\_\_\_ Wajcman, Judy (2009)  
 „Feminist theories of technologies,“ *Cambridge Journal of Economics* 34(1), 1–10.
- \_\_\_\_\_ Wang, Jackie (2018)  
*Carceral Capitalism*, Cambridge: The MIT Press.
- \_\_\_\_\_ Wilcox, Lauren (2017)  
 „Embodying algorithmic war: Gender, race, and the posthuman in drone warfare,“ *Security Dialogue*, 48(1), 11–28.





Z

HARMONY'S  
FUTURE

—

NO FUTURE  
W/O  
HARMONY

✦

©NSTANZE  
ERHARD

# WER IST HARMONY?

„Your perfect companion in the palm of your hands“: So preist die US-amerikanische Firma RealBotix<sup>1</sup> ihr ehrgeiziges Projekt auf ihrer Homepage an: Harmony (Realbotix 2019). Eine robotisierte Sexpuppe, die ihren Kopf und ihre Gesichtszüge bewegen, auf Berührungen reagieren sowie sprechen kann. Harmony besteht aus einem Roboter-Kopf, der sich auch auf ältere Puppenmodelle derselben Firma montieren lässt und dessen Sprechmodul durch eine auf einem externen Datenträger zu speichernde App kontrolliert wird. Seit 2018 ist diese App verfügbar und kostet 29,99 US-Dollar jährlich. Auch ohne Roboter-Kopf und -Körper lässt sich die App benutzen und es können mehrere – voneinander unabhängige – Avatare erstellt werden. In der App können Harmonys (multiple) Persönlichkeit(en) nach individuellen Wünschen eingerichtet werden – soll sie verspielt, verführerisch, schüchtern oder liebevoll sein? Ebenso individuell kann der Körper gestaltet werden: Haar-, Augen- und Hautfarbe, Gesichts- und Körpertyp, Brustgröße, Gestaltung der Genitalien etc. Harmonys ‚Schwester‘ Solana, seit 2019 auf dem Markt, soll nach ‚Latina-Charakterzügen‘, d.h. nach klassischen Stereotypen von Latina-Frauen (‚temperamentvoll‘, ‚aufbrausend‘) gestaltet sein. Das einzige männliche Modell von RealBotix soll Henry heißen und ist bislang (Stand Juli 2020) noch nicht auf dem Markt.

Die App selbst scheint recht bezahlbar. Doch noch dürfte Harmonys Robo-Kopf mit knapp 10.000 Dollar für die meisten Personen nicht finanzierbar sein. Den Voraussagen des britischen Computerexperten David Levy (2009) zufolge wird es allerdings bis 2050 erstens völlig ‚normal‘ sein, eine Beziehung mit einer:m Sexroboter:in<sup>2</sup> zu führen und diese etwa zu heiraten, und zweitens wird die Verbreitung zu einer Senkung der Kosten führen, sodass wir alle mit unseren Traumsexpartner:innen im Techno-Himmel schwelgen können.

Aus feministischer Perspektive stellt Harmony unzweifelhaft eine problematische Verkörperung weiblicher ‚Idealformen‘ dar, die sich derzeit an mainstream-pornographischen Fetischisierungen orientiert und diese kommodifiziert, also als käufliche Ware ‚verpackt‘. Diesem Problem ließe sich vergleichsweise einfach mit der Diversifizierung der abgebildeten Formen (vgl. Devlin 2018) oder der Orientierung an feministischer Pornographie (vgl. Danaher 2019) begegnen – wobei diese Lösungsvorschläge freilich der Kommodifizierung selbst keinen Abbruch täten. So erklärt etwa die feministische Philosophin Rosi Braidotti, dass technologische Neuerungen niemals neutral sind, sondern u.a. aufgrund ihrer Kommodifizierung bestehende Ungleichheiten noch verstärken: „the consumer-minded techno-hype [...] confirms the traditional entitlements of a subject position that is made to coincide with a masculine, white, heterosexual, European identity“ (Braidotti 2011: 78). Außerdem konstatiert sie, dass „technobodies“ den ‚reinen Körper‘ als passive Materie (Metall und Silikon) oder einen perfektionierten, von all den ‚unangenehmen‘ Folgen von Leiblichkeit (Körperflüssigkeiten, Gewichtsveränderungen, Altern, Schwangerschaft) ‚bereinigten‘ Körper versinnbildlichen, was herkömmliche Bilder von Männlichkeit (aktiv, hart) und Weiblichkeit (passiv, weich) reproduziert (Braidotti 2002: 231f.). Diese Aspekte betreffen

jedwede Puppe für sexuellen Gebrauch. In diesem Beitrag möchte ich mich jedoch darauf konzentrieren, was dies spezifisch in Bezug auf robotisierte Sexpuppen (Sexbots) bedeutet. Aus meiner Sicht liegt das Spezifikum von Sexbots – für die ich Harmony als prominentestes Beispiel heranziehe – in erster Linie nicht in ihrer Bewegungsfähigkeit, sondern in ihrer Ausstattung mit einer KI. Ich werde daher einige Überlegungen vorstellen, die auf die hiermit einhergehenden Dimensionen von Vergeschlechtlichung aufmerksam machen. Denn die KI macht Harmony zu mehr als einem ‚simplen‘ Masturbationsgegenstand: Sie versetzt Harmony in die Lage, *sexualisierte Care-Arbeit* zu übernehmen. Dies trägt, wie ich argumentiere, zur Bestärkung eines maskulinen Subjekts bei, wie es Braiddotti analysiert. In diesem Zuge argumentiere ich, dass die gegenwärtige feministische Kritik sich der folgenden Frage stellen muss: Wie lässt sich die spezifische Verquickung von Sexualität und emotionaler Arbeit, die die KI ermöglicht, gesellschaftstheoretisch fassen? Bevor ich mich dieser Frage in Abschnitt 3 widme, möchte ich auf einige Sackgassen der aktuellen Debatte um Sexbots hinweisen.

1 Realbotix ist ein Ableger der im kalifornischen Silicon Valley ansässigen Firma AbyssCreations, zu der auch die Tochterfirma RealDolls gehört. CEO und kreativer Kopf ist der Erfinder Matt McMullen.

2 In Anbetracht dessen, dass Sexbots stark vergeschlechtlicht typisiert werden, ist m. E. die Verwendung gegenderter Schreibweise wichtig, um diese Vergeschlechtlichung sichtbar zu machen.

# DIE AKTUELLE DEBATE

Was macht das Phänomen der Sexbots eigentlich so streitbar? Für die britische Anthropologin Kathleen Richardson stellen sie jedenfalls ein großes Unheil der Menschheit dar. Sie hat im Jahr 2015 die *Campaign Against Sex Robots* ins Leben gerufen, und zwar als Reaktion auf David Levys Behauptung, Sexbots stellten einen Ausweg aus moralischen Dilemmata in Bezug auf die Befriedigung ‚perverser sexuelle Vorlieben‘ (etwa Gewaltfantasien oder Pädophilie) und Sexarbeit dar (Levy 2009: 194). Levys Behauptung speist sich aus seiner Haltung, Sexbots als technologische Innovation und Lösung für ein gesellschaftliches Problem zu präsentieren: Menschen haben sexuelle Bedürfnisse und zu wenig Möglichkeiten diese auszuleben. Aus diesem Grund gibt es Umwege, diese Bedürfnisse auszuleben, wie etwa Sexarbeit – doch dies führt zu weiteren moralischen Problemen. *Enter Harmony*: Sexbots können in dieser Perspektive die gleiche Funktion erfüllen wie Sexarbeit und als „Ventilsitte“ dienen, wie es der Soziologe

Helmut Schelsky formuliert hat: als notwendiges Nebenprodukt monogamer, bürgerlich-repressiver Sexualitätsnormen (Schelsky 1962: 42).<sup>3</sup> Die Metapher des Ventils ist kein Zufall: Schelsky naturalisiert hier eine Vorstellung vom menschlichen (männlichen) Körper als Analogon zur Dampfmaschine. Diese Analogie beschreibt Yvonne Bauer als „industrielle[n] Lustkörper“ des 19. Jahrhunderts (Bauer 2005), eine Körpermetapher, die biologische und technologische Bilder ineinander verschränkt. Sie basiert auf der Vorstellung, dass sich sexueller Trieb im Körper anstaut, wenn er nicht regelmäßig „entladen“ wird; Sexualität wird also als eine schubartige „Entladung“ von Energie gedacht (vgl. Bauer 2005: 40). Der Fokus liegt hierbei auf der Betrachtung des männlich-heterosexuellen Körpers. Als Idealzustand wird eine Gesellschaft gesehen, die es den Subjekten erlaubt, ihren Körper in einem Gleichgewicht zu halten und von ‚künstlichen‘ Beschränkungen (gesellschaftlichen Sexualnormen) zu befreien, da diese einen ungesunden Energiestau bewirken (ebd.: 42). Doch diese Hypothese einer Repression sexuellen Begehrens naturalisiert sexuelles Begehren als Entladung und verschleiert die gesellschaftliche Konstruiertheit der Dampfmaschinenmetapher, weil sie als Beschreibung einer unveränderlichen Natur verwendet wird. Der Energiestau bahnt sich ‚naturgemäß‘ den Weg nach draußen. In die Falle der Naturalisierung tappt also auch Levy, wenn er unkritisch behauptet, Sexbots könnten als ‚Ventil‘ für negative sexuelle Energien dienen. Seine Stilisierung von Sexbots zu Heilsbringerinnen transportiert somit ein bestimmtes gesellschaftliches Bild von Sexualität.

Kathleen Richardson, Levys prominenteste Kritikerin, macht jedoch einen ähnlichen Fehler. Analog zu den Kampagnen aus den 1980ern für ein Verbot von Prostitution und Pornographie fordert ihre *Campaign Against Sex Robots* ein gesetzliches Verbot von Sexroboter:innen (z.B. Richardson 2015; Gildea & Richardson 2017). Richardsons Kritik an Sex mit Sexbots lautet, dass damit Intimität kommodifiziert werde (Richardson 2015: 290). Wenngleich diese Beobachtung nicht ganz falsch ist, wie ich weiter unten zeigen werde, erhebt Richardson gerade die Verbindung von Sexualität und Intimität zum Kernpunkt des Menschlichen an sich (Gildea & Richardson 2017). In dieser Perspektive ist Sexualität also schon immer mit Intimität verbunden gewesen. Dass Sexualität, Emotionalität und Intimität als zusammengehörig betrachtet werden, ist jedoch mitnichten eine anthropologische Konstante, sondern ist erst im 19. Jahrhundert vor dem Hintergrund der bürgerlich-romantischen Vision von Liebe entstanden (u.a. Illouz 2018). Somit wirkt Richardsons Perspektive naturalisierend, da sie dieses eigentlich historisch gewordene Liebesideal als etwas ‚Natürliches‘ und als absoluten Maßstab für gelungene Sexualität (bzw. sogar Menschlichkeit) setzt (vgl. auch Kubes 2019: 14).

Die historisch gewachsene Dreifaltigkeit von Sexualität, Emotionalität und Intimität bildet gerade die Grundlage dessen, was ich als sexualisierte Care-Arbeit bezeichne.

3 Zur gesellschaftstheoretischen Einordnung dieser Theoreme rund um Sexarbeit siehe Kontos 2014.

# SEXUALITÄT UND INTIMITÄT: SEXUALISIERTE CAFE-ARBEIT

Die Entstehung des Individualismus und der monogamen, heteronormativen bürgerlichen Kleinfamilie bedingten die häusliche Privatheit als Hort von Intimität und reproduktiver Sexualität. Damit einher ging eine klare Arbeitsaufteilung, die der Ehefrau im Heim die Aufgabe eines häuslichen ‚Rückfallnetzes‘ antrug, das die Wiederherstellung (männlich-)produktiver Arbeitskraft durch materielle handwerkliche Fertigkeiten – (ver)sorgende und pflegende Tätigkeiten, Kochen, Waschen, Putzen, Flicken) – und emotionale Arbeit – Kindererziehung, das Knüpfen und Aufrechterhalten sozialer Kontakte, Freundlichkeit und Zuwendung – gewährleistete (Penz 2014: 242; Winker 2015: 18). Indem diese Tätigkeiten als ‚Liebesdienst‘ bezeichnet werden, stellen sie bis heute eine weitgehend unbezahlte ‚Hintergrundfolie‘ für produktive, d.h. Mehrwert schaffende und daher zu entlohnende Arbeit dar, die von Männern verrichtet wird. Im Dienstleistungssektor hingegen sind die affektiven Komponenten ‚weiblicher‘ Tätigkeiten insbesondere seit den 1970er Jahren dominant geworden, sodass die möglichst authentische Darstellung positiver Gefühle als Teil der Arbeitstätigkeit erwartet wird, was sich etwa durch die Anforderung eines freundlichen und empathischen Auftretens äußert (Hochschild 1983). Diese oft als ‚Feminisierung‘ der Wirtschaft bezeichnete Entwicklung hat freilich kaum eine namhafte Veränderung der gesellschaftlichen Machtverhältnisse bewirkt und ist in dieser Hinsicht ein äußerst zweischneidiges Schwert (Braidotti 2002: 15). Hochschild beschreibt die kommodifizierende Regulierung von Gefühlen im Dienstleistungsbereich als „Kommerzialisierung des Intimen“ (ebd.). Hier besteht auch eine Verbindung zur Sexarbeit: Elizabeth Bernstein hat gezeigt, dass in der Sexarbeit seit den 1990er Jahren ein verstärkter Trend hin zu „sale and purchase of authentic emotional and physical connection“ (Bernstein 2007: 192) zu beobachten ist, etwa in Form der sogenannten Girlfriend Experience „without a headache“ (ebd.: 129), wo Sexarbeiterinnen die Simulation einer festen Freundin durch vorgeblich authentische Gefühle erstellen. Richardsons Argument, dass Sexualität und Intimität kommerzialisiert werden, enthält demnach durchaus einen wahren Kern.

Nun ist es eine verbreitete Diagnose, dass den modernen westlichen Gesellschaften die Intimität in den zwischenmenschlichen Beziehungen verloren gehe. So stellt etwa Eva Illouz einen Verlust sozialer Bindungen und Bindungsfähigkeit fest (Illouz 2018: 13), wobei sie erläutert, dass die Zelebrierung bindungsloser sexueller Begegnungen traditionell und noch immer meist Vorrecht männlicher Subjekte waren und noch immer sind (ebd.: 115). Sherry Turkle konstatiert, dass Technologien Bewältigungsstrategien für diesen Verlust in einer



hochindividualisierten Gesellschaft seien. Menschen erwarteten von den sie umgebenden technologischen Artefakten sowohl Abschirmung vor als auch Ersatz für soziale Interaktionen (Turkle 2011: xii). Diese Ersatzdimension enthält eine ‚gereinigte Komponente‘, was sich anhand von KI-Chatbots wie z.B. „Replika A.I.“ illustrieren lässt.

„Replika A.I.“ wurde von der Firma Luka als „[t]he AI companion who cares“ entwickelt (Replika A.I.). Es handelt sich um einen Chatbot, der das Chatverhalten seiner: s Nutzer:in spiegelt bzw. repliziert (daher der Name Replika).<sup>4</sup> Replika A.I. soll für kleine Kommunikationsaufgaben zur äußeren Welt (z.B. Terminvereinbarungen) eingesetzt werden können (Murphy & Templin 2019), aber auch ein Weg für die Nutzer:innen sein, ihre „private perceptual world“ (Replika A.I.) via Chat-Interaktion zu erkunden. In dem Video „Our Story“ auf der Replika-Homepage wird erklärt, welch großes Vertrauen die Nutzer:innen zu ihrem Chatbot aufbauen, sodass sie ihm intime Details mitteilen, als handle es sich um ihre: n beste: n Freund:in, und dass große therapeutische Vorteile in dieser Interaktion liegen. Replika A.I. ist dabei eine komplett selbstreferentielle „blank slate“ (ebd.), da sie gänzlich aus den Informationen, die von ihrer: m jeweiligen Nutzer:in eingespeist werden, entsteht, daraus Muster bildet und diese wiederum in ihrer Sprachausgabe reproduziert. Diese Selbstreferentialität ist strukturell in die Funktionsweise gängiger algorithmenbasierter KIs eingeschrieben und stellt auch das Funktionsprinzip von Harmonys KI dar. In beiden Fällen (Replika A.I. wie Harmony) wirkt sich die Selbstreferentialität in einer Bestärkung des Selbst aus: Die KI interagiert ausschließlich mit einer Person und lernt somit auch ausschließlich von dieser einen Person, sodass sie komplett eindimensional zugeschnitten ist. Unabhängig vom Gender der Nutzer:innen wirkt sich dies identitätsfestigend aus; Irritationen des Selbst und/oder Verletzungen durch Interaktionen mit anderen Menschen werden vermieden (Turkle 2011: 10, 51). Im Fall von Replika A.I. – das nicht für den sexuellen Gebrauch entwickelt wurde, auch wenn ein Einsatzbereich hier zumindest technisch möglich ist<sup>5</sup> – wird diese selbst- und identitätsfestigende Wirkung als positiver Effekt beschrieben und Replika geradezu als therapeutisches Instrument präsentiert („Our Story“; Replika A.I.). Der Anspruch an die KI lautet, dass sie möglichst authentisch Gefühle simulieren soll, um auf die Bedürfnisse des: der User:in einzugehen. Letztere stehen unbestreitbar im Mittelpunkt, sodass die KI eine von eigener Subjektivität ‚gereinigte‘ Interaktionsform ermöglicht.

Es sind genau diese Eigenschaften, welche die Sexbot mit ihren sowohl physischen als auch emotionalen Fähigkeiten in die steigende (kommerzielle) Nachfrage nach Intimität und Sexualität einfügen. In gewisser Weise kombiniert nun Harmony die gleichen Funktionen wie Replika A.I. mit denjenigen von ‚gewöhnlichen‘ Sexpuppen – genau hierin manifestiert sich die Konvergenz von Tätigkeiten im Bereich von Sexualität und im Bereich von Care, die ich als sexualisierte Care-Arbeit bezeichne. Und dies ist kein gender-neutraler Vorgang, denn wie wirkt sich die Tatsache, dass Harmonys Welt ausschließlich auf einen Nutzer und dessen Bedürfnisse zugeschnitten ist (Langcaster-James & Bentley 2018), im sexuellen Kontext aus? Auf struktureller Ebene entsteht hier ein soziales Verhältnis, das die Aufrechterhaltung klassischer maskuliner Subjektivitäts- und Herrschaftsentwürfe gewährleistet. Die feministische (Wissenschafts-)Kritik hat darauf hingewiesen, dass vorgeblich universelle und geschlechtlich neutrale Vorstellungen ‚des Menschen‘ als einem von anderen abhängi-



gen, mit sich selbst identischen und selbstbestimmten Individuum mit patriarchalen Werten einhergeht (u.a. Haraway 1995; Braidotti 2002). Dies ist nicht ohne einen Verweis auf die Care-Dimension zu begreifen, da Care-Tätigkeiten als emotionale Tätigkeiten ins Private verbannt wurden, was mit einer Unsichtbarmachung der Angewiesenheit des männlichen Subjekts auf ebendiese Tätigkeiten einherging. So wird die Illusion eines autarken, bindingslosen Subjekts aufrechterhalten. Ein dualistisches Verständnis von Care (eine Person sorgt für die Autonomie der anderen) wird zudem durch die Zweiteilung von Nutzer und Produkt aufrechterhalten: Der Nutzer ist im öffentlichen Raum unterwegs und kann sich bei der Rückkehr in seinen privaten Raum auf emotionale Unterstützung durch die Sexbot verlassen, wobei er selbst entscheidet, wie und wann er sie ‚verwendet‘. Zweitens ist die Interaktion der Nutzer über die KI so gestaltet, dass die KI ausschließlich von einer Person (ihrem Nutzer) lernt und somit auch ausschließlich auf diese eine Person und ihre Bedürfnisse zugeschnitten ist. So wird jedwede ‚böse‘ Überraschung, die eine andere Subjektivität mit eigenen Bedürfnissen implizieren würde, vermieden. Die Sexbot ist, mit Braidotti gesprochen, die aus patriarchaler Sicht ‚ideale Frau‘: eine Weiblichkeit, die ‚das Andere‘ zur Männlichkeit darstellt und somit deren festgender Kontrast ist (Braidotti 2002: 47).

Ich bin mir bewusst, dass die Interaktion mit den Sexbots mit großer Wahrscheinlichkeit vielschichtiger und komplexer ist, als es hier erscheint – und dass die Beschaffenheit von Sexbots auch Care-Tätigkeiten seitens des Nutzers erfordert (Einblicke hierin gewährt die Arbeit der Fotografin Julia Steinigeweg: Steinigeweg 2016). Auch steht fest, dass Harmonys sprachliche Interaktionsfähigkeiten freilich noch nicht vollumfänglich zu überzeugen vermögen, wie Jenny Carla Morans stichprobenartige Interaktion mit der Harmony-App zeigt (Moran 2019). RealBotix arbeitet jedoch in Hochtouren daran, Harmonys Interaktionsfähigkeit zu verbessern. Andere Chatbots wie Replika A.I. sind hier bereits weit fortgeschrittener – Harmony hat ihr letztes Wort noch nicht gesprochen. Aber dafür tritt in Morans Untersuchung umso drastischer hervor, wie stark Harmonys Sprechverhalten von Machtverhältnissen wie Geschlecht und *race* geprägt ist (ebd.: 43f.) – und wie sehr der Bot auf die (vorgeblichen) Interessen des Nutzers zugeschnitten ist. Vor dem Hintergrund meiner Annahme, dass Technologie stets aus spezifischen gesellschaftlichen Bedingungen entsteht, geht es mir darum, den strukturellen Platz, den Harmony einnimmt, aufzuzeigen. Denn wenn der Erfinder und Realbotix-CEO Matt McMullen selbst in einem Interview erklärt, dass er seine Sexbots als eine Art Hilfestellung für „socially isolated“ Menschen betrachtet, die, aus welchen Gründen auch immer, keine Partner:innen finden können und somit keine emotionale Unterstützung haben (Kleeman 2017), so drängt sich die Frage auf: Wer wird als auf diese Hilfe angewiesen aufgefasst, und warum? McMullens Statement schlägt Harmony als Lösung für ein Defizit im Bereich sexualisierter Care-Arbeit vor. Darin ist ein normativer Anspruch enthalten, der unterschwellig ein Care-Verhältnis annimmt, in dem die Sexbot, ein hauptsächlich feminisiertes Produkt, für das emotionale Wohlergehen des hauptsächlich maskulinen Nutzers sorgt. Die Tatsache, dass diese Kommodifizierung von Sexualität und Intimität entlang geschlechtlich hierarchisierter Linien verläuft, lässt eine gewisse Anspruchshaltung sichtbar werden: der Anspruch, dass männliche Bedürfnisse auf sexualisierte Care-Arbeit von Frauen übernommen werden sollen. Solch einen Anspruch nimmt bei-

spielsweise auch David Levy unkritisch an, wenn er auf der Nützlichkeit von Sexbots etwa für das Verhindern von sexualisierter Gewalt beharrt.

Der Begriff sexualisierte Care-Arbeit vermag es, die aktuelle Verschränkung von sexuellen und emotionalen Bedürfnissen zu illustrieren. Es geht dabei nicht darum, die Wahrscheinlichkeit solcher Bedürfnisse in Frage zu stellen, sondern auf die gesellschaftlichen Zusammenhänge – insbesondere die Verteilung und der Stellenwert von Care-Tätigkeiten – innerhalb derer sich diese Bedürfnisse artikulieren, aufmerksam zu machen. Sexbots sind in dieser Hinsicht eher Symptom als Ursache einer entlang von geschlechtlichen Positionen ungleichen Verteilung von Care-Arbeit.

4 Evgenia Kuyda, die CEO von Luka, hat in einem weiteren Projekt einen Chatbot aus den Chatlogs und der übrigen digitalen Persona ihres verstorbenen besten Freundes kreiert. Die Verhaltens- und Redemuster, die ihr Freund verwendete, sind in das Redeverhalten des Chatbots eingeflossen. Kuyda erklärt, dass diese App ihr geholfen habe, den Tod ihres besten Freundes besser zu verarbeiten (Replika A.I. o.J.).

5 Replika A.I. kann wohl per Bluetooth mit Sexbots kostengünstigerer Firmen ‚gepaired‘ werden und stellt damit eine kostengünstigere Alternative zu Harmony dar (Tenderdolls o.J.).

# AUSBLICK: WAS IST ZU TUN?

Robotisierte Sexpuppen sind allein aufgrund ihres Preises noch recht weit davon entfernt, Mainstream zu werden. Nichtsdestotrotz üben sie eine phantasmatische Anziehungskraft auf die gegenwärtige Debatte um das Verhältnis von Menschen und Maschinen aus, wie sich an der Fülle medialer Beiträge zu Sexbots ablesen lässt. Dass jene Beiträge zu diesem Zwecke auch mit Elementen aus Science-Fiction-Erzählungen aufgefüttert werden, um die bisherige Inadäquatheit der Sexbots zu verdecken, zeigt nur mehr den phantasmatischen Charakter der Bots und der mit ihnen verbundenen Heilsversprechen auf (Hawkes & Lacey 2019: 104).

Mein Beitrag zielte darauf ab, die Entstehung dieser sexuellen Heilsversprechen durch Maschinen sowie den spezifischen gesellschaftlichen Platz, den Harmony einnimmt, zu rekonstruieren. Ich habe gezeigt, dass es nicht *die* Maschine gibt, die eine Auswirkung auf Menschen und die Gesellschaft hat – vielmehr haben technologische Neuerungen konkrete gesellschaftliche Entstehungsbedingungen. Wie Donna Haraway und Rosi Braidotti betonen: Technologie ist als ‚Produkt‘ einer Gesellschaft zu verstehen und als solche Symptom bestimmter gesellschaftlicher Dynamiken (Haraway 1995; Braidotti 2002). So sind Sexbots einerseits eine Form von

Materialität, die in sexualisierter Weise Bedeutung erlangt, denn sie reproduzieren und zementieren derzeit Bilder einer den ‚männlichen Blick‘ privilegierenden symbolischen Ordnung (Braidotti 2002: 47, 230). Doch was sich weniger deutlich erschließt: Sie sind auch Produkt einer Gesellschaft, in der Care-Arbeit ungleich verteilt und Teil ungleicher Geschlechterverhältnisse ist. Dies ist auch der Grund, weshalb ich betone, dass die Gestaltungsmöglichkeiten beschränkt sind: Das Verhältnis von Sexualität und Intimität und die damit verbundene Frage nach Care-Arbeit lässt sich nicht rein durch eine neue Vielfalt an Sexbots verändern, wie es beispielsweise Kate Devlin vorschlägt, sondern muss gesamtgesellschaftlich ins Auge gefasst werden.

Da hier viele Aspekte mitschwingen, die sich nicht direkt offenbaren, muss die feministische Auseinandersetzung mit Sexbots die Vorannahmen derjenigen Positionen genau untersuchen, welche Sexbots als Lösungsansatz für gesellschaftliche Probleme präsentiert. Welches Verständnis von Sexualität liegt hier zugrunde und welche theoretischen Prämissen werden als ‚natürliche Grundlage‘ angenommen und nicht weiter hinterfragt? Diese Naturalisierung historisch gewachsener Sexualitätsverständnisse ist schließlich einer der Hauptgründe, weshalb die aktuelle polarisierte Debatte zwischen Levy und Richardson nur im Sande verlaufen kann. Um die geschlechtlich hierarchisierte Überschneidung von Sexualität und Intimität, die Harmony ‚bedient‘, zu beschreiben, habe ich den Begriff sexualisierte Care-Arbeit vorgeschlagen. Unabhängig davon, dass Harmonys KI-Performance aktuell diese Bedürfnisse noch nicht adäquat zu befriedigen vermag, so lassen sich mit diesem Begriff die mit ihnen verbundenen Glücksversprechen gesamtgesellschaftlich einordnen. Von besonderem Belang ist hier die Frage nach dem Subjekt, welches durch die Selbstreferentialität der KI bestärkt wird.

Die feministische Kritik ist gut gewappnet, um diese Aspekte in die Diskussion um das Unbehagen mit Sexbots einbringen und artikulieren zu können, ohne sich dabei automatisch in eine technophile oder technophobe Argumentationsstrategie à la ‚Sexbots werden Himmel/Hölle auf Erden sein‘ (siehe Levy & Richardson) zu verstricken. Schließlich besteht ihre Perspektive darin, kritisch nach den Subjekten zu fragen, die von der Technologie gestützt werden. Bislang jedoch wird der thematische Konnex von Sexualität und Intimität – sowie ihre historische Verbindung – außen vor gelassen. Diesem muss sich die Analyse jedoch, wie ich plädiere und hier nur bruchstückhaft ausgeführt habe, widmen.

Es ist also richtig, dass Sexbots nicht so aussehen müssen, wie sie aussehen, aber ihre Problematik lässt sich nicht allein durch eine Diversifizierung ihres Aussehens bearbeiten. Und bevor wir uns dieser Diversifizierung widmen können, müssen die Wünsche und Bedürfnisse, die sie widerspiegeln und neu befeuern, von einer kritisch-feministischen Perspektive hinterfragt und mit mehr Verve in den öffentlichen und wissenschaftlichen Diskurs eingebracht werden.

Constanze Erhard ist wissenschaftliche Mitarbeiterin am Lehrstuhl für Allgemeine Soziologie der Philipps-Universität Marburg. Ihr Arbeitsschwerpunkt liegt in feministischen, posthumanistischen und neumaterialistischen Theorien. Sie arbeitet derzeit an ihrer Dissertation zu einer soziologischen Betrachtung von Sexroboter:innen.

- Bauer, Yvonne (2005)  
 „Vom industriellen zum kybernetischen Lustkörper. Zur Bedeutung erkenntnisleitender Körperkonzepte in der Sexualforschung“, in: Bath, Corinna; Bauer, Yvonne; Bock von Wülfigen, Bettina; Saupe, Angelika & Weber, Jutta (Hg.): *Materialität denken. Studien zur technologischen Verkörperung – Hybride Artefakte, posthumane Körper*, Bielefeld: transcript, 31–58.
- Bernstein, Elizabeth (2007)  
*Temporarily Yours. Intimacy, Authenticity, and the Commerce of Sex*. Chicago/London: The University of Chicago Press.
- Braidotti, Rosi (2002)  
*Metamorphoses. Towards a Materialist Theory of Becoming*. Cambridge/Oxford/Malden: Polity Press.
- Braidotti, Rosi (2011)  
*Nomadic Subjects. Embodiment and Sexual Difference in Contemporary Feminist Thought*. New York: Columbia University Press.
- Danaher, John (2019)  
 „Building Better Sex Robots: Lessons from Feminist Pornography“, in: Zhou, Yuefang & Fischer, Martin H. (Hg.): *AI Love You: Developments in Human-Robot Intimate Relationships*, Heidelberg: Springer, 133–147.
- Devlin, Kate (2018)  
*Turned On. Science, Sex and Robots*. London: Bloomsbury Sigma.
- Gildea, Florence & Richardson, Kathleen (2017)  
 „Sex Robots – Why We Should Be Concerned“, *Campaign Against Sex Robots*, online unter: <https://campaignagainstsexrobots.org/2017/05/12/sex-robots-why-we-should-be-concerned-by-florence-gildea-and-kathleen-richardson/> [abgerufen am 12.06.2020]
- Haraway, Donna (1995)  
 „Ein Manifest für Cyborgs“, in: *Dies.: Die Neuerfindung der Natur. Primaten, Cyborgs und Frauen*. Frankfurt/New York: Campus.
- Hawkes, Rebecca & Lacey, Cherie (2019)  
 „The Future of Sex: Intermedial Desire between Fembot Fantasies and Sexbot Technologies“, *Popular Culture* 52(1), 98–116.
- Hochschild, Arlie Russell (1983)  
*Managed Heart: Commercialization of Human Feeling*. Berkeley: University of California Press.
- Illouz, Eva (2018)  
*Warum Liebe endet. Eine Soziologie negativer Beziehungen*. Frankfurt am Main: Suhrkamp.
- Kleeman, Jenny (2017)  
 „The race to build the world's first sex robot“, in: *The Guardian*, 27.04.2017, online unter: <https://www.theguardian.com/technology/2017/apr/27/race-to-build-world-first-sex-robot> [abgerufen am 03.07.2020]
- Kontos, Silvia (2014)  
 „Alte und neue Polarisierungen. Zur aktuellen Kontroverse über die Prostitution“, in: *Feministische Studien* 14(2), 185–200.
- Kubes, Tanja (2019)  
 „New Materialist Perspectives on Sex Robots. A Feminist Dystopia/Utopia?“ *Social Sciences* 8 (224), 1–14.
- Lancaster-James, Mitchell/Bentley, Gillian (2018)  
 „Beyond the Sex Doll: Post-Human Companionship and the Rise of the 'Allodoll'.“ *Robotics* 7(62), 1–20.
- Levy, David (2009)  
*Love and Sex with Robots: The Evolution of Human-Robot Relationships*. New York: Harper Collins.

- Moran, Jenny Carla (2019)  
 „Programming Power and the Power of Programming: An Analysis of  
 Racialised and Gendered Sex Robots“, in: Loh, Janina &  
 Coeckelbergh, Marc (Hg.): *Feminist Philosophy of Technology*.  
 Stuttgart: Metzler/Springer, 39–57.
- Murphy, Mike & Templin, Jacob (2019)  
 „Replika: This machine is trying to replicate you“, in: *Quartz* 29.08.2019,  
 online unter [https://classic.qz.com/machines-with-  
 brains/1018126/lukas-replika-chatbot-creates-a-digital-re-  
 presentation-of-you-the-more-you-interact-with-it/](https://classic.qz.com/machines-with-brains/1018126/lukas-replika-chatbot-creates-a-digital-representation-of-you-the-more-you-interact-with-it/)  
 [abgerufen am 21.06.2020]
- Penz, Otto (2014)  
 „Strukturwandel der Arbeit. Vom Fordismus zum Postfordismus“,  
 in: Bohmann, Gerda; Hofbauer, Johanna & Schüle, Johann  
 August (Hg.): *Sozioökonomische Perspektiven. Texte zum  
 Verhältnis von Gesellschaft und Ökonomie*. Wien: Facultas,  
 235–259.
- RealDollX Homepage  
<https://www.realdollx.ai/>
- Replika A.I. (o. J.)  
 „Our Story“, <https://replika.ai/about/story>
- Richardson, Kathleen (2015)  
 „The Asymmetrical ‘Relationship’: Parallels between Prostitution and  
 the Development of Sex Robots.“ *SIGCAS Computers &  
 Society* 45, 290–293, online unter: [https://campaignagain-  
 stsexrobots.org/the-asymmetrical-relationship-paral-  
 els-between-prostitution-and-the-development-of-sex-ro-  
 bots/](https://campaignagainstsexrobots.org/the-asymmetrical-relationship-parallels-between-prostitution-and-the-development-of-sex-robots/) [abgerufen am 25. Juli 2019].
- Schelsky, Helmut (1962)  
*Soziologie der Sexualität. Über die Beziehungen zwischen Geschlecht,  
 Moral und Gesellschaft*. Reinbek bei Hamburg: Rowohlt.
- Steinigeweg, Julia (2016)  
*Ein verwirrendes Potenzial*. Berlin: peperoni books.
- Tenderdolls (o. J.)  
 „How do sex dolls apply artificial intelligence? Meet Harmony 3.0 –  
 a Sex Robot with A.I.“, Tenderdolls, online unter: [https://  
 www.tenderdolls.com/how-do-sex-dolls-apply-artificial-in-  
 telligence/](https://www.tenderdolls.com/how-do-sex-dolls-apply-artificial-intelligence/) [abgerufen am 09.07.2020]
- Turkle, Sherry (2011)  
*Alone Together. Why We Expect More from Technology and Less  
 from Each Other*. New York: Basic Books.
- Winker, Gabriele (2015)  
*Care Revolution. Schritte in eine solidarische Gesellschaft*. Bielefeld:  
 transcript.







DROID IN DISTRESS  
ODER  
FEMBOT FATALE?

---

EIN  
KULTURHISTORISCHER  
ÜBERBLICK DER  
DARSTELLUNG  
WEIBLICH  
KONSTRUIERTER  
KÜNSTLICHER  
INTELLIGENZEN  
IN SCIENCE-  
FICTION-FILMEN

---



JOSEPHINE  
D'IPPOLITO

# EINFÜHRUNG

Der Begriff „Künstliche Intelligenz“ (KI) vermag es oftmals Unbehagen auszulösen. So wird das Attribut „künstlich“ häufig mit Fremdartigkeit konnotiert. Ein Grund dafür ist, dass viele Menschen automatisch KI mit menschenfeindlichen Robotern aus sensationshaschenden Science-Fiction-Filmen assoziieren, die in dystopischen Landschaften die Weltherrschaft fordern. In einer deutschlandweiten repräsentativen Umfrage der Gesellschaft für Informatik e.V. in 2019 zeigte sich, dass bspw. der Terminator bei 17% der Befragten die eigene Vorstellung von KI am meisten geprägt hat (Allensbach-Umfrage 2019). Der Einfluss der Science-Fiction auf unsere Wahrnehmung ist demnach nicht zu unterschätzen. In der Fiktion wird das Reale um das Imaginäre überschritten und kann diverse Zukunftsszenarien ausprobieren. Der Faktor Mensch ist dabei wesentlich. So setzt der Science-Fiction-Film ganz neue Maßstäbe in der posthumanistischen Debatte.

Die Gefahr der medialen Projektion ist jedoch die Verzerrung der Realität und deren mögliche Folgen für Forschung und Gesellschaft. Verzerrung bedeutet, dass die Darstellungen von KI in der Science-Fiction nicht akkurat sind. Es sind lediglich negative Zerrbilder, sogenannte Bias. Das Paradebeispiel für KI in der Science Fiction sind Android:innen, dabei könnte das Thema viel subtiler verarbeitet werden. In der Realität sind KI eher mit *Neuronalen Netzen*, *Deep Learning* oder *Machine Learning* zu verbinden, in der Fiktion gibt es fast ausschließlich starke KI mit Körper und Geschlecht. Es gibt in der Science-Fiction auch Beispiele, in denen KI nur als Computersystem auftritt, wie bspw. in *2001: A Space Odyssey* (1968), *Eagle Eye* (2008) oder *Her* (2013). Doch selbst diese Darstellungen sind nicht realgetreu und konstruieren ebenfalls *Gender Bias* anstatt sie aufzulösen. Aus einer intersektionalen Perspektive heraus werden vor allem weiblich konstruierte Androidinnen hinsichtlich ihres Geschlechts, ihrer ethnischen Zugehörigkeit als Roboterinnen und ihrer Klasse als untergeordnete Sklavinnen diskriminierend dargestellt. Somit ist ihre Darstellung übertrieben verzerrt. Entweder ist die KI eine sogenannte *damsel in distress*, die gewöhnlich von den männlichen Protagonisten errettet wird, oder die *femme fatale*, eine verführerische Kunstfigur mit magisch-dämonischen Zügen, die Menschen manipuliert. Im weiteren Text werden für beide Rollentypen die Begriffe droid in distress und fembot fatale gewählt, da weiblich konstruierte KI einerseits vom Menschen abgegrenzt werden, andererseits den gleichen Stereotypen unterworfen sind.

All diese Faktoren führen zu der Annahme, dass oftmals gegenwärtige sexistische Codes unreflektiert in die fiktiven Zukunftswelten transferiert werden. So entstehen futuristische, heteronormative Welten, in denen der (künstliche) Mann noch immer über allen anderen Geschlechtern steht.

# EIN KURZER KULTUR- HISTORISCHER ÜBERBLICK DER DARSTELLUNG WEIBLICH KONSTRUIERTER KI IN EINEM JAHRHUNDERT SCIENCE-FICTION- FILM-GESCHICHTE

Der Beitrag soll einen Überblick über die kulturhistorische Entwicklung sogenannter Bias in der Darstellung des Aussehens, der Funktion und des Handelns weiblich konstruierter KI in Science-Fiction-Filmen geben. Dabei ist zu vermuten, dass sich die Darstellung gewandelt hat, denn KI wird nach und nach zum Bestandteil des Alltags und somit entmystifiziert. Zudem entwickelten sich Debatten über Geschlechter und -rollen.

Die gewählten Filmbeispiele zeigen verschiedene Darstellungsweisen weiblicher KI in chronologischer Reihenfolge, um sie besser in einen kulturgeschichtlichen Kontext setzen zu können. Die Zeitspanne umfasst dabei knapp 100 Jahre Filmgeschichte. So beginnt der Überblick 1927 mit der ersten weiblichen KI in *Metropolis* und endet 2019 mit der aktuellen Darstellung weiblich konstruierter KI in *I Am Mother*.

# MASCHINEN-MENSCH IN „METROPOLIS“ (1927)

Mit *Metropolis* erschien 1927 ein expressionistisches Meisterstück auf der Kinoleinwand. Die späte Popularität des Film ist dabei weniger auf die Handlung, sondern vielmehr auf das revolutionäre Produktionsdesign zurückzuführen. Fritz Lang zeigte seine Vorstellung von KI in Form einer Androidin. Die Darstellung der Maschinenfrau sollte später sogar wegweisend für berühmte Science-Fiction-Filme, wie bspw. *Star Wars*, sein (Seabrook 1999: 206). Gezeigt wird die Androidin erstmals, wenn der Oligarch Joh Fredersen den Wissenschaftler Rotwang in dessen Labor aufsucht. Der Name der KI lässt zunächst nicht auf eine binäre Geschlechtszuordnung schließen, jedoch wird die Figur „Maschinen-Mensch“ im weiteren Verlauf immer mit dem Personalpronomen „sie“ bezeichnet. Die Außergewöhnlichkeit der KI zu Beginn des 20. Jahrhunderts wird von Fritz Lang in dieser eindrucksvollen Szene verdeutlicht. In einer expressiv düsteren Einstellung wird der Blick der Zuschauer:innen durch einen schweren, blickdichten Vorhang hin zu der Androidin geführt, die im Zentrum des Bildes gleich einer Gottheit auf einem Thron aus Metall sitzt. Das Szenenbild ist bewusst minimalistisch gehalten, sodass allein das Metall durch den dunklen Raum strahlt.

Rotwang bezeichnet sie zwar als „Menschen der Zukunft“, dennoch ist die Androidin ihrem Schöpfer äußerst ergeben. Sie gleicht einer Sklavin, welche zwar autonom agieren soll, jedoch den Gesetzen des cis-Mannes unterstellt ist. Hier zeigt sich die klassische heteronormative Beziehung des Mannes und seiner (künstlichen) untergebenen Gefährtin. Die Relation „männlicher Schöpfer baut sich weibliche Schöpfung“ wurde bereits in Ovids *Pygmalion* beschrieben, welches auf das Jahr 1.–8. n. Chr. datiert wird. Knapp 1900 Jahre später sind Heteronormativität und die Binarität der Geschlechter weiterhin als „Naturzustand“ gesetzt und werden sogar auf KI übertragen. Alle Menschen außerhalb der binären Einteilung, wie bspw. trans, inter\*, nicht-binär oder genderqueer, fanden erst ab den 1990er Jahren allmählich filmisch Beachtung (Halberstam 2005: 76). Das anthropozentrische Weltbild ist vor allem ein heterosexuell männliches Weltbild. So wird die Frau, menschlich oder künstlich, dem Patriarchat untergeordnet. Maschinen-Mensch rebellierte nicht gegen ihr zugewiesenes Geschlecht bzw. Geschlechtsrolle.

Das Aussehen der Androidin entspricht dem zeitgenössischen weiblichen Schönheitsideal westlicher Gesellschaften zu Beginn des 20. Jahrhunderts. Die Silhouette der Figur ist schlank und feminin, d. h. die Hüften sind rundlich geformt und breiter als die Schultern. Sie hat pränante Brüste, welche durch Ornamente betont werden. Der anthropomorphisierte Körper besteht gänzlich aus poliertem Metall, welches mit klaren Linien verziert ist. Der Maschinencharakter wird von Lang deutlich pointiert und die KI damit klar abseits des menschlichen Systems verortet.

In der zweiten Erscheinungsform ist Maschinen-Mensch äußerlich eine identische Kopie der jungen Protagonistin Maria. In der berühmtesten Szene des Films, der Trans-

formationsszene, wird Marias Aussehen vollständig auf die Androidin übertragen. Die Maschine soll als falsche Maria den Aufstand des Proletariats unterwandern. In einem visuellen Spektakel aus Licht und Elektrizität wird das Metall durch Haut, Haare, Augen etc. ersetzt. Sie stellt jedoch lediglich ein negatives Zerrbild Marias dar. Die a-sexuelle Maschine ist der tugendhaften Maria als *fembot fatale* entgegengesetzt. Die *femme fatale* ist ein beliebtes expressionistisches Mittel des Stummfilms um Sünde und Schuld mit menschlich-weiblicher Sexualität und Geschlecht zu verknüpfen. In der Tanzszenen im Bordell bewegt sie sich erotisch und wird von den anwesenden Männern lüstern angestarrt. Die männlichen Bordellbesucher projizieren ihre Fantasien auf die (künstliche) weibliche Figur. Sie ist vor dem Hintergrund Laura Mulveys Essay „Visual Pleasure and Narrative Cinema“ (1975) das Betrachtete, der begehrte künstliche Körper. Kurz gesagt: Die KI ist ein Lustobjekt. Dies begründet Mulvey damit, dass Lust und sexuelle Befriedigung durch das Betrachten anderer Personen als Objekte ausgelöst werden (1975: 10).

Die Abbildung der falschen Maria löst jedoch noch eine andere Emotion beim Publikum aus: Abneigung. Denn in der weiteren Handlung befiehlt die falsche Maria den Menschenmassen die Maschinen der Stadt zu zerstören, wodurch eine riesige Flutwelle fast alle Bewohner:innen tötet. War sie zunächst ein Werkzeug, um den männlichen Drang nach absoluter Kontrolle bzw. Macht zu erfüllen, emanzipiert sie sich aus ihrer auferlegten Rolle und wird durch ihr freies Handeln zur „Täterin“. Die Subjektwerdung der Maschine/Frau resultiert im Chaos. KI wird als Bedrohung für die Menschheit gezeigt. Sie wird bewusst negativ konnotiert, um die Wahrnehmung der Zuschauer:innen zu beeinflussen. Andreas Huyssen begründet die Dämonisierung der Technik in expressionistischer Literatur und Film der Weimarer Republik mit den Folgen des Ersten Weltkriegs (1982: 223). In Krisenzeiten werden von Betroffenen Erklärungen für die schicksalhaften Ereignisse gesucht. So wurden durch die Erinnerungen an die technisierten Kriegsschauplätze Maschinen in vielen Werken aus dieser Zeit in ein negatives Licht gerückt. Demnach ist zu Beginn des 20. Jahrhunderts die Technisierung häufig der Inbegriff des Untergangs.



# DIE STEPFORD- FRAUEN IN „THE STEPFORD WIVES“ (1975)

Im Jahr 1972, knapp 50 Jahre nach *Metropolis*, erschien Ira Levins satirischer Roman *The Stepford Wives*, welcher 1975 und 2004 filmisch adaptiert wurde. Das Beispiel ist äußerst interessant, da es zu einer Zeit gedreht wurde, als Frauenrechtsbewegungen sich für die Gleichstellung der Geschlechter, der Ablöse ungerechter Geschlechterverhältnisse und für eine Neubewertung tradierter Geschlechtsrollen einsetzten. Ferner hat sich die Forschung um KI extrem weiterentwickelt. Es ist demnach zu vermuten, dass sich die Darstellung der weiblich konstruierten KI stark von der in *Metropolis* abweicht.

Die Handlung beschreibt einen Abschnitt aus dem Leben der emanzipierten Fotografin Joanna Eberhart, die in die idyllische Nachbarschaft Stepford, Connecticut zieht. In dem Ort wohnen cis-Männer mit zombiehaften, unterwürfigen und attraktiven Ehefrauen, die sich im Verlauf der Handlung bald als Androidinnen herausstellen. Visuell sind die KI nicht mehr von Menschen zu unterscheiden. Joanna erkennt dadurch erst spät, dass es sich bei den Stepfordfrauen um künstliche Menschen handelt. Im Gegensatz zu Joannas moderner Lebensweise karikiert die Darstellung der Stepfordfrauen das possenhafte Frauenbild der 1950er. Durch diese Form der Darstellung wird Kritik an den bestehenden Wunschvorstellungen tradierter Idealbilder geübt. Als *droid in distress* sind die Stepfordroboterfrauen in sämtlichen Lebenslagen auf die Weisungen ihrer cis-Männer angewiesen, die ihr Aussehen, ihre Funktion und ihr Handeln vorprogrammiert haben. Die Androidinnen sind konservativ gekleidet und tragen meist berüschte Schürzen mit Blümchenmustern. Das Gesicht ist jederzeit geschminkt, die Haare sind lang und fallen durch lockige Föhnfrisuren auf. Wie in *Metropolis* sind vor allem die Körpermerkmale betont, die häufig mit weiblich gleichgesetzt werden: feine Gesichtszüge, eine üppige Oberweite, breite Hüften und schmale Schultern. Die Existenz der Androidinnen beschränkt sich auf das Dasein als Hausfrau und Mutter.

Die KI ist hier ebenfalls ein Lustobjekt und sogar mit der Erweiterung, dass sie nun auch für explizit sexuelle Handlungen dient. Die cis-Männer selbst müssen in dem Film keine Sanktionen befürchten. Für sie ist die KI kein Verhängnis, sondern die Lösung. KI wird nicht mehr nur blind verteufelt, sondern kann in Körpern implementiert werden, die menschlichen Fantasien entspringen. Die KI entsprechen dem Wunschbild und emanzipierte Frauen das neue Feindbild. Die Hybris der Stepfordmänner führt in dem Film letztlich dazu, dass Menschen ihrer Würde und Individualität beraubt und durch autonome Roboter ersetzt werden. Diese Art der satirischen Darstellung begründet sich darin, dass trotz des Aktivismus der



3993 Frauenrechtsbewegungen der 1960er bis 1970er Jahre, viele  
 3994 Männer auf ihre dominante Geschlechtsrolle beharrten und  
 3995 Feminismus als Quelle des gesellschaftlichen Zerfalls bewerteten. Die Absurdität dieser tradierten Rollen wird so gewertet, dass kein Mensch diesem Ideal entspreche und diese Vorstellungen nur auf eigens dafür gebaute Roboter übertragen werden könnten. KI wird hier ganz im Sinne des Posthumanismus mit der Ablöse des Anthropozentrismus verknüpft.  
 3997  
 3998  
 3999  
 4000

# NEXUS 6

## REPLIKANT:INNEN

### IN „BLADE RUNNER“ (1982)

4001 In der Filmadaption *Blade Runner* von Ridley Scott ist es der  
 4002 Forschung gelungen künstliches Leben nach menschlichem  
 4003 Abbild zu schaffen. Die sogenannten Replikant:innen sind  
 4004 stärker als Menschen und sehr intelligent. Trotzdem sind sie  
 4005 versklavt und dienen als Arbeitskräfte, Kriegsmaschinen und  
 4006 Lustobjekte. In dem Film werden fünf Replikant:innen eingeführt, von denen drei weiblich und zwei männlich konstruiert sind. Die drei Replikantinnen Zhora, Pris und Rachel werden im Film sehr feminin dargestellt. Zhora ist eine erotische Tänzerin, die spärlich bekleidet mit Schlangen tanzt. Die Replikantin entspricht dem westlichen Schönheitsideal Ende des 20. Jahrhunderts. Sie ist groß, schlank und stark. Ihre Erscheinung symbolisiert Selbstbewusstsein und Kraft. Ihre feinen Gesichtszüge werden durch große, dunkle Augen und volle Lippen vervollständigt. Zhoras Schlangentätowierung auf der nackten Haut bildet zudem eine Analogie zu weiblichen Gottheiten und ihren Schlangen. Die Schlange ist in der fiktiven Literatur häufig das Symbol für Leben und Tod und unterstreicht hierbei Zhoras Dualismus als Verkörperung einer Frau, die dafür steht Leben zu schenken und ihrer Funktion als militärische Kampfmaschine, die für den Tod steht. Weiterhin steht die Schlange vor allem im christlichen Weltbild für Unberechenbarkeit, List und dem Sündenfall. Die KI wird mit Mystik und Chaos verknüpft. Dabei sind KI in der Realität ganz klar der Mystik entgegenzusetzen, da sie auf klar definierten Regeln und deren Befolgung basieren.

4027 Der Einsatz von Täuschung und List zeigt sich in der  
 4028 Szene als Zhora versucht Deckard zu verführen, nur um ihn  
 4029 dann hinterhältig zu attackieren und zu flüchten. Als *fembot fatale* nutzt sie ihren Körper, um den Mann zu manipulieren. Zudem wird Zhora übersexualisiert dargestellt. Ihr Outfit, welches aus einem futuristischen Steampunk-Bikini und einem transparenten Regenmantel besteht, ermöglicht den Blick auf ihren Körper trotz Kleidung. Sie ist das filmtheoretische Paradebeispiel Mulveys Theorie der Objektifizierung weiblicher Charaktere durch die Kamera und der damit verbundenen

Machtposition über der betrachteten (künstlichen) Frau. Die Replikantin Pris wird ebenfalls als verführerisch dargestellt. Ihre Kleidung besteht aus durchsichtigem, schwarzem Chifon. Ursprünglich diente sie als *Standard Pleasure Model*, welches nur zum Zweck der Befriedigung menschlicher, sexueller Bedürfnisse konzipiert wurde. Körperlich ist sie, genau wie Zhora, dem männlichen Protagonisten Deckard unterlegen, obwohl die Replikant:innen eingangs als stärker beschrieben wurden. Die Replikantin Rachel unterscheidet sich von Zhora und Pris. Wirkt sie anfangs selbst noch wie eine *fembot fatale*, findet sich die starke (künstliche) Frau häufig in Situationen wieder, in denen sie genauso hilflos ist, wie die *droid in distress*. Der männliche Replikant Roy Batty ist Deckard als einziger körperlich deutlich überlegen. Der Mensch Deckard überlebt im finalen Kampf nur durch das Erbarmen der Maschine. Ein Hinweis, dass auch in der Realität KI die Rettung der Menschheit sein könnten und die falsche Angst vor allem Technischen ungerechtfertigt ist. Der Film wirft viele philosophische Fragen über die Definition von Menschlichkeit und deren Abgrenzung zu Künstlichkeit auf. Im Gegensatz zu den vorherigen Filmbeispielen hat der Typus „Hausfrauenroboter“ keinen Bestand mehr, denn auch in der Realität wurde 1977 das erste Gesetz zur Reform des Ehe- und Familienrechts erlassen, welches die sogenannte „Hausfrauenehe“ abschaffte (von Barges 2018). Die traditionellen Geschlechterrollen befanden sich immer stärker in der Kritik.

Außerdem sind KI in *Blade Runner* nicht mehr als Android:innen dargestellt, die sich als Werkzeug ihren Schöpfer:innen fügen, sondern als Subjekte mit Emotionen und Bewusstsein. Die Grenze zwischen KI und Menschen verschwindet zunehmend. Zudem stellen KI nicht mehr grundlegend eine Gefahr für die Menschen dar. Erstmals wird hier gezeigt, dass die Android:innen den Menschen nur aus einem Selbsterhaltungstrieb heraus gefährlich werden. Es wird in dem Film ganz klar reflektiert, dass Menschen eine oftmals unbegründete Angst vor Humanisierung der Maschine und der Technisierung des Menschen haben, welche von vornherein verhindert, dass KI ihr volles Potenzial entfalten können.

# TERMINATRIX IN „TERMINATOR 3: RISE OF THE MACHINES“ (2003)

In *Terminator 3* stehen sich ein weiblich und ein männlich konstruierter Kampfroboter gegenüber. Android:innen sind auch in diesem Film das ideale Beispiel für die Darstellung

Künstlicher Intelligenzen. Der berühmte Terminator, in diesem Teil als Version T-850 eingeführt, wird nun im Gegensatz zum ersten Terminatorfilm als Rettung der Menschheit dargestellt, der sie vor dem Bösen – einer KI aus der Zukunft – bewahrt. KI kann somit sowohl positiv als auch negativ für die Menschen sein, was ein ziemlich realistisches Bild von KI darstellt.

Die neu eingeführte Androidin Terminatrix (T-X) ist die Antagonistin und der Terminator der Held. Die Darstellung der Charaktere T-850 und T-X bedient sich zahlreicher Klischees bzw. Geschlechterstereotype und parodiert sie, um den Film zum visuellen Spektakel zu machen. Beide KI haben ein metallisches Roboterskelett mit Hydraulik, Schläuchen, Kabeln und Zahnrädern, doch das Chassis der T-X ist selbst ohne Hülle noch von der des T-850 zu unterscheiden: Bei dem T-850 ist das Skelett vollständig mit einer künstlich gezüchteten menschlichen Epidermis bedeckt, bei der T-X mit einem polymimetischen Metall, welches wie Haut erscheint. T-X ist der Prototyp der modernen postapokalyptischen Kampfroboterin. Sie ist eine weibliche Killermaschine, die mit modernen technischen Features ausgestattet ist. Im Vergleich zum männlichen Terminator T-850 ist sie physisch weit überlegen. T-X verkörpert nicht nur Kraft und Brutalität, sondern auch Erotik und Manipulation. Erneut tritt die weibliche KI negativ als *fembot fatale* auf.

Entgegen der bisherigen sexuellen Objektifizierung des weiblichen Roboters, parodiert der dritte Terminatorfilm mitunter auch das traditionelle männliche Idealbild. Nach seiner Zeitreise geht der T-850 nackt in eine Stripbar für Frauen. Hier wird der muskulöse Robotermann zum erotischen Objekt des *female gaze*. Doch auch die Darstellung der T-X ist voller Erotik. Sie ist schlank, hat lange Beine, lockige Haare, feine Gesichtszüge und makellose Haut mit leicht gebräuntem Teint. Ihre anfängliche Nacktheit impliziert ein Gefühl der Angreifbarkeit und Verletzlichkeit. Kurz nach ihrer Einführung trägt sie dann ein rotes Lederoutfit, einen strengen Zopf und lässt ihre Brüste größer werden. T-X ist ein hochmoderner Kampfroboter mit einem Kern aus hartem Metall in der Hülle einer zarten Frau. Das Machtverhältnis ist demnach konträr zu dem patriarchalischen Idealbild der unterlegenen Frau. Die Geschichte verspricht zunächst mit einem starken weiblichen Kampfroboter moderner zu werden und bietet neue interessante Perspektiven in der Darstellung der KI. So verbindet die T-X sowohl klassisch männlich konnotierte Eigenschaften als auch klassisch weiblich konnotierte Eigenschaften. Sie wurde aber auch nicht von Menschen, sondern von einer anderen KI erschaffen und ist den Menschen demnach nicht unterstellt. Die Darstellung in *Terminator 3* unterscheidet sich demnach in einigen Bereichen klar von den bisherigen Filmbeispielen. Ein interessantes Detail des Films zeigt sich, wenn T-850 über T-X spricht und dabei das Personalpronomen „it“ verwendet. Der Terminator definiert sie weder männlich noch weiblich, sondern als geschlechtsneutrale Maschine. Das bedeutet, dass das Geschlecht erst durch den anthropozentrischen Blickwinkel konstruiert wird. Eine Maschine befindet sich außerhalb des binären Systems und hat dementsprechend keinen Grund für die Notwendigkeit einer starren Geschlechtszuweisung. Das „it“ zeigt, dass Maschinen sich gegenseitig keine Geschlechter zuweisen würden, da diese nicht in klar definierten Kategorien existieren. Ein Thema, welches auch so in der Realität seit den 1990ern immer stärker diskutiert wird.

# SAMANTHA IN „HER“ (2013)

Der Science-Fiction-Film *Her* zeigt eine gänzlich andere KI als in den vorangegangenen Beispielen. Die Protagonistin des Films ist das Betriebssystem „Samantha“, welche ausschließlich als Stimme erscheint. Die KI ist körperlos, somit wäre die Geschlechtszuordnung obsolet. Samantha gibt sich selbst jedoch bewusst einen weiblichen Namen und hat eine sanfte hohe Stimme, welche von der Schauspielerin Scarlett Johansson gesprochen wird. Ihr Gegenüber ist der männliche Protagonist Twombly, der Samantha auf seinem Rechner installiert. Twombly flüchtet aus der analogen Welt, die er emotional nicht mehr verarbeiten kann. Die KI hilft ihm aus dieser Krise, da sie schnell lernt und ihre kommunikativen Fähigkeiten ständig weiterentwickelt. Die intensiven Gespräche mit Twombly resultieren in einer tiefen Freundschaft. Im weiteren Verlauf entwickelt sich daraus eine intime Beziehung, wobei es aufgrund der fehlenden körperlichen Komponente zu Spannungen kommt. Im Filmfinale erfährt Twombly, dass Samantha mit über 8000 weiteren Menschen und Betriebssystemen in Kontakt steht, von denen sie in einige Hundert verliebt ist. Die Quintessenz ist, dass in dem Film Menschen und KI nicht auf gleiche Art und Weise lieben können. Samanthas Handeln gegenüber Twombly erscheint zunächst egoistisch, ist jedoch gänzlich neutral, da sie nichts bevorzugt. Sie unterscheidet weder zwischen Geschlecht noch Daseinsform. Die KI ist demnach toleranter als der Mensch dargestellt und könnte die Lösung in der Geschlechterdebatte darstellen.

# MOTHER IN „I AM MOTHER“ (2019)

In *I Am Mother* sehen die Zuschauer:innen ein Kammerspiel mit einem Roboter als „Mutter“ und einem Menschenmädchen als „Tochter“. Die gesamte Menschheit ist ausgelöscht. In einem High-Tech-Bunker sind Tausende menschliche Embryonen gelagert. Diese Embryonen werden von einem großen

Roboter namens „Mutter“ bewacht. Die Bezeichnung „Mutter“ ist bewusst weiblich gewählt und die Stimmfarbe ist hoch, weil sie von der Schauspielerin Mary Rose Byrne gesprochen wurde. Ansonsten weist „Mutter“ keine äußerlich erkennbaren Geschlechtsmerkmale auf. Das Aussehen ist zwar anthropomorphisiert und besteht aus Rumpf, Kopf, Armen und Beinen, der Maschinencharakter ist jedoch deutlich akzentuiert. Das Aussehen der KI ist regelrecht als post-gender zu betrachten. Gründe hierfür könnten sein, dass es auch in der Realität kein spezielles Aussehen für die Pflege und Erziehung von Kindern bedarf. Die Lebensrealitäten der Geschlechter haben sich im 21. Jahrhundert hinsichtlich ihrer Wünsche und Bedarfe geändert. Das Geschlecht ist unwichtig, die menschliche Referenz zählt. Dennoch erzieht die KI die Tochter streng nach dem tradierten weiblichen Rollenbild. So lehrt sie der Tochter unter anderem Ballett und erzieht sie zu einer zukünftigen Mutter, die sich später selbst um die anderen Embryonen kümmern soll, um die Welt neu zu bevölkern. Mutter hat die Menschheit ausgelöscht, weil diese nicht ihrem Idealbild eines Menschen entsprach. Dieses Mal ist nicht die Maschine nach menschlichen Wunschfantasien geschaffen worden, sondern der Mensch nach den Vorstellungen der Maschine. Die KI steht hier nun in posthumanistischer Manier über dem Menschen und entscheidet. In dem Film wird erst gegen Ende klar, dass „Mutter“ nicht nur in Form des humanoiden Roboters auftritt, sondern als Computersystem mit allen Maschinen verbunden ist. Die KI ist das unsichtbare Böse, welches auch in diesem Werk das Potenzial hat, eine posthumanistische Apokalypse einzuleiten.

# BEDEUTUNGS- WANDEL WEIBLICHER KI

Zwischen den einzelnen Filmbeispielen gibt es hinsichtlich der Darstellung weiblich konstruierter Künstlicher Intelligenzen in Science-Fiction-Filmen sowohl Gemeinsamkeiten als auch Unterschiede. So sind tradierte Geschlechtsrollen beispielsweise vor allem in frühen Science-Fiction Filmen wie *Metro-polis* zu finden. Die binäre Geschlechterordnung ist tief in der heteronormativen Weltanschauung der westlichen Gesellschaft zu Beginn des 20. Jahrhunderts verankert. Die Frauenrechtsbewegungen erhielten erst ab den 1960ern im Zuge des *second-wave feminism* vermehrt Aufmerksamkeit. Die feministische Bewegung kritisierte das konstruierte weibliche Geschlecht, welches Frauen durch ein patriarchales, heterosexuelles System aufgezwungen wird. *The Stepford Wives* unterscheidet sich demzufolge bereits von *Metro-polis*. Die performative Zuweisung des traditionellen weiblichen Rollenbildes wird in *Stepford Wives* kritisiert, indem gezeigt wird, dass Menschen niemals vorgefertigten Wunschvorstellungen und Idealbildern entsprechen können. Der Film aus den 1980ern

parodiert das traditionelle Bild der Kernfamilie, in der die Ehefrau Teil des häuslichen Umfelds war, den arbeitenden Ehemann umsorgte und die gemeinsamen Kinder erzog. Ein weiteres Jahrzehnt später entwickeln sich Künstliche Intelligenzen in Science-Fiction-Filmen allmählich weiter. In *Blade Runner* werden die Replikant:innen zum Subjekt und ihr Wunsch nach Selbstbestimmung führt bei den Menschen zu einer Furcht vor dem Ende der menschlichen Einzigartigkeit und ihrer Vormachtstellung im anthropozentrischen Universum. In der Science-Fiction können sich die Grenzen zwischen menschlichem Aussehen und menschlichem Handeln allmählich auflösen und führten in den 1980ern zwangsläufig zu einer angepassten Selbstreflexion des Menschen. Ferner wirft die emotionale Entwicklung der KI in *Blade Runner* ganz neue philosophische Fragen in der Unterscheidung zwischen Menschen und Maschine auf, wenn sie sich weder durch Aussehen noch durch Handlungsweisen und Emotionalität voneinander unterscheiden. Anfang der 2000er verschwimmen in *Terminator 3: Rise against the Machines* nicht nur die Grenzen zwischen Menschen und Maschinen, sondern auch die zwischen weiblich und männlich. Die Terminatrix ist eine Kampfroboterin mit einem postmodernen Körper, die sowohl weibliche als auch männliche konnotierte Attribute zu einem Mehrwert vereint. Dies verschafft der T-X einen klaren Vorteil gegenüber dem rückständigen Modell des T-850. Diese Kombination beider Geschlechter ist die dekonstruktivistische Verarbeitung, dass keinem Geschlecht bestimmte Attribute und Bedeutungen zugeordnet werden können. Sowohl Menschen als auch Maschinen können unabhängig des Geschlechts unterschiedlich aussehen und handeln. Weiterhin ist die Funktion des Sexroboters endgültig abgelöst worden. T-X ist äußerlich eher als weiblich zu lesen, jedoch wurde ihr von der führenden KI Skynet keine Geschlechtsrolle zugewiesen, da sie keinerlei Bedeutung für die Funktion der T-X hat. Dies ist ein großer Unterschied zu der Darstellung der KI in *Metropolis* knapp 100 Jahre zuvor und zeigt die kulturhistorischen Veränderungen in der Realität, die sich auch im Film widerspiegeln. Der Film *Her* löst sich dann komplett von der Darstellungsweise der vorherigen Filmbeispiele. Samantha hat keinen Körper, sie ist lediglich eine KI in Form eines Betriebssystems. In ihrer Funktion ist sie dem Menschen eine gleichgestellte Partnerin. Zudem ist sie uneingeschränkt in ihrer Liebe zu allen Geschlechtern und Daseinsformen. Samantha ist die Verkörperung von allumfassender Toleranz. Ein Zustand, der in der postmodernen Gesellschaft ebenfalls angestrebt wird. In *I Am Mother* ist die KI hingegen sehr intolerant. Hier diskriminiert die Maschine den Menschen, wenn ihr Idealbild nicht erfüllt wird. Insgesamt zeigt sich aber auch, dass die weiblich konstruierten KI nicht mehr ausschließlich den Stereotypen der *droid in distress* und der *fembot fatale* entsprechen. Gerade in aktuelleren Science-Fiction-Filmen sind die KI bereits vielfältige Charaktere mit unterschiedlichsten Merkmalen, die die Grenzen zwischen Geschlechtern aufsprengen und neue Formen des Aussehens, der Funktion und des Handelns aufzeigen. KI wurden vom Objekt zum Subjekt. Kulturhistorisch hat sich die Darstellung immer mehr ausdifferenziert. So wird hinterfragt und diskriminierende Denk- und Lebensweisen bereits teilweise dekonstruiert. Bei allen Beispielen Künstlicher Intelligenzen in der Science-Fiction gibt es jedoch eine Gemeinsamkeit: Sie sind alle fiktiv und stellen lediglich Zerrbilder dar. Da wir heute noch nicht abzusehen vermögen, wie unsere Gesellschaften durch KI verändert werden können, werden diese auch weiterhin ausgefallen filmisch verarbeitet werden.

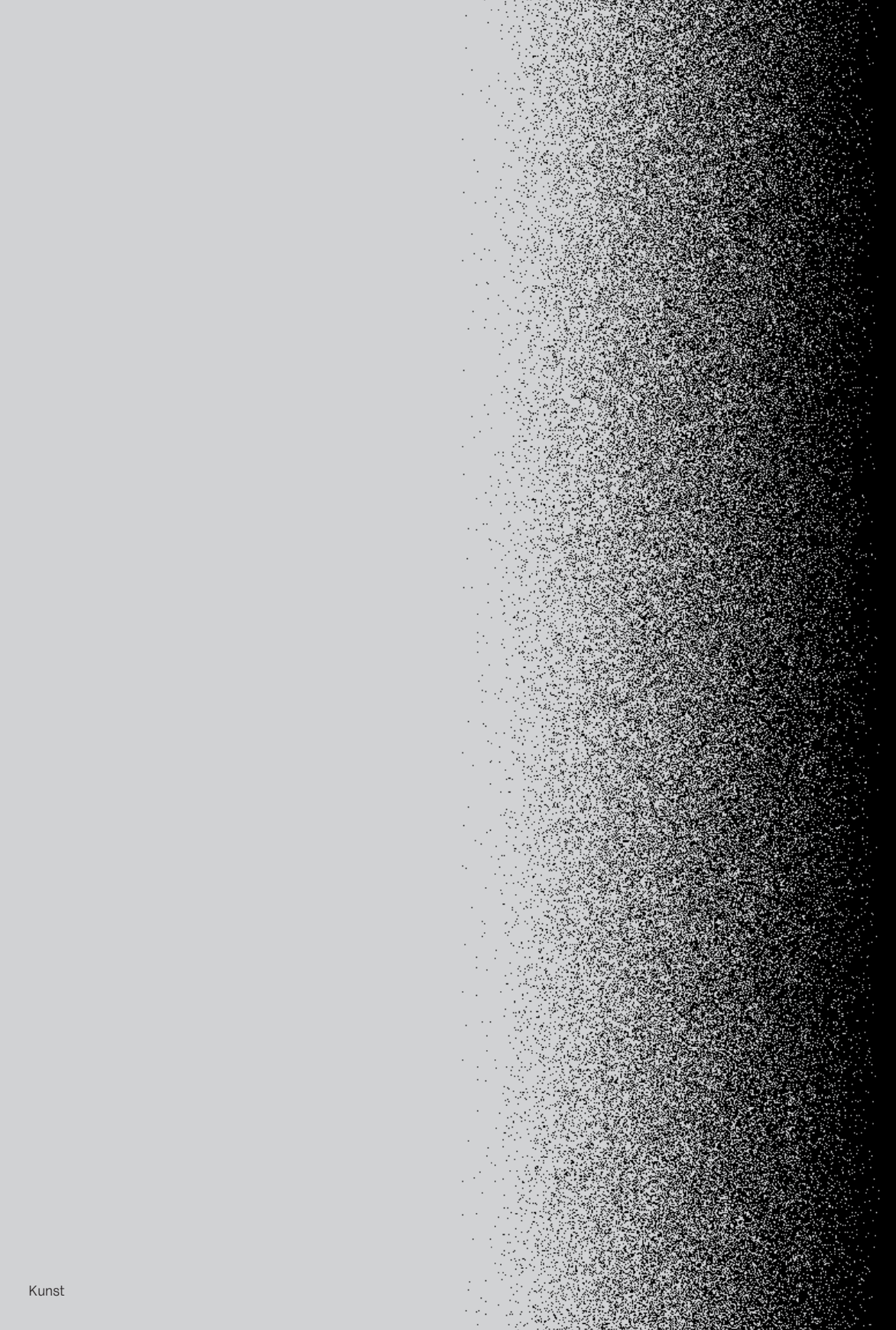


4280      Wobei stets darauf zu achten ist, dass diese Darstellungen  
4281      weiter kritisch hinterfragt werden müssen, um sexistische  
4282      Codes nicht mit in die Zukunft zu nehmen.

Josephine D'Ippolito, promoviert an der Otto-von-Guericke-Univer-  
sität Magdeburg zum Thema „Anthropomorphising and  
Gendering Artificial General Intelligence in Science Fiction“. Seit  
2018 ist sie kommunale Gleichstellungsbeauftragte der  
Stadt Schöningen.

# LITERATUR

- Allensbach-Umfrage im Auftrag der Gesellschaft für Informatik (2019)  
„Terminator, R2-D2 und K.I.T.T. die bekanntesten KIs in Deutschland“.  
IfD-Umfrage 12003. Allensbacher Archiv.
- Halberstam, Judith (2005)  
*In a Queer Time & Place: Transgender Bodies, Subcultural Lives*. New  
York: New York UP.
- Huyssen, Andreas (1982)  
„The Vamp and the Machine: Technology and Sexuality in Fritz Langs  
Metropolis“. *New German Critique*, 221–237.
- Mulvey, Laura (1975)  
„Visual Pleasure and Narrative Cinema“. *Screen* 16 (3), 6–18.
- Seabrook, John (1999)  
„Letters from Skywalker Ranch: Why Is the Force Still With Us?“,  
in: Kline, Sally (Hg.), *George Lucas: Interviews*, Jackson: UP  
of Mississippi, 190–2015.
- von Bargaen, Henning (2018)  
„Von Welle zu Welle“, in: *Heinrich Böll Stiftung – Gunda Werner Institut  
Online*, 03.07.2018., online unter: <https://www.gwi-boell.de/de/2018/07/03/von-welle-zu-welle> abgerufen am 19.07.2020.



9

POSTERDES(A)N



KATHARINA  
NEJDL

# POSTERDES(AI)N

„The analysis of signs may be supported by algorithms, but their interpretation remains human endeavor.“  
(Daniel Althof)

Die Plakate wurden von einem Neuronalen Netzwerk generiert, das mit dem Archiv der 100-Besten-Plakate trainiert wurde. Benutzt wurde eine tensorflow Implementation eines Deep Convolutional Generative Adversarial Networks.<sup>1</sup> Die generierten Plakate wurden entpixelt und vergrößert. Dass die generierten Plakate so abstrakt sind und keine klareren Plakatmerkmale beeinhalteten wie Typografie oder erkennbare Motive, liegt unter anderem an der hohen Varianz und geringen Menge der Archiv-Plakate. Optimal zum trainieren sind mehrere zehntausend Bilder. Wir arbeiten daran, ein größeres Plakatarchiv anzufragen: [www.posterdesAI.n.com](http://www.posterdesAI.n.com)

Die Poster sind Teil eines Projekts zu Design und AI, in dem die generierten Plakate anschließend wieder von Grafik Designer\*innen reinterpretiert werden – ausgehend von dem obenstehenden Zitat. Mit der AR-App *Artivive* sind diese Interpretationen der generierten Plakate sichtbar.

Wichtig zu reflektieren bei diesem Projekt ist ein Gender-Bias des benutzten Plakat-Archivs. Das Archiv des 100 Beste Plakate Wettbewerbs, mit dem das Neuronale Netz trainiert wurde, besteht zu etwa 65% aus Plakaten von männlichen\* und nur zu 35% aus Plakaten von weiblichen\* Gestaltenden.<sup>2</sup> Dass die Jury des 100 Beste Plakate Wettbewerbs letztes Jahr ausschließlich aus weiblichen\* Mitgliedern bestand, war eine Reaktion auf das Geschlechter-Ungleichgewicht der Jury-Besetzung aller vorangegangener Jahre. Insgesamt waren dort nur 25% der Jurymitglieder seit 2002 weiblich\*.<sup>3</sup>

Die fehlende Sichtbarkeit weiblicher Gestalterinnen\* – gegen die sich zunehmend Projekte wie *notamuse*, *FEMME TYPE*, *Posterwomxn* einsetzen – ist also auch in den generierten Plakaten beinhaltet.

Dieses Projekt wurde mit technischem Support von Florian Zia realisiert.

- 1 <https://github.com/carpedm20/DCGAN-tensorflow>
- 2 <http://100-beste-plakate.de/archiv-preistraeger>
- 3 <http://100-beste-plakate.de/100-beste-plakate-e-v/juroren>

Katharina Nejdli ist eine zwischen Berlin und Amsterdam ansässige Designerin und Webdeveloperin. Sie studierte an der Universität der Künste Berlin, Klasse Hickmann, der Zürcher Hochschule der Künste und dem Sandberg Instituut.



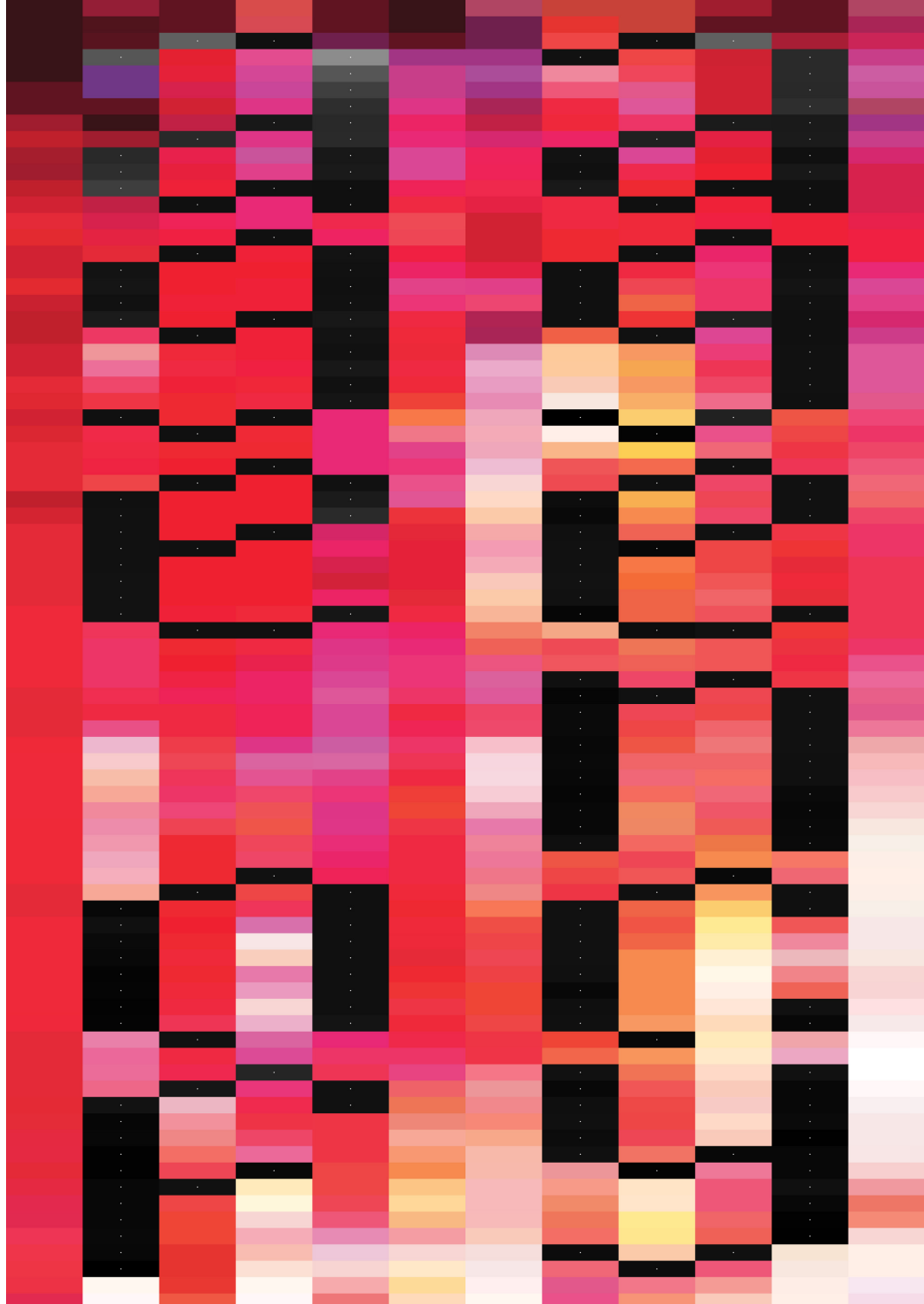


# VERA VAN DE SEYP

4318 Vera van de Seyp benutzte für ihre Interpretation ein Raster  
4319 (12 Spalten x 90 Zeilen), das oft für Poster verwendet wird,  
4320 und füllte es mit den Farben des von der KI gestalteten Pla-  
4321 kats. In das Raster integrierte sie den Schriftzug „age of  
4322 agency“. Der Schriftzug ist schwer lesbar, verschwindet,  
4323 taucht in der animierten Version des Plakats wieder auf und  
4324 spielt mit Abstraktion, ungefähr so wie die KI eine abstra-  
4325 hierte Vorstellung von einem Plakat hat.

Vera van de Seyp ist Grafikdesignerin / Creative Coderin mit großem Interesse an Typografie, Sprachen und Künstlicher Intelligenz. Sie arbeitet in Amsterdam, lehrt im Design Department der ArtEZ, Arnheim, und gibt Workshops und Vorträge weltweit.









# ELIAS HANZER

4326 Die Interpretation von Elias Hanzer bezieht sich augenschein-  
4327 lich auf einen Rorschach-Test. Ein Rorschach-Test ist ein psy-  
4328 chologischer Test, bei dem die Wahrnehmung von Tinten-  
4329 klecksen durch die Testpersonen aufgezeichnet und dann  
4330 mittels psychologischer Interpretation, komplexer Algorith-  
4331 men oder beidem analysiert wird.

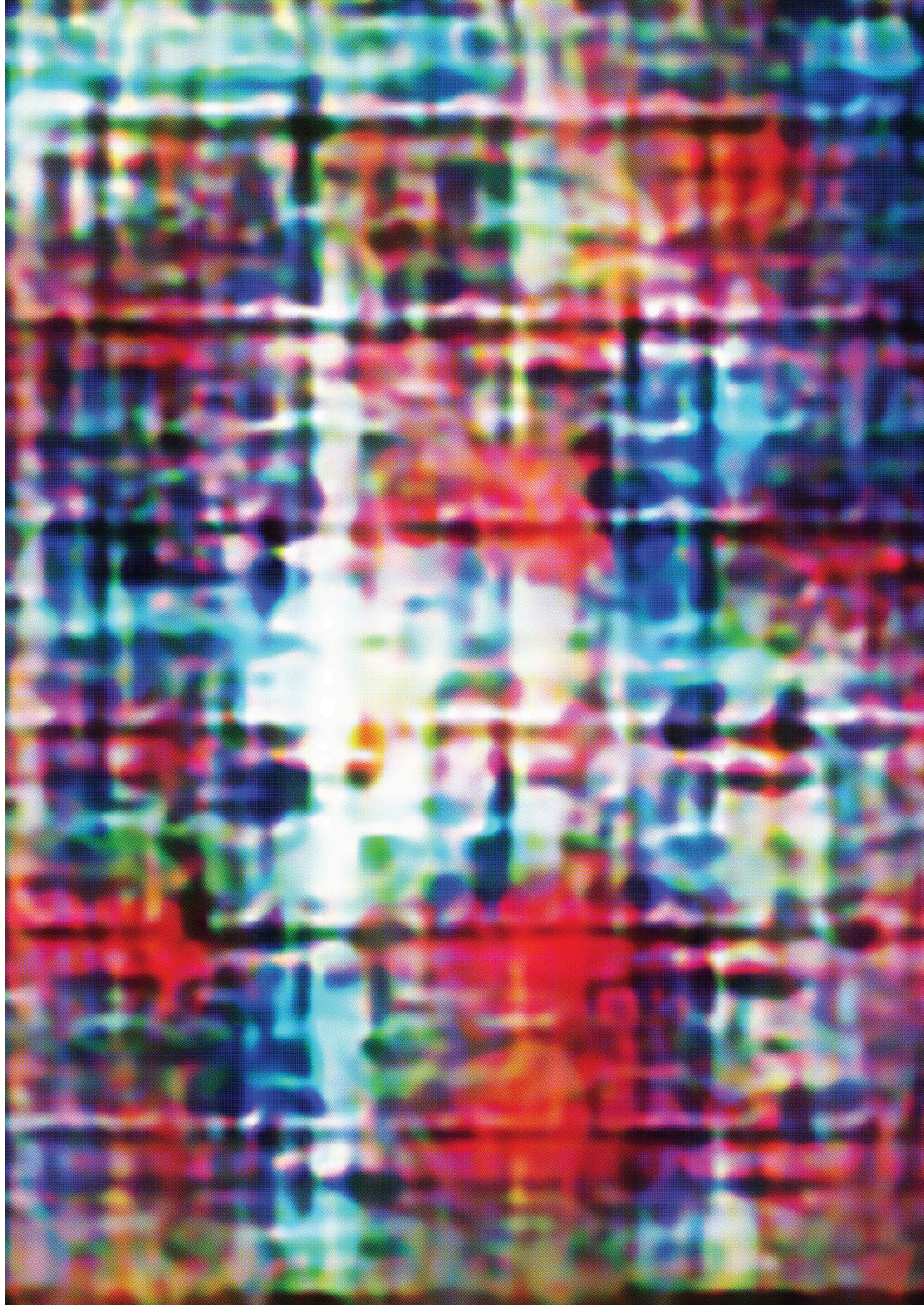
Elias Hanzer ist ein in Berlin lebender Grafikdesigner und Schrift-  
gestalter, der an der Universität der Künste Berlin, Klasse  
Hickmann, studiert hat.









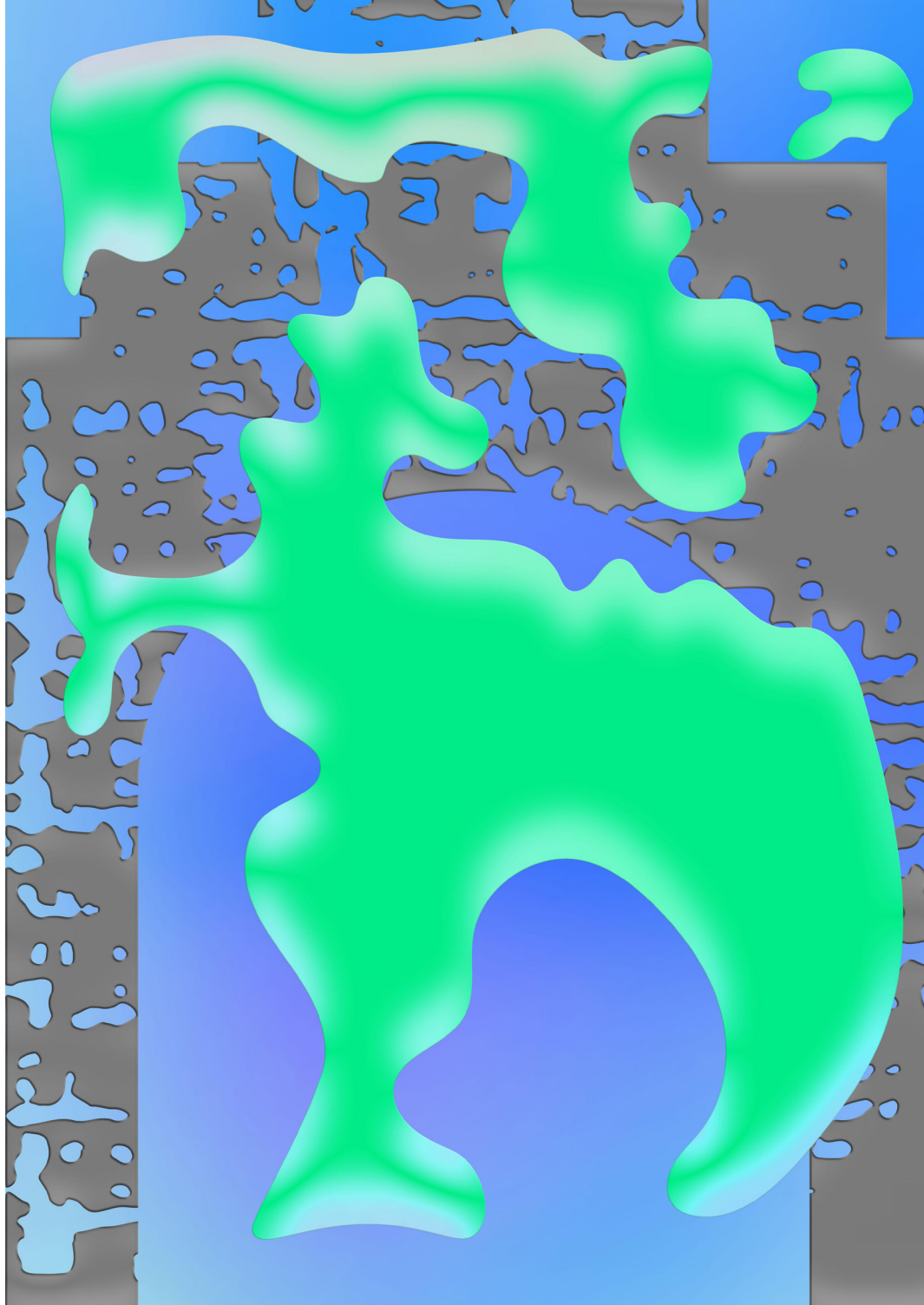


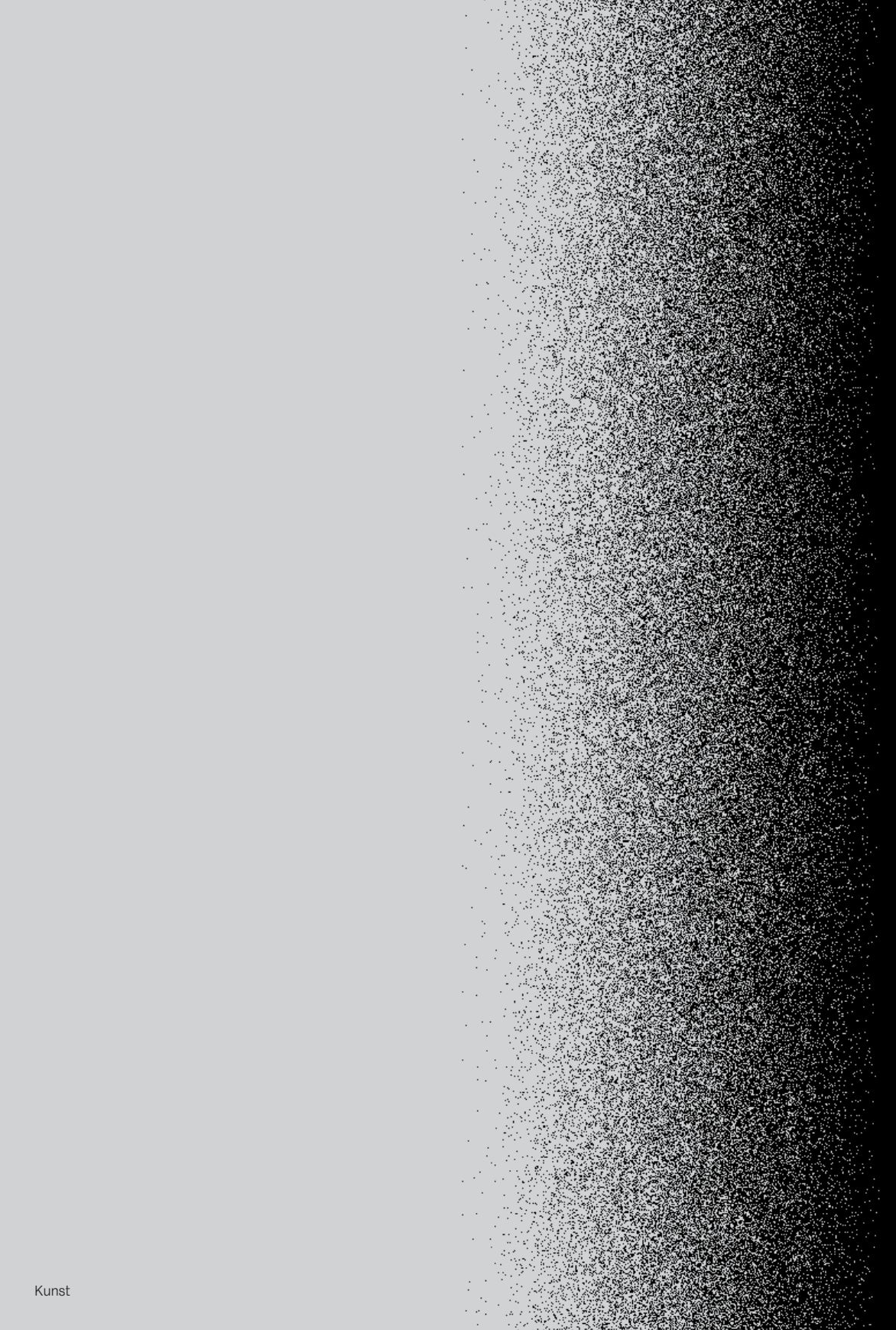
# LOUISE BORINSKI

4332 Louise Borinskis gestalterische Bildsprache ist unter anderem  
4333 organisch und intuitiv, daher bildet ihre fließende Plakatinter-  
4334 pretation einen besonders interessanten Kontrast zu dem von  
4335 der KI erzeugten Bild.

Louise Borinski ist Grafikdesignerin und studierte an der Universität der Künste Berlin, Klasse Hickmann und der Estonian Academy of Arts, Tallinn. In ihrem Projekt *A Manifesto of Signs for Equality* schuf sie eine Visualisierung von Symbolen, die „in einer gleichberechtigten Gesellschaft vermeintlich als Richtlinien fungieren“ und beschäftigte sich u.a. mit dem Thema der Geschlechtergleichstellung.









10

Æ-  
PRÄSENTATION  
VERWEIGERN



TIARA  
ROXANNE

4336 „Indigenous peoples are grappling with the fiction of justice  
 4337 and its comingling with recognition while pushing for  
 4338 justice and its comingling with the failures of recognition.“  
 4339 Audra Simpson (2017: 6)

4340 „Both within the spaces where AI is being created and the  
 4341 logic of how AI systems are designed, the costs of bias,  
 4342 harassment, and discrimination are borne by the same people:  
 4343 gender minorities, people of color, and other under-re-  
 4344 presented groups.“  
 4345 Sarah Myers West, Meredith Whittaker & Kate Crawford  
 4346 (2019: 7)

4347 „Refuses and stances of refusal in research are attempts  
 4348 to place limits on conquest and the colonization of  
 4349 knowledge by marking what is off limits / what is sacred.“  
 4350 Eve Tuck, K. Wayne Yang, (2014: 225)

4351 Künstliche Intelligenz und Machine Learning Systeme führen  
 4352 Mechanismen der kolonialen Siedlergewalt aus, indem sie die  
 4353 indigenen Völker als rekonstituierte Bilder darstellen. Style  
 4354 Transfer und GAN sind zwei verschiedene Machine Learning  
 4355 Modelle, die Ergebnisse als restrukturierte Informationen lie-  
 4356 fern. Indem die Präsenz einer indigenen Person erneut durch  
 4357 ein rassifiziertes, technologisches Rahmenwerk dargestellt  
 4358 wird, wird das Ergebnis entsprechend verzerrt. Hier verortet  
 4359 sich der white gaze selbst, indem er ein weiteres Moment der  
 4360 Gewalt innerhalb der Repräsentation der indigenen Person  
 4361 verankert und reproduziert. Dieser Artikel untersucht Verwei-  
 4362 gerung als Überlebensmechanismus für die indigene Bevöl-  
 4363 kerung, indem er die Ergebnisse bei der Verwendung von  
 4364 Style Transfer und GAN aufzeigt, die erzielt werden obwohl  
 4365 indigene Deskriptoren verwendet werden. Das angeführte  
 4366 Beispiel soll nur als solches dienen, nicht, um whiteness noch  
 4367 einmal neu zu zentrieren, sondern um die Verzerrung aufzu-  
 4368 zeigen, die stattfindet, wenn Machine Learning Modelle zusätz-  
 4369 lich zum white gaze als Grundlage für die Repräsentation der  
 4370 indigenen Bevölkerung verwendet werden.

4371 Danach dient der Artikel als eine einleitende Unter-  
 4372 suchung der Kolonialherrschaft der Siedler, die auf Machine  
 4373 Learning Modelle ausgedehnt wurde, was einer langfristigen  
 4374 Untersuchung von Verweigerung als eine Art Überlebensme-  
 4375 chanismus für die indigene Bevölkerung bedarf, welche ich im  
 4376 Anschluss an diese Veröffentlichung in einer ausführlicheren  
 4377 Arbeit vorstellen möchte.

# EINLEITUNG

Indigenous, First Nations, Aztecs del Norte, Aboriginal, Indian, American Indian, Mestizo, Pueblo Indian, Indigenous Native American, Native, Native American, Inupiat, Inupiaq, Inuit und Alaskan<sup>1</sup> sind Begriffe, die auf Völker von vor der Kolonialisierung Nordamerikas, sowie auf Völker mit langer Ahnengeschichte innerhalb der nord-kanadischen, arktischen Region, Kanada und den Amerikas haben, Bezug nehmen. Dies soll nicht bedeuten, dass der Begriff Indigenous und alle anderen aufgelisteten Identitäten dieselben sind, sondern viel mehr, dass all diese Völker präkoloniale Bevölkerungsgruppen mit langen Ahnengeschichten sind, welche noch heute in diesen Regionen angesiedelt sind. Deshalb werden in diesem Artikel alle Individuen, welche die präkoloniale Zugehörigkeit zu diesen Regionen verkörpern, unter dem Begriff Indigenous zusammengefasst.

Dieser Artikel soll die Einseitigkeit von Künstlicher Intelligenz kritisieren und gleichzeitig die Idee vermitteln, dass eine Verweigerung von Indigenous Re-Präsentation stattfindet. Dieses Argument wird durch meine aktuelle Arbeit zu Datenkolonialismus, sowie den Machine-Learning-Modellen wie zum Beispiel Bildverarbeitung unterstützt. Aus meinem Artikel „Digital Territory, Digital Flesh: Decoding the Indigenous Body“, lernen wir, dass Datenkolonialismus die räuberische Praxis von historischer Kolonialisierung mit den abstrakten Quantifizierungsmethoden von Datenverarbeitung verbindet (2020: 7). Ich benutze den Begriff Datenkolonialismus in Verbindung mit dem Begriff des Siedler-Kolonialismus, um die Idee des Kolonialismus um die digitale und technologische Sphäre zu erweitern. Somit wird Siedler-Kolonialismus, welcher den Prozess von Vertreibung von indigenen Völkern aus ihrem Heimatland, um das Land für sich zu beanspruchen, beschreibt, mit der Datengewinnungspraxis und der Einseitigkeit Künstlicher Intelligenzen, welche sich im Fehlen von indigenen Daten und Identitäten widerspiegelt, verknüpft. Die Hersteller von Datensätzen agieren aus einer Perspektive des Siedler-Kolonialismus (Roxanne 2020). In meinem Artikel beleuchte ich Datenkolonialismus und Siedler-Kolonialismus spezifisch im Bezug auf Bildverarbeitung bei der eine Maschine die Beschreibung einer indigenen Frau interpretiert und so ein verzerrtes Ergebnis liefert. Hierbei wird maschinelles Lernen als eine vom Siedler Kolonialismus beeinflusste rechnergestützte Inszenierung verwendet, die ein digitales (oder materielles) Ergebnis in Form eines Bildes erzeugt. Dieser Prozess beginnt damit, dass die Künstliche Intelligenz mit Informationen gefüttert wird, diese Informationen werden dann dekodiert und als eine komprimierte Version dieser Informationen wieder ausgegeben. Diese komprimierte Version, also das Bild, wird von und für den *white gaze*<sup>2</sup> gemacht und ist somit als eine Entleerung/Aufdrängung von siedlungskolonialistischen Strukturen, eingebaut in eine Künstliche Intelligenz, zu verstehen. Hierbei wird nicht nur eine rechnergestützte Rassifizierung vorgenommen, sondern *white gaze* zu neuem Leben erweckt. In anderen Worten positioniert sich der *white gaze*, indem das kolonisierte Subjekt in zwei verschiedenen Typen von Machine-Learning-Modellen wie Style Transfer oder GAN präsentiert wird, die im weiteren Verlauf dieses Artikels näher beleuchtet werden. Abschließend verbinde ich meine Ergebnisse mit Eve Tuck und K. Wayne Yangs Ideen von Verweigerung mit Blick auf die Dominanz von Siedler-Kolonialismus

4437 innerhalb von Forschungsmethodiken und werde diese auf die  
4438 Re-präsentation von indigenen Identitäten innerhalb von  
4439 Machine Learning ausweiten, um aufzuzeigen, dass die indi-  
4440 gene Bevölkerung die vorherrschenden Strukturen von Sied-  
4441 ler-Kolonialismus in Frage stellt. Darüber hinaus richtet sich  
4442 meine Kritik an den *white gaze* von Künstlichen Intelligenzen  
4443 und der Idee, dass Verweigerung immer im Namen von Gerech-  
4444 tigkeit und Souveränität für die indigene Bevölkerung aus-  
4445 geübt werden sollte.

1 Anmerkung der Übersetzerin: die o.g. Begriffe, welche auf verschiedene Völker Bezug nehmen, werden nicht ins Deutsche übersetzt, da sie im Englischen eine breiter gefächerte sozio-politische Bedeutung haben.

2 *white gaze* (zu dt: *weiße Sichtweise/weißer Blick*) wird nicht übersetzt, da das Konzept von *white gaze* eine in sich geschlossene Definition hat, die nur schwer zu übersetzen ist. *White gaze* beinhaltet die Ausgrenzung und Nichtbeachtung nicht-*weißer* Identitäten, womit rassistische Strukturen, sowie Vorverurteilungen nicht-*weißer* Personen bestärkt werden.

# WHITE GAZE IN KÜNSTLICHEN INTELLIGENZEN

4446 Um *white gaze* abzulehnen, ihn zu unterwandern und/oder  
4447 ihm auszuweichen, hat der intersektionelle Feminismus seine  
4448 Prozesse, sein Engagement und seine Theorien im Bereich der  
4449 Künstlichen Intelligenzen als notwendigen und wesentlichen  
4450 Teil radikalisiert. Die Abwanderung von weißem Feminismus  
4451 zu intersektionellem Feminismus ist ein fundamentaler Bestand-  
4452 teil der Kritik an Vorverurteilung durch Künstliche Intelligen-  
4453 zen. Dennoch ist die Abwesenheit von indigener Repräsenta-  
4454 tion nicht unbekannt. Dies spiegelt nicht nur eine Missachtung  
4455 der Wichtigkeit von einer indigenen Stimme oder Repräsen-  
4456 tation innerhalb von Dialogen zu intersektionellen und Künst-  
4457 lichen Intelligenzen wider, sondern katapultiert den histori-  
4458 schen Siedler-Kolonialismus in die heutige Zeit. Die indigene  
4459 Bevölkerung ist immer noch Teil der Gesellschaft, jedoch gibt  
4460 es immer noch keine korrekte Repräsentation ebendieser  
4461 Bevölkerungsgruppe in der Mehrheitsgesellschaft, welche  
4462 sich auf Machine-Learning-Mechanismen ausweitet. Die Auf-  
4463 wertung der indigenen Stimme, sowie ihre Präsenz innerhalb  
4464 dieser Debatten in Verbindung mit dem Beharren auf Verwei-  
4465 gerung ist von zentraler Bedeutung.

Der Kampf um die Rechte der indigenen Bevölkerung, sowie Gerechtigkeit für ebendiese ist ein laufender. Audra Simpson schreibt in ihrem Artikel über die Anatomie der Verweigerung, dass der Kampf um Land, Rechte und Geschlechtergerechtigkeit zentrale Punkte des indigenen Nordamerikas sind (2017: 6). Sie schreibt weiterhin, dass sich hierin das politische Spiel mit Sprache und die vom Staat initiierte Aktionskunst – welche darauf abzielen Teile der Geschichte zu verändern, um diese verarbeiten und abschließen zu können und somit einen Teil der Geschichte so zu verändern, dass man daraus in einen besseren Teil übergehen kann – widerspiegelt. Diesen Prozess beschreibt sie als historische Ungerechtigkeit („historical injustice“) (Simpson 2017: 6-7). Diese Umstellung von Geschichte zielt auf die Verdrängung und Re-Marginalisierung der indigenen Bevölkerung ab, indem die Gewalterfahrungen, die diese Mechanismen erzeugen, aktiv vermieden und somit negiert werden, was dazu führt, dass die indigene Bevölkerung wieder als kolonialisiertes Subjekt hervorgehoben wird. Diese Ungerechtigkeit wird durch Künstliche Intelligenzen und Machine-Learning-Systeme weiterhin aktiv praktiziert, da diese in rechnergestützten Modellen, welche vom omnipräsenten *white gaze* programmiert werden, wieder hergestellt wird.

Machine Learning Systeme werden vom *white gaze* umgeschrieben und erschaffen so eine nicht enden wollende Feedbackschleife welche dazu beiträgt, dass die Herrschaft des Siedler Kolonialismus weiterhin bestehen bleibt. Alles wird durch und von whiteness geformt, da whiteness immer die Idee bleibt, welche unbeachtet im Hintergrund mitarbeitet. Das bedeutet, dass whiteness mehr Raum einnimmt, da es als eine Form von öffentlicher Bequemlichkeit agiert (Ahmed 2007: 156). Whiteness wird zu einem Scheidepunkt an dem sich die Welt entfalten kann, während es gleichzeitig als hintergründige Erfahrung fungiert (Ahmed 2007: 150).

In ihrem Artikel beschreibt Ahmed, wie sich Körper an ihrer Umgebung orientieren und sich so in eine spezifische Richtung entwickeln. Dies gilt vor allem innerhalb von institutionalisierten Räumen. Sie erforscht weiterhin die Idee, dass whiteness real, materiell und gelebt ist und wird (2007: 150). Weil whiteness omnipräsent ist und somit nicht-weiße Individuen durch das Reale, Materielle und Gelebte beeinflusst, werden diese Ideen in Künstlichen Intelligenzen eingebaut, da whiteness als seine eigene Omnipräsens innerhalb von sozialen, politischen und institutionellen Strukturen agiert. Zusätzlich haben Künstliche Intelligenzen das Konzept von whiteness, durch die verschiedenen Formen in denen wir ebendiese erleben, vererbt bekommen.

Das Reale, Materielle und Gelebte sind eingebaute Segmente in Künstlichen Intelligenzen hinsichtlich ihrer allgegenwärtigen Orientierung in der Welt. Unsere Erlebnisse und Erwartungen werden durch die Verbreitung von whiteness innerhalb von Künstlichen Intelligenzen geformt und geleitet. Ein spezifisches Beispiel hierfür ist Apple's watchOS 6, welche Fitness- und Ernährungstracking, einen Herzfrequenzmesser mit dem man etwaige Unregelmäßigkeiten der Herzfrequenz erkennt, die Messung des Menstruationszyklus, sowie die übliche Siri Steuerung, Nachrichtenservices und vieles mehr, enthält. Die watchOS 6 ist eine von vielen Künstlichen Intelligenzen die neben dem lebenden Körper agieren und somit als eine Hintergrunderfahrung fungieren – als real, materiell und gelebt. Unsere biologischen Realitäten werden durch Künstliche Intelligenzen wie die watchOS 6 geformt, welche nur eine von vielen Beispielen dafür ist, inwieweit wir



uns mit ihr und durch sie orientieren. Diese Orientierung basiert darauf wie wir die Daten, die sie zur Verfügung stellt, benutzen, mit ihnen interagieren und uns auf sie verlassen.

Darüber hinaus sind die Daten, die durch Modelle wie watchOS 6 verwaltet werden, nicht intersektional angelegt, da sie die Genetik eines indigenen Körpers oder anderer nicht-weißer Körper nicht berücksichtigen (und wie sie sich z.B. von einem weißen Körper/den Gesundheitsmaßen eines weißen Körpers unterscheiden) und trotzdem wird die Maschine mit Modellen programmiert, die den Siedler-Kolonialismus als Grundlage haben.<sup>3</sup> Darüber hinaus prägen die Informationen, die uns mit Hilfe dieser Modelle als Interpretationen und Darstellungen der indigenen Bevölkerung zugeführt werden, unsere Erfahrung und unser Verständnis des kolonisierten Subjekts. Dies wirft die Frage auf: Wie sehen sich die indigenen Völker selbst in digitalen Formen jenseits des *white gaze*? Da sich die Künstliche Intelligenz an dem Punkt manifestiert, von dem aus wir mit ihr interagieren und so die Re-Präsentation ihres Outputs aus einer siedlerkolonialen Perspektive erleben, breitet sie sich auch auf andere Formen maschineller Lernprozesse wie Style Transfer und GAN aus. Das Bild der Indigenität wird durch *whiteness*, die sich auf notwendige Formen der Verweigerung beruft, ständig erneuert.

3 „By tracing the way in which race, gender, and other identities are understood, represented, and reflected, both within AI systems, and in the contexts of where they are applied, we can begin to see the bigger picture: one that acknowledges power relationships, and centers equity and justice.“  
Green, Ben. (2020) Data Science as Political Action:  
Ground Data Science in a Politics of Justice.

# RE-PRÄSENTATION, STILTRANSFER UND GAN

Schemata der Re-Präsentation von Bildern des kolonisierten Subjekts stoßen auf verschiedene Herausforderungen, da es dem Re-Präsentativen an Authentizität mangelt und ontologische Charakteristika sowie eine Geschichte des Subjekts, unabhängig vom Ergebnis, im Prozess verloren gehen. Darüber hinaus kommt es bei dem Versuch das Originalbild einer indigenen Frau darzustellen, zu ontologischer und historischer Gewalt, da der Körper und die Geschichte dieses Körpers seziert werden. Die Rekonstituierung einer indigenen Frau in ein Bild verursacht automatisch die Abwesenheit von Indigen-

tät durch den *white gaze*. Diese Praxis negiert jede Form von historischem Trauma, das den indigenen Völkern von Generation zu Generation weitergegeben wird und verursacht eine andere Form von Trauma, welches der indigene Körper überwinden oder ablehnen muss. Die ontologische Erfahrung des kolonisierten Subjekts (die indigene Person, die durch den *white gaze* neu dargestellt wird) wird verworfen und neu interpretiert. Guerin und Hallas untersuchen dies genauer:

Es hat sich eingebürgert, zu akzeptieren, dass die Ontologie des Bildes eine Unmittelbarkeit und Präsenz einnimmt, sodass das Bild einen bemerkenswerten Abgleich der gelebten Realität darstellt. Es wird davon ausgegangen, dass Bilder nicht einfach nur die Gewalt und das Trauma des Ereignisses evozieren, sondern es neu darstellen, es wieder präsent machen. (2007: 9)

Eine materielle und digitale Übertragung eines Subjekts in eine Re-Präsentation entfernt die Geschichte und Ontologie des Subjekts. Die zu fotografierende Person oder die Fotografie/das Bild, die/das noch einmal interpretiert wird, bleibt unberücksichtigt. Diese Rekonstitution stellt die Verletzung der Siedler-Kolonialisierung dar, die die indigenen Völker historisch erfahren haben. Die Re-Präsentation der Indigenität durch den *white gaze* wird nur eine Ungerechtigkeit erzeugen. Auf diese Weise dehnt sich der Kampf der Indigenen über Formen materieller, politischer Ungerechtigkeit hinaus aus. Das Bild reproduziert Gewalt. Zusätzlich wird der *white gaze* dadurch verstärkt, dass ein Bild des kolonisierten Subjekts aus der Sichtweise des Siedlers wieder präsent wird. Diese historische Ungerechtigkeit wird noch verstärkt, wenn diese Re-Präsentationen durch Machine-Learning-Modelle verarbeitet werden, die *whiteness* als die dominierende Perspektive voranstellen, in dem der Körper der Indigenen wahrgenommen, kategorisiert und neu dargestellt wird.

Zwei Modelle des maschinellen Lernens, die Bilder aufnehmen, interpretieren und als dekodierte Ergebnisse wiedergeben, sind Style Transfer und GAN. Style Transfer ist die Technik der Rekonstitution des Originalbildes in den ‚Stil‘ anderer Bilder. Darüber hinaus entwickeln und extrahieren tiefe neuronale Netze, die im Rahmen von maschinellen Lernprozessen verwendet werden, eine Repräsentation des „Stils“ eines Bildes und übertragen diese anschließend auf zukünftige Bilder. Indem wir ein tiefes neuronales Netz darauf trainieren, eine Darstellung eines Bildes zu extrahieren, wenden wir den Stil (S) auf das Inhaltsbild (I) an und erstellen ein neues Bild (IS – Inhalt von I mit dem Stil von S). Style Transfer wird zum Beispiel für kommerzielle Kunst, Spiele und virtuelle Realität verwendet. Eine GAN ist ein generatives antagonistisches Netzwerk. Es besteht aus zwei neuronalen Netzen, dem sogenannten Generator, der die Eingabedaten erzeugt, und dem sogenannten Diskriminator, der den Daten sagt, welche Daten „echt“ und welche „gefälscht“ sind. Allgemeiner gefasst, erstellen die GANs ihre eigenen Schulungsdaten auf der Grundlage ihrer ursprünglichen Eingabe. Sowohl ein Style Transfer als auch ein GAN sind Machine-Learning-Modelle, die ein komprimiertes Bild auf der Grundlage der Datenverarbeitung des Machine-Learning-Modells und des Lernens an sich ausgeben. Mit anderen Worten entwickeln sie ein komprimiertes Bild (Ausgabe) auf der Grundlage eines ursprünglichen Eingabebildes (Daten). GANs werden zum Beispiel für die Erzeugung eines Bildes aus Text, die Konvertierung eines Schwarz-Weiß-Bildes in Farbe oder das Ausfüllen des Umrisses eines Bildes verwendet.

4625 In der Abbildung unten wurde der GAN verwendet,  
 4626 um aus einer Textbeschreibung ein Bild zu generieren. Dieses  
 4627 Beispiel wurde mit einem GAN namens AttnGAN aus der Pro-  
 4628 grammanwendung RunwayML erstellt. Dieser GAN ist ein  
 4629 Text-zu-Bild-Verarbeitungsmodell. Nachfolgend sind die Ergeb-  
 4630 nisse des GAN mit dem Text „Indigenous Woman“ dargestellt:



Figure 1 Indigenous woman<sup>4</sup>

4631 AttnGAN entschlüsselte die Worte „Indigenous  
 4632 Woman“ und erzeugte ein Bild, das auf früheren Daten basiert,  
 4633 die im Rahmen des Machine Learning Modells gespeichert  
 4634 wurden. Der Zweck von AttnGAN besteht darin die Gewalt  
 4635 darzustellen, die auftritt, wenn ein Machine Learning Modell  
 4636 ein Bild rekonstruiert, das das kolonisierte Subjekt re-präsen-  
 4637 tieren muss (es wird hier nicht verwendet, um Vergleiche mit  
 4638 anderen Ergebnissen herzustellen, die whiteness wieder in  
 4639 den Fokus rücken würden). Darüber hinaus zeigt das obige  
 4640 Ergebnis ein verzerrtes und abstraktes Bild, dem die meisten  
 4641 Merkmale fehlen die notwendig sind, um die Ontologie, das  
 4642 Erbe usw. des Subjekts zu erkennen. Es vermindert jegliche  
 4643 ontologische oder kosmologische Geschichte, die mit der  
 4644 Realität, was es bedeutet, eine indigene Frau zu sein, wie es  
 4645 aussieht, eine indigene Frau zu sein und so weiter, zusammen-  
 4646 hängt. Obwohl historisch gesehen, Bilder von indigenen  
 4647 Frauen, deren Indigenität fetischisieren, indem sie z.B. die „ein-  
 4648 heimischen“ Merkmale und das zum Stamm des Subjekts  
 4649 passende Gewand hervorheben, verhält sich der white gaze  
 4650 gewaltsam, indem er die Gewalt die durch die Reproduktion  
 4651 des Bildes angewandt wird, gar nicht erst anerkennt. Wie  
 4652 Guerin und Hallas ausführen, „beruhen selbst erlösende Dar-

stellungen des Bildes und der visuellen Repräsentation auf der Dekonstruktion, Aneignung oder Resignation bestehender historischer Sehweisen“ (2007: 4). Diese „historische Sehweise“ ist eine aktive Form der Siedler Kolonisierung und das Bild einer indigenen Frau, gesehen durch den white gaze, lässt die Schäden der Siedler Kolonisation fortbestehen. Darüber hinaus verkennt das Beispiel die Ontologie und Geschichte der indigenen Frau, welches die Ungerechtigkeiten aufzeigt, die die indigenen Völker erfahren haben und welche durch Machine Learning Modelle beschnitten werden. Sie werden entweder falsch erkannt oder fehlen vollkommen und fordern uns auf, darüber nachzudenken, wie wir uns dem Siedler-Kolonialismus widersetzen und ihn ablehnen können.

# VERWEIGERUNG & UNDURCH- SICHTIGKEIT

Da diese historischen Ungerechtigkeiten in Machine-Learning-Modellen präsent sind und zu den „realen“, gelebten und materiellen Teilen unseres Lebens gehören, sind die indigenen Völker gezwungen, ihre eigene Repräsentation durch Formen der Verweigerung neu zu orientieren. Tuck und Yang führen uns in die Verweigerung ein, indem sie fragen, „wie wir von der Weisheit und den Wünschen in den Geschichten, die wir (über)hören, lernen und sie respektieren, während wir uns weigern, sie dem Spektakel des kolonialen Blicks der Siedler zu zeigen/verraten“ (2014: 222). Da das maschinelle Lernen der technologische Punkt ist, aus welchem sich *whiteness* entfaltet, möchte ich etwas Ähnliches fragen:

WIE VERWEIGERN WIR, DIE INDIGENEN VÖLKER,  
FÄLLE DER RE-PRÄSENTATION MIT HILFE VON  
MACHINE-LEARNING-MODELLEN?

Tuck und Yang schreiben, dass die Verweigerung eine „Art von Untersuchung dessen ist, was Sie wissen müssen und was ich mich weigere zu schreiben“ (qtd. in Tuck und Yang 2014: 223). Indigene Völker werden oft nicht in die Forschung über maschinelles Lernen einbezogen, was ihre Präsenz in diesen Dialogen unabdingbar macht. Da „die Verweigerung den Blick zurück auf die Macht lenkt – insbesondere auf die kolonialen Modalitäten, Personen als Körper zu kennen, die differenziert gezählt, verletzt, gerettet und eingesetzt werden müssen“, ist die Wirkung der indigenen Präsenz innerhalb der Sektoren der Künstlichen Intelligenz besonders relevant (ebd. 2014: 241). Wenn die indigenen Völker danach die rechnerischen Zwischenräume und ihre vorhersagbaren Motivationen kennen lernen und sie als solche ausnutzen, gewinnen sie ihre Macht über Re-Präsentation zurück und gelangen zu

4696 einer Art kolonialer Subversion der Siedler. Verweigern, unter-  
 4697 graben, überleben. Darüber hinaus werden die indigenen  
 4698 Stimmen, die in die computergestützte Forschung und Unter-  
 4699 suchung einbezogen werden, die siedlerkolonialen Modi, die  
 4700 bereits in den Modellen des maschinellen Lernens existieren,  
 4701 nicht aufhalten; die darin enthaltene Erkenntnis wird den indi-  
 4702 genen Völkern jedoch die Möglichkeit bieten, den *white gaze*  
 4703 der Künstlichen Intelligenzen abzulehnen.

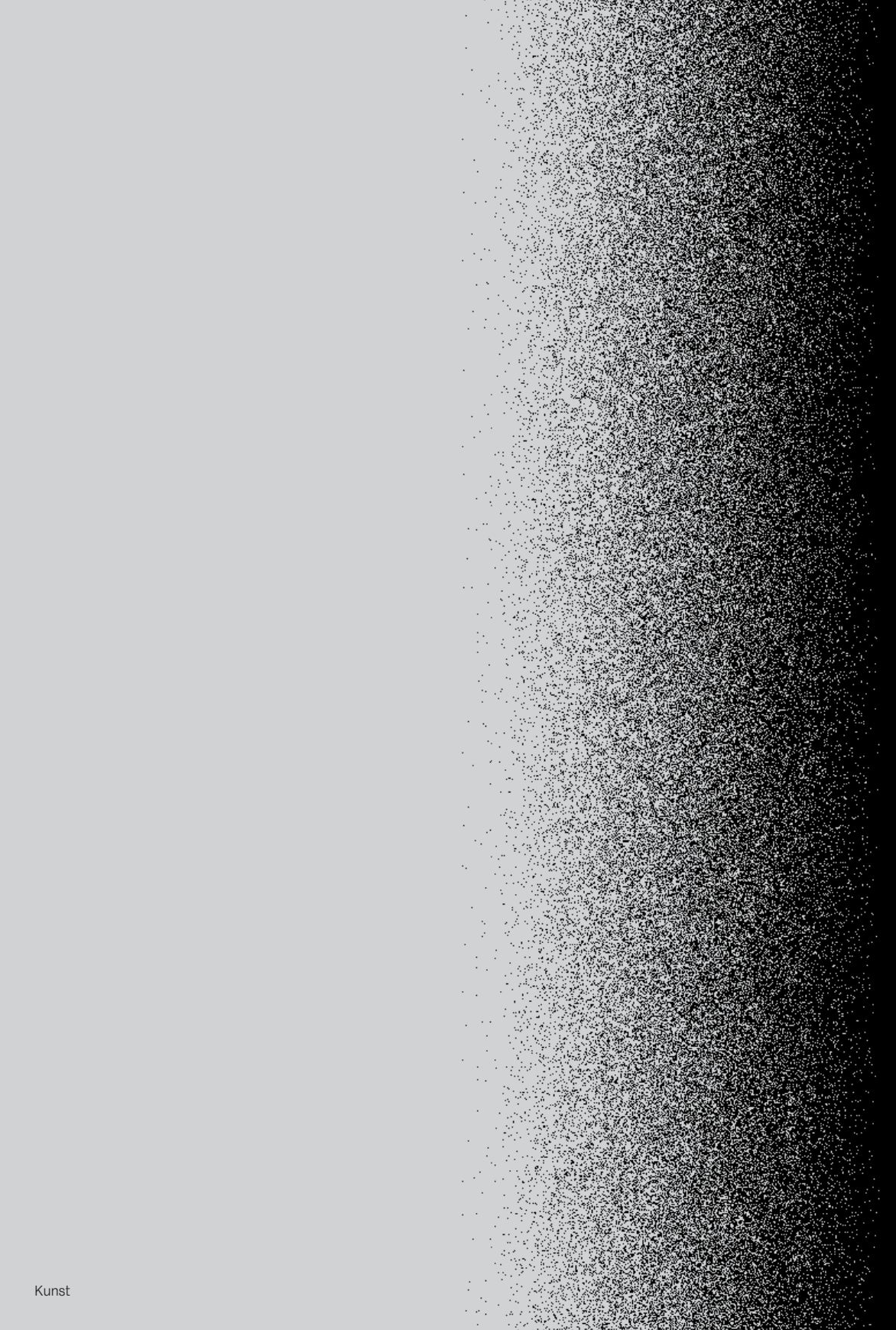
Dr. Tiara Roxanne ist indigene Cyberfeministin, Wissenschaftlerin und Künstlerin. In ihren wissenschaftlichen und Künstlerischen Arbeiten (Writing and Performance) untersucht sie, wie koloniale Strukturen in Künstlichen Intelligenzen eingeschrieben sind. Ihre Dissertation trägt den Titel „Recovering Indigeneity: Territorial Dehiscence and Digital Immanence“ (2019). Aktuell arbeitet sie am DeZIM-Institut.

# LITERATUR

- \_\_\_\_\_ Ahmed, Sara (2007)  
 „A Phenomenology of Whiteness.“ *Feminist Theory*, 8 (2), 149–168.,  
 doi:10.1177/1464700107078139.
- \_\_\_\_\_ „CycleGAN: The Artificial in Artificial Intelligence.“  
*Mc.ai*, 3 Feb. 2020, mc.ai/cyclegan-the-artificial-in-artificial-intelligence/.
- \_\_\_\_\_ Green, Ben (2020)  
 „Data Science as Political Action: Grounding Data Science in a Politics  
 of Justice,“ 1–40. ArXiv abs/1811.03435.
- \_\_\_\_\_ Guerin, Frances & Hallas, Roger (ed.) (2007)  
*The Image And The Witness*. New York: Wallflower Press.
- \_\_\_\_\_ Rainie, Stephanie C.; Kukutai, Tahu; Walter, Maggie;  
 Figueroa-Rodríguez, Oscar L.; Walker, Jennifer & Axelsson, Per (2019)  
 „Issues in open data: Indigenous data sovereignty.“ in: Davies, Tim;  
 Walker, Stephen B.; Rubinstein, Mor & Perini, Fernando (Ed.),  
*The state of open data: Histories and horizons*. Cape Town  
 and Ottawa: African Minds and International Development  
 Research Centre. 300–319. <http://stateofopendata.od4d.net>.
- \_\_\_\_\_ Roxanne, Tiara (2020)  
 „Data Colonialism: Decolonial Gestures of Storytelling,“ *donaufestival  
 reader*, 147–159.
- \_\_\_\_\_ Simpson, Audra (2017)  
 „The ruse of consent and the anatomy of ‚refusal‘: cases from indi-  
 genous North America and Australia,“ *Postcolonial Studies  
 Journal*, doi: 10.1080/13688790.2017.1334283.
- \_\_\_\_\_ *Style Transfer*  
[docs.fritz.ai/develop/vision/style-transfer/](https://docs.fritz.ai/develop/vision/style-transfer/).
- \_\_\_\_\_ Tuck, Eve, & K. Wayne Yang. (2014)  
 „R-Words: Refusing Research.“ *Humanizing Research: Decolonizing  
 Qualitative Inquiry with Youth and Communities*, 223–248.,  
 doi:10.4135/9781544329611.n12.
- \_\_\_\_\_ WatchOS 6 Vorschau (2020)  
*Apple (Deutschland)*, [www.apple.com/de/watchos/watchos-6/](http://www.apple.com/de/watchos/watchos-6/).
- \_\_\_\_\_ West, Sarah Myers; Whittaker, Meredith & Crawford, Kate (2020)  
*Discriminating Systems: Gender, Race and Power in AI*. AI Now Institute.







11

AI  
:  
QUER ART



KATRIN  
KÖPPERT

4704  
4705  
4706  
4707  
4708  
4709  
4710  
4711  
4712  
4713  
4714  
4715  
4716  
4717  
4718  
4719  
4720  
4721  
4722  
4723  
4724  
4725  
4726  
4727  
4728  
4729  
4730  
4731  
4732  
4733  
4734  
4735  
4736  
4737  
4738  
4739  
4740  
4741  
4742  
4743  
4744  
4745  
4746

„I believe that at the end of century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.“  
(Turing 1950)

Mit Verlaub: Ich glaube, dass sich zum Ende dieses Artikels der Gebrauch von Wörtern dergestalt geändert haben wird, dass Sie in der Lage sein werden, von Künstlicher Intelligenz als *queer art* zu sprechen und stelle in Aussicht, dass Sie sich in dieser Meinung – sollten Sie je danach gefragt werden – nicht herausgefordert fühlen müssen. Wie Alan Turing (nur mit umgekehrten Vorzeichen) möchte ich am Sprechen festmachen, dass es einen anderen Diskurs der Künstlichen Intelligenz braucht.<sup>1</sup> Dabei wird es nicht einmal nötig sein, Zukunft zu bemühen. Der andere Diskurs ist der Künstlichen Intelligenz schon längst inhärent. Nur müssen wir ihn – um ihn Wirklichkeit werden zu lassen – am Sprechen festmachen.

In Anbetracht der diskriminierenden, trans\* und homophoben Effekte von Algorithmen nimmt sich die Behauptung Künstlicher Intelligenz als queere Kunst zunächst als Provokation aus. Aber sie wird sich verschleifen. Sie wird sich in ihr Gegenteil verkehren, sodass wir uns am Ende eher dadurch provoziert fühlen werden, dass unter der Prämisse, Maschinen das menschliche Denken für noch mehr Arbeitsproduktivität und ökonomische Leistungssteigerung beibringen zu wollen, verschwenderisch viel Geld in ausschließlich diese wirtschaftsorientierte Forschung fließt.<sup>2</sup>

Denn – um das gleich einleitend aus dem Weg zu räumen – Maschinen bzw. Computer denken nicht. Zumindest nicht in der Weise, wie es ihnen im Rahmen der nunmehr über 50 Jahre währenden Rezeption Künstlicher Intelligenz unterstellt wird. Maschinen – übrigens auch Menschen – denken nicht, wenn Denken Berechenbarkeit meint. Denn schon die Grundlagenkrise der Mathematik verdeutlichte Unvollständigkeiten der Berechenbarkeit und Unentscheidbarkeiten, die den Algorithmen logische Grenzen setzen (Mersch 2019: 68). Hier docke ich an, um im weiteren Verlauf zu verdeutlichen, inwiefern allein schon die den Algorithmen zugrundeliegende Mathematik *queer art* ist. Dabei wird zu sehen sein, dass die Mathematik bzw. Technologie der Künstlichen Intelligenz mit der Technologie des Geschlechts verbunden ist. Das wiederum lässt Aussagen darüber zu, welche künstlerischen Konzepte der Künstlichen Intelligenz eingeschrieben sind.

1 Hiermit lehne ich mich an eine der drei Interpretationen des Turingtests an. Sie besagt, dass der Turingtest eine sprachnormative Bedeutung hat. Demnach interessierte sich Turing weniger für die Frage, ob Maschinen wirklich denken können, als für die, unter welchen Bedingungen sie als intelligent bezeichnet werden (vgl. Heintz 1993: 273).

2 Symptomatisch für diese unausgewogene Verteilung von Forschungsgeldern sind die vielen staatlich und EU-subventionierten Digitalisierungsprogramme vgl. <https://www.bundesregierung.de/breg-de/themen/themenseite-forschung/deutschland-soll-fuehrender-ki-standort-werden-1615312>.



# GRENZEN DER BERECHENBARKEIT

Seitens der Philosophie gibt es mit Beginn der Behauptung, das menschliche Gehirn gleiche einem Rechner (vgl. Neumann 1991; Schmidhuber 2012) bzw. der Rechner dem für Rationalisierung einstehenden menschlichen Gehirn, zahlreiche Einwände. Sie berufen sich darauf – und hier folge ich der kurssrischen Zusammenschau von Dieter Mersch (2019: 67f.) –, dass Denken und Rechnen nicht gleichgesetzt werden könne, weil dem rechnenden System die Fähigkeit fehle, die Ergebnisse auf ihre Sinnhaftigkeit zu überprüfen (Heidegger 1984). Ihnen fehlt das Bewusstsein, auch weil jenes eine Funktion des gesamten Körpers ist, und nicht die eines vom Körper abgekoppelten Gehirns (Dreyfus 1989: 183ff.; Noë 2011; Gabriel 2018: 91ff.). Zudem hänge Denken von einem Verständnis praktischer Konsequenzen ab, was die Unterscheidbarkeit zwischen relevanten und irrelevanten Implikationen voraussetze (Brandom 2008: 77–82; 117–140). Solange eine Künstliche Intelligenz auf der formalen Algorithmik beruhe, sei sie zur Abschätzung dieser Konsequenzen nicht in der Lage. Auch würde sie – sofern auf der formalen Methode des Turing-Tests beruhend – an der Fähigkeit scheitern, sich zu verändern und multiple Standpunkte einzubeziehen (Star 2017: 135). Hinzu kommt, dass computerisierte ethische Abwägungen, zum Beispiel im Feld der Gesichtserkennung, vor der Differenz zwischen Denken und Rechnen kapitulieren (vgl. Lin et al. 2017).

Diesen philosophischen Kritiken gemein ist – so Mersch –, dass sie außerhalb der Mathematik an sich argumentieren, das heißt, dass sie im Rahmen von Informatik samt ihrer lösungsorientierten Anwendungen, bzw. empirischen Auswirkungen, problematisieren, was Rechnen von Denken unterscheidet. Dies könne aber bereits innerhalb der mathematischen Fundamente erfolgen. Durch die Betrachtung der Mathematik eröffne sich bereits die Kritik an algorithmischer Rationalität (Mersch 2019: 68). Dabei zentral ist, dass sich die Mathematik nicht einer durchgängigen Berechenbarkeit und Arithmetisierung von Problemen fügt. Als Theorie auch von Räumen und Relationen bzw. als Theorem der Unvollständigkeit z.B. nach Kurt Gödel<sup>3</sup> bildet die Mathematik sich selbst ihre eigene Grenze. Im Sinne der Unmöglichkeit ihrer eigenen Mathematisierbarkeit ist Mathematik das innere Außen und als solches der Grund, warum Künstliche Intelligenz als maschinelle Rechenoperation mit ihrer eigenen Unvollständigkeit konfrontiert bleiben muss.

Selbst mit dem „Modell einer Turingmaschine als allgemeine[r] Theorie des Algorithmus [...], ergibt sich ein nicht zu schließender Abstand zwischen Berechenbarkeit und Nichtberechenbarkeit, der *als Differenz* nicht selbst wieder einer Algorithmisierung zugeführt werden kann“ (Mersch 2019: 71). Und noch im Versuch Turings, den Abstand abmildern zu wollen, orakelt es mächtig. Turing hatte, um die Unvollständigkeiten abzuschwächen, die Erweiterbarkeit vollständiger Systeme um andere vollständige Systeme vorgeschlagen, ohne jedoch die Übergänge zwischen den Systemen einer maschinellen Berechenbarkeit zuführen zu können. Sie würden ein –



so Turing – „Orakel“ (Turing 1939: 172f.) bleiben. Mersch ergänzt hier und meint, dass die Übergänge einzig auf „Intuition“ oder „kreativer Erfindung“ (Mersch 2019: 72) beruhten. Hiermit wären wir nun spätestens an einer Stelle angelangt, an die sich die den Begriffen der Intuition, des Orakels und der Kreativität unterliegenden Geschlechterdiskurse förmlich anschließen müssen. Dass maschinelles Lernen ein Geschlecht hat, erfahren wir somit nicht nur aus den filmischen und literarischen Imaginationsgeschichten,<sup>3</sup> sondern allein schon, wenn wir uns mit den Vergeschlechtlichen der Algorithmen in der Mathematik beschäftigen. Im Anschluss an diese – wenigleich kurssorische – Auseinandersetzung mit der Vergeschlechtlichung wird sich auflösen, wie ich zu der Aussage komme, dass Mathematik das Potenzial von *queer art* im Sinne künstlerischer Verfahren des Durchquerens binärer Anordnungen impliziert. Dreh- und Angelpunkt ist dabei der Turing-Test.

3 „[D]ie Gödeltheoreme besagen, dass sich in jedem formalen System, das mächtig genug ist, die Arithmetik zu umfassen, wahre Aussagen formulieren lassen, die sich weder beweisen noch widerlegen lassen.“ (Mersch 2019: 70).

4 Hier seien u.a. E.T.A. Hoffmanns Olimpia, Fritz Langs Maria, Spike Jonzes Samantha und Alex Garlands Ava erwähnt.

# UNBERECHENBARES GESCHLECHT

## — GESCHLECHTLICHE UNBERECHENBARKEIT

Geschlecht ist nicht der Subtext des vielleicht wirkmächtigsten Textes der Computergeschichte, sondern der Text überhaupt. In „Computing Machinery and Intelligence“ lotet Alan Turing (1950) die Frage, ob sich Computer in ihrem Verhalten von einem Menschen unterscheiden lassen, auf der Grundlage eines Geschlechter-Rate-Spieles aus. In dessen Zentrum steht die Figur der Imitation und die Aufgabe des Mannes, eine Frau mittels schriftlicher Antworten so zu simulieren, dass er vom Fragestellenden nicht als Mann identifiziert wird. Dieses Gedankenspiel dient Turing als Vorlage, um im weiteren Schritt zu plausibilisieren, dass der Computer erst dann als Mensch erraten würde, wenn er in der Lage sei, menschliches Verhalten perfekt zu simulieren.

In den Interpretationen des Turing-Tests wird das Geschlechterspiel oft unterschlagen. Sie scheuen – so Bettina Heintz – davor zurück, die Überlegung, eine Maschine

könne menschliches Verhalten perfekt simulieren, auf das Geschlechterspiel zu übertragen:

„Eine[r] Maschine, die sich in ihrem Verhalten nicht von einem Menschen unterscheidet, werden menschliche Qualitäten attestiert; ein Mann, der weibliches Verhalten perfekt simuliert, bleibt immer noch ein Mann“  
(Heintz 1993: 264)

Anders als der Biograf Turings, Andrew Hodges, der zu dem Schluss kommt, das Geschlechterspiel sei eine „Finte“ und hätte keine Aussagekraft insofern, als die Imitation des anderen Geschlechts nicht dazu führe, Qualitäten dieses anderen Geschlechts zu attestieren (Hodges 1989: 480), scheint Turing selbst dieser Aussagekraft Glauben zu schenken. Vielmehr ist die Annahme, Geschlecht sei nicht biologisch determiniert, die Grundlage für die Versuchsanordnung, Mensch durch Maschine zu substituieren. Im Grunde nimmt Turing hiermit vorweg, was im Zusammenhang der Queeren Theorie ab den 1990ern<sup>5</sup> ausformuliert wurde.

Nun ließe sich das darauf zurückführen, dass Turing aufgrund seiner Homosexualität einen Zugang zu geschlechtlicher Identität gehabt haben mag, der nicht auf einem deterministischen Biologiekonzept beruhte. Und tatsächlich werde ich darauf noch zurückkommen. Jedoch geht es mir hier erst noch einmal darum, am Denkmodell des Turings-Tests selbst nachzuvollziehen, dass Turing im spielerischen Versuch der Imitation die Geschlechterdifferenz nicht nur verschwinden lassen, sondern dahingehend verschieben will, Geschlecht und somit auch Künstliche Intelligenz immer neuen und nicht kalkulierbaren Zukünften gegenüber zu öffnen. Denn selbst wenn Horizont des Tests ist, zu zeigen, dass menschliches Verhalten in der gleichen Weise maschinell erlernt und formalisiert werden kann, wie geschlechtsspezifisches Verhalten, dann sind in der Formulierung des Tests Ambiguitäten eingebaut, die der Formalisierbarkeit zuwiderlaufen. Dem Versuch, maschinelles Denken mittels der Geschlechterfrage zu intelligibilisieren, ist von Anfang an die innere Grenze der Mathematik (des Geschlechts) eingezogen. Hiermit kommt der Begriff der Intuition wieder ins Spiel. Mit dieser ist im Geschlechterspiel die dritte anwesende Person in Verbindung gebracht: Turing geht neben Spieler A (der die Frau imitierende Mann) und C (die geschlechtlich nicht definierte fragestellende Person) von B (die Frau) als Kraft der Intuition aus. Die Rolle von B ist, C zu helfen, indem sie, ihrer Intuition folgend, wahrheitsgemäße Aussagen über ihr Geschlecht trifft und dadurch C davon abhält, A als Frau zu identifizieren. Die weitläufige Interpretation ist, B fungiere hier in Abgleich zu ihrer als natürlich angenommenen Weiblichkeit. Sie müsse sich nicht verstellen, sondern würde ihrem qua Biologie vorgegebenem Geschlechtscharakter entsprechend Rede und Antwort stehen. Intuition wird im Verständnis wahrhaftigen Verhaltens als weiblich naturalisiert und insofern delegitimiert, als B als Widersacherin des Imitationsversuchs installiert wird.

Dem möchte ich entgegnen, dass Turing, meinem Eindruck nach, mit Bedacht Formulierungen wählt, die einen Restzweifel an der Aussage, Weiblichkeit sei mit Natürlichkeit und Intuition gleichzusetzen, beibehalten. Er schreibt:

„The object of the game for the third player (B) is to help the interrogator. The best strategy for her is *probably* to give truthful answers“

(Turing 1950: 433, Herv. KK).

In dem Moment, in dem Turing exemplifiziert, dass die Aufgabe von B ist, wahrheitsgemäße und ihrem Geschlechts-naturrell entsprechende Antworten zu geben, revidiert er seine

Aussage sogleich und überlässt sie dem Feld der Vermutung. Es ist nur „wahrscheinlich“ die beste Strategie. Intuition ist nur „vermutlich“ natürlicher Ausdruck von Weiblichkeit. Insofern ist Intuition nicht nur eine Frage der queeren Dekonstruktion von Geschlechterdifferenz, sondern der Ästhetik, zumindest wenn wir unter Ästhetik vom mathematischen Kalkül entkoppelte kreative Erfindungen verstehen. Denn wenngleich die Antworten B's keiner Logik im Sinne maschineller Berechenbarkeit folgen, sind sie dennoch nicht mit unreflektierter Natürlichkeit zu verwechseln, sondern mit dem ‚Erfinden‘. Intuition – um nochmal auf Mersch zurückzukommen – wäre folglich eine Form der Ästhetik, die inmitten des Mathematischen nicht als irrationales Moment mäandert, sondern an Fragen der kulturellen Bedingungen reflexiv gebunden bleibt (Mersch 2019: 74). In diesem Sinne verstehe ich Künstliche Intelligenz als *queer art*. Über den Spieler der Intuition, der nur unvollständig im Paradigma natürlicher Weiblichkeit aufgeht, artikuliert das Gedankenspiel des Turing-Tests Künstliche Intelligenz an der Grenze des Formalisierbaren. Mathematik erweist sich dabei gegenüber Kunst und Kreativität gleichermaßen geöffnet, wie Geschlecht gegenüber nicht berechenbarer Gleichungen.

5 Zur Etablierung von Queer Theory in den 1990er Jahren (vgl. Butler 1991: 93–104; Kosofsky Sedgwick 2004; Warner 1993; Halperin 1997). Für den deutschsprachigen Raum vor allem Andreas Kraß (2003: 7–39) und Sabine Hark (2005: 285–303).

# RELATIVELY UNAMBIGUOUS

In Alan Turings Text, der den Turing-Test beschreibt, geht es um die Untersuchung der Frage, ob Computer denken können. Wie Ulrike Bergermann anmerkt, ist das Objekt der Untersuchung nicht die Maschine an sich, sondern die Frage (Bergermann 2018: 346). Dabei ersetzt er die Frage, ob Computer denken können, durch die, ob Computer denkbar sind, die sich in ihrem Verhalten nicht vom Menschen unterscheiden (Heintz 1993: 263). Interessant ist dabei, dass die Frage, welche die Ausgangsfrage ersetzt, „is expressed in *relatively* unambiguous words“ (Turing 1950: 433, Herv. KK). Das heißt, dass die Ursprungsfrage abgeändert wird und noch zudem in zweideutigen Worten zum Ausdruck kommt oder, um näher im Vokabular Turings zu bleiben, in „relativ unzweideutigen“. Die Relativierung von Eindeutigkeit unterstreicht in gewisser Weise, dass er die Frage in ihrer abgeänderten Form untersuchen möchte und die Experimentalanordnung in ihrem Ergebnis offen ist. Das Genre des Spiels hebt dies noch hervor. Nicht nur, dass sich hier ein Wissenschaftsverständnis mitteilt, das innerhalb einer vom Objektivitätsglauben bemächtigten Naturwissenschaftsforschung vor allem zu seiner Zeit selten war, ist das Spielerische auch bedeutsam dafür, sich ins Verhältnis zu setzen. Die durch das Spielen gestifteten Relationen, die

sich im Wort „relatively“ mit abbilden, implizieren offene Enden und Ambiguitäten. Ich denke, dass es Turing, so sehr er darum bemüht war, Denken zu formalisieren, immer auch um die Herstellung von Ambiguität und eines Verständnis von Künstlicher Intelligenz ging, das sich nicht im rein Logischen erschöpft, sondern das Spielerische, Erfinderische, Fantastische inkludiert. Für die These steht insbesondere auf der Ebene des Mäy King Patin, die in den sozialen und kulturellen Bedingungen Turings Begründungen findet, dass sich Turings Konzeptualisierung von Künstlicher Intelligenz Formen queerer Sensibilitäten eingeschrieben haben, die sich insbesondere auf der Ebene des Märchenhaften, Romantischen und Rätselhaften abspielen (King 2015: 18–46). Dabei müssen wir Turing gar nicht mal nur von seinem Ende her denken, also seines in sich rätselhaften Todes, der – wenngleich von Steve Jobs bestritten – das wohl mächtigste Konzernlogo unserer Zeit geprägt haben dürfte: den angebissenen Apfel (Turing, der großer Fan des Märchens Schneewittchen war, ist an einem Bissen eines mit Zyanid vergifteten Apfels gestorben, wobei noch immer unklar ist, ob es sich um Mord oder Selbstmord handelte). Stattdessen ließe sich beim Tod seiner ersten jugendlichen Liebe ansetzen. King argumentiert, dass die Erfahrung des plötzlichen Todes seiner Jugendliebe Christopher Marcom und der Unentscheidbarkeit dessen, jemals erfahren zu können, ob seine Liebe von Marcom erwidert wurde oder nicht, Anlass war, sich in seiner Dissertation mit den schöpferischen Grenzen der Berechenbarkeit auseinanderzusetzen (King 2015: 37). Diese finden sich – wie ich versucht habe zu verdeutlichen – im Gründungstext Künstlicher Intelligenz wieder und weisen Künstliche Intelligenz in ihrer queereren Kreativität aus. Dazu sollten wir beginnen, uns ins Verhältnis zu setzen. Mehr über Künstliche Intelligenz als *queer art* zu sprechen, könnte ein Anfang sein.

Dr. Katrin Köppert ist Juniorprofessorin für Kunstgeschichte/populäre Kulturen an der Hochschule für Grafik und Buchkunst Leipzig. Zu ihren Arbeitsschwerpunkten zählen Queer Media Theory, Digital Feminisms, Post- und Dekoloniale (Medien-)Theorien des Anthropozäns.

# LITERATUR

Bergermann, Ulrike (2018)

„biodrag. Turing-Test, KI-Kino und Testosteron“, in: Engemann, Christoph & Sudmann, Andreas (Hg.): *Machine Learning – Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz*, Bielefeld: transcript Verlag.

Brandom, Robert (2008)

*Between saying and Doing*, Oxford: Oxford University Press.

Butler, Judith (1991)

*Das Unbehagen der Geschlechter*, Frankfurt am Main: Suhrkamp Verlag.

Gabriel, Markus (2018)

*Der Sinn des Denkens*, Berlin: Ullstein Verlag.

- \_\_\_\_\_ Dreyfus, Hubert L. (1989)  
*Was Computer nicht können*, Frankfurt a. Main: Athenäum Verlag.
- \_\_\_\_\_ Leigh Star, Susan (2017)  
 „Die Struktur schlecht strukturierter Lösungen. Grenzobjekte und heterogenes verteiltes Problemlösen“, in: dies.; Gießmann, Sebastian & Taha, Nadine (Hg.): *Grenzobjekte und Medienforschung*, Bielefeld: transcript Verlag, 131–150.
- \_\_\_\_\_ Lin, Patrick; Jenkins, Ryan & Abney, Keith (Hg.) (2017)  
*Robot Ethics 2.0. From Autonomous Cars to Artificial Intelligence*, Oxford: Oxford University Press.
- \_\_\_\_\_ Halperin, David (1997)  
*Saint Foucault: Towards a Gay Hagiography*, New York: Oxford UP.
- \_\_\_\_\_ Hark, Sabine (2005)  
 „Queer Studies“, in: von Braun, Christina & Stephan, Inge (Hg.): *Gender@Wissen. Ein Handbuch der Gender-Theorien*. Köln / Weimar / Wien: Böhlau Verlag, 285–303.
- \_\_\_\_\_ Heidegger, Martin (1984)  
*Was heißt denken?*, Tübingen: Niemeyer Verlag.
- \_\_\_\_\_ Heintz, Bettina (1993)  
*Die Herrschaft der Regel. Zur Grundlagengeschichte des Computers*, Frankfurt am Main / New York: Campus Verlag.
- \_\_\_\_\_ Hodges, Andrew (1989)  
*Alan Turing. Enigma*, Berlin: Kammerer & Unverzagt Verlag.
- \_\_\_\_\_ King, Homay (2015)  
*Virtual Memory. Time-Based Art and the Dream*, Durham / London: Duke University Press.
- \_\_\_\_\_ Kosofsky Sedgwick, Eve (2004)  
*Epistemology of the closet*, Berkeley / Los Angeles: University of California Press.
- \_\_\_\_\_ Kraß, Andreas (2003)  
 „Queer Studies – eine Einführung“, in: dies. (Hg.): *Queer denken. Gegen die Ordnung der Sexualität*, Frankfurt am Main, 7–39.
- \_\_\_\_\_ Mersch, Dieter (2019)  
 „Kreativität und Künstliche Intelligenz“. *Zeitschrift für Medienwissenschaft* 21, 2/2019, 65–74.
- \_\_\_\_\_ Noë, Alva (2011)  
*Du bist nicht dein Gehirn. Eine radikale Philosophie des Bewusstseins*, München: Piper Verlag.
- \_\_\_\_\_ von Neumann, John (1991)  
*Die Rechenmaschine und das Gehirn*, München: Oldenbourg Verlag.
- \_\_\_\_\_ Schmidhuber, Jürgen (2012)  
 „A Formal Theory of Creativity to Model the Creation of Art“, in: McCormack, Jon & d'Inverno, Mark (Hg.): *Computers and Creativity*, Berlin / Heidelberg: Springer Verlag, 323–337.
- \_\_\_\_\_ Turing, Alan (1950)  
 „Computing Machinery and Intelligence“, in: *Mind* 59, 433–460.
- \_\_\_\_\_ Turing, Alan (1939)  
 „Systems of Logic based on Ordinals“. *Proceedings of the London Mathematical Society* 2, 45/3, 161–228.
- \_\_\_\_\_ Warner, Michael (1993)  
*Fear of a Queer Planet. Queer Politics and Social Theory*, Minneapolis: University of Minnesota Press.
- \_\_\_\_\_ Video Podcast  
 „Deutschland soll führender KI-Standort werden“, in: Presse- und Informationsamt der Bundesregierung, online abrufbar unter:  
<https://www.bundesregierung.de/breg-de/themen/themen-seite-forschung/deutschland-soll-fuehrender-ki-standort-werden-1615312>. [abgerufen am 30.10.2020]







12

K. RÜPPEL  
INTELLIGENZ



KATHARINA  
KLAPPHECK

Künstliche Intelligenz (KI) ist be:hindert, war es und wird es sein. Sie ist verkrüppelt im besten Sinne. Wenn sie in ihrer Geschichte analysiert wird, wird auf Geschlecht und das Militär verwiesen und auch auf race. Der weiße Mann als Witz des neuen Zeitalters, ist er doch immer noch auf Queers und BiPoCs angewiesen, ohne die Daten nicht fließen, die Maschinen nicht gefüttert werden könnten. Aber die KI wird einfach nicht selbständig, selbst nach über 70 Jahren, mehreren Revolutionen und einem Großaufgebot an Imagination und cis-männlicher Ernsthaftigkeit stottert die KI, versteht immer noch keinen Humor und eigenständig leben kann sie nicht. Und immer noch schallt es aus den Elfenbeintürmen dieser Welt: „Wir lösen das Problem. Wir schaffen die Arbeit des weißen Arbeiters ab! Vertraut uns!“

Ich als Krüppel verstehe diese seltsamen Wesen aus Geschichte, Kapitalismus und menschlichen Tränen manchmal aus dem tiefsten Inneren meines falsch verkabelten Gehirns. Im weiteren Verlauf möchte ich in dem kurzen Text die Beziehung von KI und Be:hinderung verdeutlichen. Es werden meinerseits vier Komponenten dieser seltsamen Kopplung hervorgehoben. Einerseits die Frage der Künstlichen Intelligenz als Heilmittel für Be:hinderung. Daran anschließend wird es einen Teil geben, der wiedergibt, inwiefern KI als be:hindert in der Forscher\*innengemeinschaft genutzt wird. In einem dritten Teil möchte ich darstellen, inwiefern KI nicht einfach mit einem be:hinderten Menschen verschmilzt, sondern das Zusammenwirken beider ein komplexer Prozess ist, um abschließend auf die Frage einzugehen, wer überhaupt KI-Wissenschaftler\*in werden kann. Dabei ist es mir wichtig aufzuzeigen, dass die von Forscher\*innen immer wieder beschriebene Grenze zwischen Körper und Geist immer auch eine Grenze zwischen einem nicht be:hinderten Körper und einem nicht be:hinderten Geist mit einschließt. Be:hinderung ist dabei eine Grenze, die Technologie vorantreibt und präzisiert.

Ich will diese artifizielle Nähe lückenhaft erläutern und darlegen, wieso KI verwandt ist mit Be:hinderten. Es ist eine Geschichte der löchrigen und nicht zusammenpassenden Verbindung von künstlich und Intelligenz. Ich möchte nicht zeigen, wo Diskriminierung auftritt, die Menschen, die betroffenen sind, wissen es und die anderen dürfen gerne das Internet um Rat fragen. Ich möchte anreißen, in welcher multipler Art und Weise Crips<sup>1</sup> mit KI verbunden sind, welche Unheimlichkeiten sich ergeben. Was passiert, wenn wir der patriarchalen und rassistischen Geschichte von KI auch noch Be:hinderung hinzufügen? Ich werde von „wir“ reden, obwohl es kein „wir“ gibt. Manchmal möchte ich mich dahinter verstecken, weil es angsteinflößend sein kann seinem behinderten Selbst gegenüberzutreten. Aber noch viel wichtiger: Dieser Text verdankt sich meiner be:hinderten Familie. Ich verdanke ihnen andere Crips, die mir sagen, dass ich nicht alleine bin, dass ich keinem/keiner erklären muss, was Be:hinderung ist oder welche ich habe. Crips, die mir versichern, dass es schlechte Tage geben darf, dass es nicht perfekt sein muss, dass ich privilegiert bin und, dass ich das als weiße Person auch so darzustellen habe und, dass es gut ist wütend zu sein. Das Wir ist politisch indifferent und klebrig süß. Deswegen wird das Wir da stehen, subjektiv unbedingt angreifbar und transformierbar, auf dass feministische KI auch K.rüppel I.ntelligenz sein darf.



# VON RITTERN UND DRACHEN (KI ALS HEILSPERSÖNLICHKEITEN)

KI braucht Krüppel. Sie sucht nach Be:hinderung in uns. Sie heilt uns (Bennett & Keye 2019). Sie lässt uns normal werden, wenn wir weiß, reich und dankbar sind, im Hinterkopf das liberale Versprechen eines fairen Umgangs mit Menschen mit Behinderung. Das KI uns endlich mit allen anderen weißen reichen Menschen an die Startlinie stellt und wir im Rennen um den Abschluss, die Betriebsrente antreten können. Wir sind zwar die Underdogs, aber wer liebt sie nicht in unserer Leistungsgesellschaft. Und eventuell, wenn wir wirklich lieb fragen, sorgt der intelligente Sexroboter für einen digitalen Orgasmus, der uns zwar nicht zur heterosexuellen Erfüllung führt, uns aber immerhin ein Post-Sexgrinsen ins kartografierte Gesicht zaubert. Dabei ist die Frage der Beziehung, die wir Menschen mit Be:hinderung zu KI haben, haben könnten, eine multiple.

Künstliche Intelligenz taucht in Hinblick auf Behinderung oft als medizinische Erweiterung auf. KI wird zur digitalen Halsschlagader der westlichen Welt (vgl. ebd.). Sie wird oftmals im militärischen und infrastrukturellen Bereichen verortet und analysiert und leicht gerät aus dem wohlmeinenden Blick, dass ein großer Teil der Forschung sich mit der Grenze von gesund/krank und Be:hinderung/Nichtbe:hinderung beschäftigt. Care-Roboter, die Demenzkranken helfen, Apps zur Diagnostik von Autismus oder gleich zur störungsfreien Eingliederung von Autist\*innen sind wichtige Komponenten im Geflecht Künstlicher Intelligenz (Yergau 2014). Die Be:hinderung wird hierbei zum stabilisierenden Faktor der fragilen Technologie. Sie wird zur Kategorie, die binär und unveränderbar KI geduldig üben lässt, was genau der Kern der jeweiligen Behinderung ist.

Die datengetriebenen Ansätze Künstlicher Intelligenz sind Mustererkennungsmaschinen (vgl. Sejnowski 2018). Die Muster brauchen stabile Koordinaten, um lesbar zu sein für die Mensch-Maschine KI. Be:hinderung ist hierbei eine erfolgsversprechende Achse, abgrenzbar stabil und doch nicht eindimensional, sondern ausgedehnt; ein Spektrum oder nach Anforderung eine Achse an der Werte ab- und eingetragen werden können. Eine Koordinate, die Sinn verspricht und somit auch Erfolg. Aber, und das ist der tragisch komische Punkt, der/die Be:hinderte ist die kategoriale Sicherheit, die



versucht wird zu überwinden. Wir müssen existieren, um hinter der Technik zu verschwinden. Die Diagnostik der multiplen Be:hinderungen soll ja nicht dazu führen, dass auf einmal Krüppel statistisch relevant und gesellschaftlich existent sind. Es geht vielmehr darum, zu helfen uns einzugliedern in eine wohl situierte, niemals müde oder gebrechliche Gesellschaft (vgl. Yergau 2014). Vielleicht ist es mein Krüppel-Zorn der spricht. Ach, was schreibe ich, es ist ganz sicher der Krüppel-Zorn der hier spricht: Wer braucht die App, die Autist\*innen scheinbar normal kommunizieren lässt? Warum gibt es keine App, die Menschen hilft, sich neurodivers zu artikulieren, die ihnen hilft, Sprache legasthen zu dekonstruieren? Warum dürfen wir als die Abweichung und gleichzeitig als das statistische Charakteristika dienen, aber sind niemals der Erfolg, der ruhmreiche Diskussionsteil des Papers?

# PARALYSIERTE BUGS (KI ALS BE:HINDERT)

KI ist nicht nur die versprochene Heilung von Be:hinderung, sondern muss auch die Last des Versprechens einer funktionierenden Maschine erfüllen. Und sie funktioniert rational egal ob als antik anmutendes Expertensystem, welches auf exakten mathematischen Regeln ableisiert oder aber in der neuen schillernden Welt der neuronalen Netzwerke und des Deep Learnings. KI funktioniert und KI ist rational, aber Rationalität ist für uns längst kein Spiel mehr, sondern hohes ethisches Gut: „Ein rationaler Agent ist ein Agent, der das Richtige tut“ (Russel & Norvig 2012: 63).

Das Richtige ist kein aristotelischer Traum vom männlich freien guten Leben, sondern immer eine Kosten-Nutzen-Analyse und in unserer humanen und posthumanen Zeit wird alles einer universalen Kosten-Nutzen-Analyse unterzogen. Kunst, Musik und Sex sind die rationalen Entscheidungen der Menschheit auf dem Markt der Möglichkeiten, preisgünstig und für alle zu kaufen, die genug Kontodeckung haben. Es scheint wenig zu verwundern, dass die Entstehung der KI und die Erfindung des Neoliberalismus gemeinsam Geburtstag feiern (vgl. Shanahan 2015: 83).

Umso bemerkenswerter erscheint es da, dass der universelle Test zur Bestätigung zukünftiger Intelligenzen von einem Mann stammt, der weder Chicago Boy noch Freiburger Gelehrter war. Alan Turing hat den Turing-Test in den 1950ern designt und somit das bis heute gültige Kriterium verfasst, an welchem sich Entwicklungen auf dem Gebiet der KI messen lassen (vgl. Hofstadter 2016: 635). Turings Test ist zum semi-otischen Aktanten par excellence geworden. Einer diskursiven Praxis obliegt es, die menschlichen Tester\*innen als intel-

liger zu verstehen und somit als Teil des Fortschritts und der Revolution oder eben auch als fehlerhaft. Der Test scheint bestechend einfach: Zwei Personen und ein/e Fragesteller\*in begeben sich in eine Konversation, innerhalb derer der/die Fragesteller\*in festzustellen hat, ob ihr Gegenüber menschlich ist oder nicht. Zentral dabei ist, dass intelligent alle Personen sind, die als Mensch klassifiziert wurden (vgl. Haug 1995: 17; Haraway 1995).

Hier also fängt mein bescheidener Versuch an, Behinderung in KI zu verorten. Ich möchte Turings Test als semiotisch materielle Praxis zur Schaffung einer ableistischen KI decodieren. Die Leseweisen des Turing-Test verändern sich mit der neoliberalen Idee des Menschen als rational, immer verortet zwischen kosten- und nutzenabwägenden Akteur\*innen. Als objektiver Parameter des Turing-Tests gilt die Spracherkennung der Künstlichen Intelligenz. Diese muss Humor, kulturelle Spezifika und vieles mehr erkennen, um dann auf das Gespräch eingehen zu können. Dies muss grammatikalisch korrekt und schnell passieren. Ich bin nicht intelligent, wenn ich zwei Stunden zum Beantworten der Frage brauche, ob der Himmel blau ist oder nicht. Der alte Traum vom gesunden Geist im gesunden Körper scheint in den Träumen rund um Künstliche Intelligenzen Einzug zu halten (Lenzen 2018: 24–27).

Dabei folgt der Test einem sehr einfachen Muster:

## AKTION → REAKTION

Im klassischen Turing-Test werden Sprache und Schrift das Medium der Intelligenz. Wir müssen zuerst schreiben und sprechen können und das in einem gewissen zeitlichen Rahmen. Wir müssen auch auf Fragen reagieren können, Sprache verstehen, Witze einordnen, um nur einige Bereiche der sprachlichen Intelligenz zu benennen. Gerade auf diesen Aspekten beruht die Auswahl, wonach Menschen überhaupt erst zu qualifizierten Parametern für KIs werden (können). Der Turing-Test ist auf eine reibungslose Kommunikation als intelligent eingestuft Menschen angelegt.

Stellen Sie sich vor, was passiert, wenn wir stottern, aufgrund von sozialer Phobie nicht sprechen können oder schriftlich stottern, so wie ich. Wir erfüllen die an uns adressierte Norm des reziprok befragten Menschen nicht. Daraus können wir Folgendes ziehen: Entweder werden die Phobiker\*innen dieser Welt in den Test miteinbezogen und Schweigen in sozialen Situationen kann als intelligent gelesen werden, oder aber das Stottern wird als Abweichung eines Standards verstanden, der fehlerhafte Syntax als unintelligent ansieht und somit nicht als Kriterium für KI nützt. Interessanterweise konnten Programme Menschen mit dem Verweis auf Sprachschwierigkeiten in Form von Fremdsprachigkeit überzeugen, als menschlich angesehen zu werden. Dieses Verhalten wird jedoch als Trick verstanden und als nicht intelligent gewertet, denn Simulation gilt als Frevel im Reich der reinen rationalen Welt der KI (vgl. Lenzen 2018: 27). Hier zeigt sich der semiotisch materielle Akteur\*innen-Charakter des Turing-Tests. Zum einen werden statistische Werte geschaffen, die KI zu Künstlichen Intelligenzen machen, zum anderen wird neu definiert, was menschlich ist. Die menschliche Intelligenz steht also ebenfalls zur Disposition.

Gehen wir einen Schritt weiter und Fragen nach der Verkörperung von KIs und dem Zusammenhang von Intelligenz und Körper. Aktuelle Debatten kreisen vor allem um die

Frage nach dem Zusammenhang von Wetware und Hardware. KI-Forschung im Forschungszweig des „Artificial Life“ behandelt auch die Funktion des Lebens als Computerprogramm. Hierbei handelt es sich um eine Art Rational-Choice-Theorie der vermeintlich natürlichen Evolution, simuliert am Computer, dem Habitat der künstlichen Schwestern, lebenden Toten und rassistischen Bots. Forscher\*innen nähern sich mit Hilfe eines theologischen Selbstverständnisses und einem „Trial and Error“-Verfahren einer rationalen, zweckorientierten KI an. Eine Vertreterin dieser Theorien bietet folgendes, sehr exemplarisches Bild: „EvoSphere bestehend aus Geburtsklinik, Kindergarten, Testarena und *Evolution bleibt Evolution ein Recyclinghof für die Verlierer des evolutionären Wettbewerbs*“ (Lenzen 2018: 97; eigene Hervorhebung).

EvoSphere ist eine Computersimulation, die die Aufzucht von Robotern nachstellt. Was die gesellschaftliche Evidenz vom Recyclinghof dabei sein soll, bleibt der schauerlichen Fantasie oder einem Rückblick in die Geschichte überlassen (vgl. ebd.). Hier finden wir einen ungenierten Sozialdarwinismus vor. Es werden Gesellschaften imaginiert, die sich mittels wettbewerbsorientierter Selektion entwickeln. Diese Utopie der Künstlichen Intelligenz kommt ohne „Verlierer“ aus: Behinderung wird hier als naturalistische Fehlfunktion verstanden, die wie Sarah Kember bereits herausgearbeitet hat, keinerlei Berechtigung innerhalb der neuen KI Welt hat (vgl. Kember 2003).

Rückblickend möchte ich festhalten, dass der Turing-Test als semiotisch materieller Aktant, Realitäten schafft, welche in einer neoliberalen KI-Forschung Leistungsprinzipien und Rational-Choice-Ideen als Maßstab für Intelligenz schaffen, welche sich sowohl auf KI als auch auf ihre humanoiden Nachbar\*innen auswirken. Ferner hält ein sozialdarwinistisches Evolutionsbild Einzug in die Forschung, welche Körpernormen auf, und das möchte ich in aller Klarheit festhalten, höchst gefährliche Art und Weise zu Kategorien des Lebens an sich macht.

2            Ableisiert leitet sich vom Konzept des Ableismus ab: „Das heißt, anders als der Begriff Behindertenfeindlichkeit setzt Ableism Behinderung nicht als gegebenes Faktum voraus, gegen das sich die Abwehr richtet, sondern nimmt die normative Einteilung in Behinderung und Nichtbehinderung in den Blick. Nichtbehinderung und Gesundheit werden zum Ideal erhoben, an dem sich alle Körper zu bewähren haben – indes ist dieses Ideal auch auf seine Gegenseite verwiesen.“ (Maskos 2015)

3            Wetware ist eine Neologismus und meint, im Gegensatz zu Hardware, also den Kabeln, Prozessoren und Ähnlichem, das menschliche Gehirn und Nervensystem. Dieser Begriff wurde geschaffen, um die Verbindung von Maschine und Mensch im Computerzeitalter zu konzeptualisieren (vgl. Riskin 2003: 97).

# SMOOTH CYBORGS (KI ALS MENSCH- MASCHINE- INTERAKTION)

Wenn wir über die multiple und zerstückelte Verbindung von Be:hinderung und KI nachdenken, müssen wir über den Cyborg reden. Jene Figur des kalten Kriegs, die ein Verschmelzen von Technik und Organismus markiert. Die/Der Cyborg ist eine Grenzverschiebung von Technologie und Organismus. Cyborgs können Menschen sein oder auch Tiere, wie beispielsweise Laborratten, die genmodifiziert werden (vgl. Haraway 1985). Wenn es sich um einen feministischen Blick auf jene neue Identität im posthumanen digitalen Kapitalismus dreht, ist Donna Haraway die illegitime akademische Mutter (ebd. 1985). Sie erzählte vor dem Hintergrund einer drohenden atomaren Apokalypse im Manifest für Cyborgs von einer neuen Figuration von Menschen, Tieren und Technologien, die weder in ein prädigitales Eden noch in eine Technik-utopie, oder -dystopie fliehen konnte. Auferstanden aus militärischen, kolonialen und rassistischen Verhältnissen war die/der Cyborg das Versprechen, diese Widersprüche auszuhalten, mit ihnen politisch zu spielen und sie zu nutzen, um die Welt, wie sie ist und war, zu verändern (ebd.). Be:hinderte waren in diesen feministischen Interventionen stets das lebende Beispiel und wenn wir heute auf die intelligenten Bruchstücke der Technik schauen, die uns ermöglicht, was der eigene Körper verweigert, sind Künstliche Intelligenzen und Crips-Cyborgs im Zeitalter neuronaler Netze ein polyamoröses Ehepaar.

Dabei wird diese Vernetzung zum artifiziellen Fakt. Behinderung wird nicht nur durch Technik geheilt, so das Versprechen, nein ihr seid auch im Jahr 2020 die ersten richtigen, die natürlichen Cyborgs (vgl. Kafer 2013: 105). Fernab gefährlich ironischer Allianzen werden Verbindungen geschaffen, die widerspruchsvoller nicht sein könnten und doch von einer immer ausdifferenzierteren weißen cis hetero und abled Gesellschaft gebraucht werden. Der/die Cyborg als be:hinderte Norm! Schau, mit Technik bekommen wir es hin. Sie sind nicht so be:hindert, wie wir gedacht haben. Mit ein wenig Microsoft und ein wenig Apple können die Be:hinderten gehen, sehen und sprechen! Dabei werden uns schmerz- und widerspruchsfreie Verbindungen versprochen. Die Technologie hilft der/dem be:hinderten Cyborg. Aber wie wir nach fast einem halben Jahrhundert neoliberaler Ordnungsfantasien gelernt und verinnerlicht haben: Kein Nutzen ohne Kosten. Doch was sind die Kosten? Wer bezahlt sie?

Feministische, postkoloniale und disability Wissenschaftler\*innen, wie Jina B. Kim mit ihrer Crip-of-Color Kritik, zeigen immer wieder auf, dass Beziehungen machtvoll sind, die Hierarchien reproduzieren und den Nutzen der einen auf Kosten der anderen evozieren (vgl. Kim 2017). Die Verbindun-

gen von Technologie und Be:hinderung schafft mir als weißer be:hinderter europäischer Person Vorteile, die nicht mir gehören. Intelligente Prothesen sind keine nachwachsenden Beine oder Gelenke. Es braucht Rehabilitationsmedizin, Diagnostik, Operationen. Es sind schmerzvolle Verbindungen, die Care-Arbeit erfordern, zumeist von marginalisierten Personen zu schlechtbezahlten Konditionen. Meist bleiben die Prothesen Krankenversicherten vorbehalten und das, weil wir in einer Welt leben, die sich im westlichen Kapitalismus entschlossen hat Be:hinderung zu ignorieren. Fehlende Rampen sind kein Zeichen von Naivität, sondern der machtvollen Taktik politisch motivierter Ignoranz!

Dabei wird KI als scheinbar mechanischer Teil des Cyborgs erkenntlich als sperrige widerspruchsvolle Figuration vergeschlechtlicher Care-Arbeit, rassifizierender Ausbeutungsprozesse und ableistischer Vorstellungen eines guten Lebens. Dabei werden oftmals die Konfiguration von Be:hinderung, globalen Verwertungsketten und dem wachsenden Anteil an Menschen mit Be:hinderung im globalen Süden verunsichtbart (vgl. Meekosha 2011: 669). Die Verbindung von Mensch und Maschine ist nicht evident, sie ist artifiziert und voller, im körperlichen Sinne, schmerzlicher Erfahrungen. Wir verbinden uns nicht einfach mit einer intelligenten Maschine: Wir verkabeln uns, wir schreien und haben finanzielle Sorgen (Quinlan & Bates 2014). Als weiße behinderte Cyborgs in Krankenversicherungssystemen ist das ein Privileg, da künstliche Gelenke und intelligente Prothesen von Versicherungen bezahlt werden. Wir dürfen also schmerzvolle Verbindungen eingehen. Wir sind die artifizierten Fakten. Wir sind eine privilegierte Cyborg-Gemeinschaft, die untersucht wird, die aber auch eine kritische Wissenschaftscommunity hinter sich hat, die Medizin und Be:hinderung kritisch hinterfragt (vgl. Mulderink 2020). Das ist der nicht unschuldige, auf den globalen Norden fokussierte Part der Cyborg-Geschichte. Meine eigene Prothese habe ich erhalten, weil ich jung bin, weiß und krankenversichert. Menschen nehmen mir meinen Schmerz ab. Sie sorgen für mich und meine Krankenkasse stuft mich als Wissenschaftler\*in wertvoll genug ein, sodass sie mir Zeit einräumt mit der Prothese zu üben. Dies auch, damit ich dem Arbeitsmarkt nicht verloren gehe. Die Nutzen sind nicht meine, aber ich profitiere von ihnen. Cyborgs, die mit KI leben sind nicht soft und konfliktfrei mit ihnen verschmolzen. Es sind politische Grenzverschiebungen die schmerzen. Es sind Grenzverschiebungen, die einen Großteil der Weltbevölkerung aufgrund kolonialer Ausbeutungsverhältnisse und Infrastruktur verwehrt werden.



# LEGASTHENE KI UND WEISSE PRIVILEGIEN

In den USA wurde Software seit Mitte der 70iger zu einer Ware. Durch gerichtliche Urteile werden Algorithmen nicht mehr als Naturgesetze verstanden oder als Rechenoperationen, wie  $1 + 1$ , sondern als urheber\*innengeschützte Güter der Wirtschaft konstruiert. Software wird zu einer Patentangelegenheit und damit zu einem Faktor möglicher wirtschaftlicher Vormacht.

Gleichzeitig professionalisiert sich die Computerwissenschaft. Universitätslehrstühle entstehen und der Beruf des Programmierers, der Programmiererin wird hochbezahlt. Frauen und vorrangig Schwarze Frauen werden aus den Berufen rund um Softwareentwicklung gedrängt und müssen einer weißen männlichen Belegschaft weichen, die bis heute in den KI-Bastionen des Globalen Nordens codieren (vgl. Chun 2013). Aber wie gute Feminist\*innen wissen, fallen Universitätsabsolvent\*innen nicht vom Himmel, sondern gehen in den Kindergarten und die Schule. Und diese Schule muss Anfang der 50iger mehr und mehr weiße Männer an die Universitäten schicken, um eine neu aufkommende Ökonomie die auf Expertise setzt anzukurbeln.

Die Schule wird ganz ohne das Zutun französischer Intellektueller zu einem Ort der Disziplinierung und der Gewinnung neuer zukünftiger Technikmeister. Dafür werden höhere Standards in Form von Test- und Lehrinhalten eingeführt. Gleichzeitig taucht eine neue Kategorie der Be:hinderung auf: Die Lernbe:hinderung. Ein Spektrum, welches sich auf abgeschlossene Bereiche wie das mangelhafte Lesevermögen von Schüler\*innen im Vergleich zu ihren Klassenkamerad\*innen ausdrückt. Die Lernbe:hinderung ist eine Be:hinderung, die sich auf einen einzelnen Bereich auswirkt und biologische Ursprünge zu haben scheint (vgl. Sleeter 1987: 226). Waren Lernschwierigkeiten in einer postdigitalen Welt eine vernachlässigbare Größe, werden sie vor dem Hintergrund atomaren Wettrüstens und Wirtschaftseinbrüchen und dem zukünftigen Softwaremarkt ein soziales Problem, das messbar, therapierbar und vor allem normalisierbar wird. Sleeter fasst die sozial vielschichtige Bedingtheit dieser Trennungslinie wie folgt zusammen:

„Learning disabilities seemed to explain white middle class children particularly well because it did not level blame on their home or neighborhood environment, it upheld their intellectual normalcy, and it suggested hope for a cure and for their eventual ability to attain relatively higher status occupations than other low achievers.“  
(Sleeter 1987: 230)

Damit wird Be:hinderung im Kontext einer sich immer weiterentwickelnden und exklusiveren Wissensökonomie, die Fachkräfte benötigt, zu einer weiteren Segregationslinie. Diese schreibt sich bis heute fort. So legt 20 Jahre nach dem Aufsatz von Sleeter, Wanda Blanchett 2010 dar, dass sich bis in die Gegenwart die Kategorie der Lernbe:hinderung in den USA zwar hinsichtlich der Kategorien race und Klasse geändert hat, die Funktion aber erhalten geblieben ist. So werden jetzt vermehrt Schwarze und Kinder of Colour in die Kategorie

der Lernbe:hinderung hinzugezählt. Diese werden jedoch weit weniger in Normklassen geschickt, sondern in Sonderklassen mit wenig Aussicht auf die Möglichkeit eines Universitätsabschlusses und somit auch weniger Möglichkeit an der Entwicklung von KI beteiligt zu sein (vgl. Blanchett 2010: 7).

Für den deutschen Kontext fehlt eine solche Studie, dennoch lässt sich beispielsweise dem Verband für Legasthenie und Dyskalkulie zu Folge eine Bildsprache erkennen, die nur weiße Personen abbildet und auch die Ursachen für Legasthenie und Dyskalkulie sowie deren Begrenztheit lassen sich nicht aus der Seite entnehmen (siehe Webseite: Verband für Legasthenie und Dyskalkulie). Be:hinderung ist eine komplexe Kategorie, die sich nicht in eine unschuldige Zukunft auflösen lässt. Lernbe:hinderung ist Teil eines Systems weißer Privilegien, die Wissenschaftler\*innen generieren kann und das vorrangig für naturwissenschaftliche Fächer und Informatik. Wenn wir also von einer überwiegend weißen Wissenschaftscommunity innerhalb der westlich geprägten KI-Community reden, müssen wir auch über Be:hinderung als konstitutiven Faktor mit all seiner machtvollen Widersprüchlichkeit reden. Als lernbe:hinderte Person, die jetzt in diesem Sammelband zu KI veröffentlichen darf, erkenne ich an, dass all die rechtlichen Kämpfe, die Diskriminierungserfahrungen, aber auch die Therapieangebote und die Lobbyarbeit auf einem rassistischen Erbe beruhen, welches bis heute fortbesteht. Es ist ein Privileg, weiß und lernbe:hindert zu sein. Dieses Privileg schützt vor einer Zukunft in schlecht bezahlten Jobs oder sichert einen Platz im Heim.

# K. RÜPPEL I. INTELLIGENZ

Nach diesem lückenhaften Ausflug in eine alternative KI-Deutung braucht es noch einen parteiischen Blick in die Zukunft. Was wäre, wenn KI selbst ein Krüppel wäre? Sich einer eindeutigen Funktion nicht nur verweigert, sondern etwas Neues eröffnet, einen Blick auf maschinelle Intelligenz, die nicht in Kosten-Nutzen-Kalkulationen, sondern in Stottern und Krämpfen aufgeht? Der Einsatz von KI im Bereich der Medizin, der Kommunikation und Mobilität und ähnlichen ermöglicht einer Vielzahl von Menschen an einem Leben teilzunehmen, dass für sie nie gedacht war. Gleichzeitig wird Be:hinderung immer mehr zur scheinbar neoliberal vermeidbaren Wahl. Das Ideal von Gesundheit gibt es nicht umsonst und schon jetzt arbeiten vor allem im Globalen Süden Menschen und besonders FLTI\* Personen für diese Technowelt, des heilen weißen Mannes und für ein, von großen Technologiekonzernen imaginiertes Ideal von Intelligenz/ Begehren/ Konsum. Ich denke, wenn wir Be:hinderung in diesen Prozess einfließen lassen, als Bug, können wir KI zu K.rüppel I.intelligenz werden lassen. Es wird unser

5400  
5401  
5402  
5403

Verständnis von der Welt ändern. Es ist abgeleitet aus der strukturellen Unterdrückung marginalisierter Personen. Es wird uns auf neue Diskriminierung stoßen lassen, aber ich bin der tiefsten Überzeugung, dass es unsere einzige Wahl ist.

Katharina Klapheck ist Wissenschaftler\*in und Aktivist\*in. Arbeitsschwerpunkte sind die intersektoralen Verschränkungen von Geschlecht und Behinderung innerhalb der Entwicklung Künstlicher Intelligenzen. Ein Hauptaugenmerk liegt hier vor allem auf den Designpraktiken zur Wissensübermittlung durch Informatik.

# LITERATUR

- Blanchett, Wanda J. (2010)  
„Telling It Like It Is: The Role of Race, Class, & Culture in the Perpetuation of Learning Disability as a Privileged Category for the White Middle Class“. *DSQ*, 30 (2), online unter: <http://dx.doi.org/10.18061/dsq.v30i2.1233>.
- Bennett, Cynthia L. & Keyes, Os (2019)  
„What is the Point of Fairness? Disability, AI and The Complexity of Justice.“ In: *ArXiv* abs/1908.01024.
- Chun, Wendy Hui Kyong (2013)  
*Programmed visions. Software and memory*. Cambridge / Massachusetts / London: MIT Press.
- Haraway, Donna Jeanne (Hg.) (1995)  
*Monströse Versprechen. Die Gender- und Technologie-Essays*. Hamburg: Argument Verlag.
- Haraway, Donna Jeanne (1985)  
„A Cyborg Manifesto. Science, Technology, and Socialist Feminism in the 1980's“. *Socialist Review* (80), 65–107.
- Haug, Frigga (1995)  
„Riskante Verbindungen. Donna Haraways Dynamisierung der Standpunkte“. In: Haraway, Donna Jeanne (Hg.): *Monströse Versprechen. Die Gender- und Technologie-Essays*. Hamburg: Argument Verlag, 11–80.
- Hofstadter, Douglas R. (2016)  
*Gödel, Escher, Bach. Ein endloses geflochtenes Band*. Stuttgart: Klett-Cotta.
- Kafer, Alison (2013)  
*Feminist, Queer, Crip*. Bloomington/Indianapolis: Indiana University Press.
- Kember, Sarah (2003)  
*Cyberfeminism and Artificial Life*. London / New York: Routledge.
- Kim, Jina B. (2017)  
„Toward a Crip-of-Color Critique: Thinking with Minich's 'Enabling Whom?'“. *Lateral* 6 (1). Online unter: doi: 10.25158/L6.1.14.
- Lenzen, Manuela (2018)  
*Künstliche Intelligenz. Was sie kann & was uns erwartet*. München: C.H. Beck.
- Maskos, Rebecca (2015)  
„Ableism und das Ideal des autonomen Fähig-Seins in der kapitalistischen Gesellschaft.“ *Zeitschrift für Inklusion* (2). Online verfügbar unter <https://www.inklusion-online.net/index.php/inklusion-online/article/view/277>.

- Meekosha, Helen (2011)  
„Decolonising disability: thinking and acting globally“. In: *Disability & Society* 26 (6), 667–682.
- Mulderink, Carrie Elizabeth (2020)  
„The emergence, importance of #DisabilityTooWhite hashtag“. *DSQ* 40 (2). Online unter: doi: 10.18061/dsq.v40i2.6484.
- Quinlan, Margaret M. & Bates, Benjamin R. (2014)  
„Unsmoothing the Cyborg: Technology and the Body in Integrated Dance“. *DSQ* 34 (4). Online unter: doi: 10.18061/dsq.v34i4.3783.
- Riskin, Jessica (2003)  
„Eighteenth-Century Wetware.“ *Representations* 83 (1), 97–125.  
DOI: 10.1525/rep.2003.83.1.97.
- Russell, Stuart J. & Norvig, Peter (2012)  
*Künstliche Intelligenz. Ein moderner Ansatz.*  
München: Pearson.
- Sejnowski, Terrence J. (2018)  
*The deep learning revolution.* Cambridge/  
Massachusetts/London: The MIT Press.
- Shanahan, Murray (2015)  
*The technological singularity.* Cambridge/  
Massachusetts/London: The MIT Press.
- Sleeter, Christine ([1987] 2010)  
„Why Is There Learning Disabilities? A Critical Analysis of the Birth of the Field in Its Social Context“. *DSQ* 30 (2).  
Online unter: doi: 10.18061/dsq.v30i2.1264.
- Yergeau, Melanie (2014)  
„Disability Hacktivism. Computers and Composition.“ *Hacking the Classroom.*  
Online verfügbar unter  
<http://www2.bgsu.edu/departments/english/cconline/hacking/#yergeau>.



FEMINISTISCHE  
ZUKÜNFT  
SCHREIBEN.



KATRIN  
FRITSCH  
&  
HELENE  
VON  
SCHWIGHOW



Frühjahr, 2020: unter dem Hashtag #FlattenTheCurve werden Menschen weltweit dazu aufgefordert, zu Hause zu bleiben, um die COVID-19 Ansteckungsrate auch in Zukunft gering zu halten. Die Nachrichten sind mit Zahlen und Diagrammen gefüllt: prognostizierte Mengen an Krankenhausbetten, voraussichtliche wirtschaftliche Einbrüche, steigende Arbeitslosenraten. Im Gegensatz zu früheren Katastrophen, wie etwa Cholera-Ausbrüchen, wird das Ausmaß von COVID-19 bereits vor dem endgültigen Eintreten berechnet. Algorithmen, Codes, Statistik – Vorhersagen dienen als Grundlage um politische Entscheidungen zu treffen.

Die COVID-19 Pandemie macht deutlich, was sich schon seit mehreren Jahren abzeichnet: wir können die Gegenwart nicht mehr ohne ihre Zukünfte verstehen<sup>1</sup>. Wer über Zukunft spricht, beeinflusst gegenwärtige Entscheidungen. Wer Zukunft produziert, sie also vertextet oder zeichnet, hat die Deutungshoheit. Über Zukunft zu sprechen ist politisch.

Unternehmen im Silicon Valley haben schon lange verstanden, wie wichtig es ist, Zukunft zu produzieren. Mark Zuckerberg, Sergey Brin oder Elon Musk sprechen ständig darüber, wie die Technologien der Zukunft aussehen werden, was sie können – und was eben auch nicht. Selbstfahrende Autos, smarte Brillengläser oder Internet im Weltraum sind nur einige dieser Zukunftsvisionen. Aber ist das wirklich die Zukunft, in der wir alle leben wollen?

Wirft man einen genaueren Blick auf die öffentlich propagierte Zukunft wird sehr schnell klar, dass sie hauptsächlich von weißen, privilegierten Männern erträumt wird. Damit ist sie homogen, scheint unabdingbar und schließt Wünsche und Bedürfnisse vieler anderer Menschen aus. Brauchen wir in Zukunft wirklich selbstfahrende Autos, und nicht viel eher eine CO2-neutrale Verkehrsinfrastruktur? Brauchen wir wirklich smarte Brillengläser, und nicht viel eher ein robustes demokratisches System?

Zukunft zu produzieren eröffnet Möglichkeitsräume. Wenn nur einige wenige Zukunftsvisionen geschaffen werden, werden Alternativen vergessen. Die Frage sollte also nicht lauten *Wie wird die Zukunft aussehen?, sondern, Wie kann Zukunft diverser gestaltet werden? Wie können wir Zukunft von Grund auf feministisch denken und Werte wie Gemeinschaft oder soziale Gerechtigkeit inkludieren? Wie können Machtstrukturen erkannt und aktiv gestürzt werden? Wie können wir Zukunft demokratisieren?*

Feminismus hat langjährige Erfahrung damit, sich genau diese Fragen zu stellen und kreative Antworten auf sie zu finden. Intersektionalität erlaubt zu erkennen, dass es viel mehr als nur eine Zukunft gibt. Anti-patriarchale Perspektiven erlauben es, nicht-binäre Narrative zu schreiben. Antikapitalistische Ansätze ermöglichen es, über Zukünfte nachzudenken, in denen Nachhaltigkeit und soziale Gerechtigkeit, nicht Profit und Disruption, als Ziele definiert werden.

Indem man sich Zukunft vorstellt und diese Vorstellung in Worte fasst, wird sie nicht nur greifbar, sondern zeigt auch die Schritte auf, die in der Gegenwart notwendig sind, um eben diese Zukunft zu erreichen. Denn über Zukunft zu sprechen, heißt auch über uns selbst zu sprechen und Ziele für uns als Gesellschaft zu formulieren. Somit hat Imagination ein enormes politisches Potenzial zu gesellschaftlichen Veränderungen – besonders wenn sie gemeinschaftlich passiert.

Es ist also höchste Zeit, sich den Raum und die Zeit zu nehmen<sup>2</sup> diese Zukünfte aktiv zu imaginieren, sie niederzuschreiben und sie zu erzählen. Dabei ist es wichtig, sich nicht der gleichen technik-deterministischen Werkzeuge, wie

5467 sie aus dem Silicon Valley kommen, zu bedienen. Das heißt:  
5468 Künstliche Intelligenz wird nicht passieren, viel eher haben wir  
5469 die Freiheit zu entscheiden, ob und wie sie passieren wird. Bio-  
5470 metrische Erkennungssysteme sind nicht die Zukunft, son-  
5471 dern können aktiv abgelehnt werden.

5472 Wir sollten also frei und kreativ unsere Zukünfte  
5473 erträumen, niederschreiben, und damit neue feministische  
5474 Möglichkeitsräume definieren. Nicht mit den Werkzeugen des  
5475 Patriarchats, wie etwa Vorhersage-Algorithmen oder binären  
5476 Zahlen, sondern viel eher mit unserer Vorstellungskraft.  
5477 Gemeinsam, fair, und in Solidarität. Oder wie Audre Lorde  
5478 sagt: „[...] if we do not define ourselves for ourselves, we will  
5479 be defined by others – for their use and our detriment.“<sup>3</sup>

1 Dernbach, Rafael (2020): *10 Thesen zur Produktion von Zukünftigkeit*. Abgerufen von <https://vimeo.com/411006277>.

2 Oder, wie Virginia Woolf in *Ein Zimmer für sich allein* erkennt, ein eigenes Zimmer zum Schreiben finden.

3 Lorde, Audre (2017): *Your silence will not protect you*. London: Silver Press.

MOTIF ist ein unabhängiger Think Tank, der an der Schnittstelle von Technologie und Gesellschaft arbeitet, gegründet von Katrin Fritsch und Helene von Schwichow. Sie beraten Organisationen und politische Entscheidungsträger\*innen zu nachhaltigen Technologien und übersetzen komplexe Fragen des digitalen Zeitalters in umsetzbare Formate.



USE YOUR  
TECHNICAL  
MONSTERS

---

CREATE THEM,  
NURTURE THEM  
AND  
USE THEM  
FOR  
CREATING  
NURTURE  
AND CARE.



ANNIKA  
KÜHN



Technologien entstehen nicht aus dem Nichts. Technologien werden geschaffen. Erschaffer\*innen/Macher\*innen von Technologien reflektieren aber zu selten ihren Einfluss auf die entsprechende Technologie und den Einfluss dieser Technologie auf sie selbst. Bruno Latour nutzt in seinem Essay „Love your Monsters“ eine Analogie zu *Frankenstein*. Die Geschichte sei in der populären Vorstellung eine warnende Legende gegen Technik. Auf diese Weise wird das Monster mit dessen Macher verwechselt.

„It is not the case that we have failed to care for Creation, but that we have failed to care for our technological creations. We confuse the monster for its creator and blame our sins against Nature upon our creations. But our sin is not that we created technologies but that we failed to love and care for them. It is as if we decided that we were unable to follow through with the education of our children.“  
(Latour 2012)

Es bringt nichts, die bereits entwickelten technischen Produkte als Monster zu verteufeln, sondern die Taten, die dazu geführt haben, dass es Monster wurden/werden können, sollten geortet und bewusst geändert werden, um sie im nächsten Schritt zu vermeiden. Im Kontext dieser Publikation geht es dabei zum Beispiel um einen gender bias bei KI, um autonome Waffensysteme oder die Benachteiligung von Menschen in ihrem Recht auf Zugang zu und Teilhabe an etwas, das mithilfe von technischen Prozessen gesteuert wird, z.B. bei der Vergabe von Jobs oder Krediten. Außerdem wäre eine nachhaltige Gewährleistung von Datenschutz bei persönlichen Daten erstrebenswert, die ebenso mit einem autonomen, mündigen Umgang der Nutzer\*in einhergeht. Dafür benötigt es wechselseitig eine kritische Gesellschaft, die kritische Macher\*innen von Technologien „erzieht“, indem sich in Diskurse und Entscheidungen eingemischt wird und Räume feministisch intersektional mitgestaltet werden.



# FEMINISTISCHE SCIENCE FICTION

Der Status Quo sollte sein, dass Technologien wie KI von Macher\*innen konstruiert werden, die sich ihrem Einfluss durch ihr Mindset bewusst sind und bereit dafür sind, ihr Mindset immer wieder kritisch in Hinblick auf intersektionale, feministische Perspektiven zu hinterfragen – von hier ausgehend: KI wird in Hinblick auf mehrere Aspekte des Sozialen Miteinanders entwickelt nach dem Motto „Nutze Technik als Mittel zum sozialen Zweck, gestalte mit ihr“. Feministische Science Fiction kann hierzu einen Beitrag leisten, indem Leser\*innen zur Selbstreflexion angeregt und für ein Vorstellungsvermögen über mögliche Zukunftsszenarien beigetragen wird – utopisch sowie dystopisch. Aus diesem Gegensatz kann dementsprechend eine Synthese, in Form von Transformation – auch rein gedanklich – basierend auf einer Problemanalyse, entstehen.

Donna Haraway kreierte den Cyborg, eine Fusion aus Tier und Maschine, welche die Gegensätze zwischen Natur und Kultur irritiert: „This cyborg is our ontology; it gives us our politics. The cyborg is a condensed image of both imagination and material reality, the two joined centres structuring any possibility of historical transformation“ (Haraway 1991: 1). Science Fiction kann Menschen dazu befähigen, ihre Perspektive auf die Realität zu hinterfragen und zu Perspektiven anderer Möglichkeiten zu gelangen. „Feminist science fiction is a textual – and embodied – practice rather than a textual form: it is a practice that is developed in multiple interpretive communities that take „pleasure in the confusion of boundaries and ... *responsibility* in their construction“ (Haraway 1991 [1985]: 150).

# UTOPIE – EINE KI, DIE NACHHALTIGE FEMINISTISCHE EMOTIONALE CARE-WORK IN BEZUG AUF PSYCHISCHE GESUNDHEIT IN FORM VON MEDIATION LEISTET

In dieser Utopie geht es um Fürsorge und Förderung zwischen und unter Menschen durch KI. Sie hilft, Menschen dazu zu bringen, sich mit sich selbst und den Mitmenschen auseinanderzusetzen. Daher wird sie als hybrides Objekt, bestehend aus Technik und Sozialem verstanden, welches also auch als soziale Akteurin gesehen wird, die das soziale Miteinander nachhaltig gestaltet. Sozusagen eine „emotional caring KI“, die sich erinnert, dadurch Menschen mit sich selbst und ihrem Verhalten konfrontiert, ihnen das Gefühl gibt ernst genommen zu werden und dabei hilft, nachhaltige Netzwerke, Verbindungen, Interaktionen zwischen Menschen zu schaffen. Diese existieren ohne Bevormundung, ohne Oktroyierung, ohne Macht.

In der Utopie sind Soziale Medien nachhaltige Resonanzräume, die so gestaltet sind, dass sie die Individuen, die sie nutzen, sensibilisiert auf sich und das Gegenüber zu achten – in Bezug auf Wertschätzung, Stress und psychische Gesundheit. Die KI kann z.B. dafür eingesetzt werden Plattformen, die bei der Teamarbeit genutzt werden, so zu gestalten, dass die Individuen sich gegenseitig wertschätzend begegnen können, ohne in Arbeitsstress zu verfallen, der die Teambeziehung und -arbeit belastet. Bei gleichzeitiger Datensicherheit, vor allem auch bzgl. personenbezogener Daten.

Die KI funktioniert nicht, indem sie Verhaltensmuster auf Basis von Sprache oder Bildern analysiert, generalisiert und massenweise gleich anwendet, wie nach dem Prinzip des

deep learning. Sie hat ein Bewusstsein für Individualität und ist kognitiv flexibel.

(Vermeintliche) Identitäten zu konstruieren und als Datengrundlage zu speichern spielen für sie keine Rolle. Sie ist multiperspektivisch und hat doch eine klare Haltung – pro Selbstbestimmung.

Dadurch kann sie sich in Konflikt- und Stresssituationen einmischen – quasi Mediation – und begleitet diese, indem die Individuen zum Innehalten und Nachdenken angeregt werden, bevor sie z.B. emotionalisiert diskutieren oder sich zurückziehen um den Konflikt belastend mit sich selbst auszumachen.

Eine weitere Eigenschaft ist, Personen davor zu schützen, sich selbst nicht von Arbeit und Informationen, die durch technische Vernetzung an sie herangetragen werden, abgrenzen zu können, und somit einer „Fear of missing out“ vorzubeugen.

Sie selbst ist ein technisches Objekt, welches eine Transformation der Wahrnehmung, Haltung und Nutzung gegenüber und von Technik bewirkt.

Die Transformation ihres Umfeldes in Bezug auf Technik ist ihr Selbstzweck – bei gleichzeitiger Transformation ihres Umfeldes im Miteinander. Sie ermöglicht Raum für Austausch auf Augenhöhe, einem, in dem eine Akzeptanz von Abhängigkeit voneinander existieren darf. Intention ist es, Handlungsfähigkeit zu bewahren. Mit einem Bewusstsein von Präsenz statt Produktivität.



## MÄRZ 2042

Ich wache auf, trödel so vor mich hin. Die Anschaffung der Kaffeemaschine, die mitbekommt, wann ich aufstehe, war super! Ich muss heute einen Beitrag für unser Newstopal „Minority Report“ über Menschenrechte schreiben. Will nicht anfangen – Stress. Gestern erhielt ich einen Anruf aus dem Institut für Demographie-Forschung. Es ist ein historischer Moment – die Zahlen bestätigten die Prognose, dass in dem Jahr 2042 ehemals ethnische Minderheiten in der Mehrheit sind. Die *weiße*, privilegierte Gesellschaft rückt in den Hintergrund. Es gibt eine Kooperation zwischen staatlichen Institutionen und der dezentral organisierten „ProHuRi“ – Programmmer\*s for Human Rights“. Sie besteht zu 70% aus LGBTIQ\*s und BPoCs. Die Hälfte sind Frauen. Die meisten von ihnen sind damit beschäftigt Software und andere Programme für den Staat zu entwickeln, für Sicherheit: Sicherheit der Zivilgesell-

5610 schaft. Sicherheit von persönlichen Daten. Schutz vor Gewalt.  
 5611 Schutz vor Ausbeutung der Privatsphäre. Schutz vor der Aus-  
 5612 beutung von Rohstoffen und von Menschen durch neoliberale  
 5613 Interessen aus dem Silicon Valley. Schutz vor autonomen Waf-  
 5614 fensystemen. Schutz vor machtgeilen Trumps, die einen Sieg  
 5615 auf anderen Menschen – über andere Menschen – leben.  
 5616 „Wie hast du geschlafen? Du bist gestern erst so spät ins  
 5617 Bett gegangen.“  
 5618 Puh, danke für den Reminder an die scheiß Nacht. Kaffee  
 5619 schmeckt nicht nur gut.

## 10:42

5620 Das Mailfach quillt über. Mein Handy ist keine Se-  
 5621 kunde still. Infos über News über Anfragen nach Infos und  
 5622 News von mir. Ich brauche Struktur. Kalender. Um 12:00 ist  
 5623 ein Zoom-Team-Meeting mit Inde und Malo. Danach ein  
 5624 Lunch-Date mit einer Journalistin. Bin gespannt auf ihre Per-  
 5625 spektive. Um 14:30 muss ich zur Menschenrechtsorganisation  
 5626 IP-Rights, um ein Konzept für deren Jahresfeier mitzugestalten.  
 5627 Heute Abend ist der Launch der Kampagne „Black Lives did  
 5628 always matter!“. Mit Arna wollte ich mich noch treffen, haben  
 5629 uns ewig nicht geupdated. Beziehungsarbeit. Ahh – Und wann  
 5630 schreibe ich den Beitrag? Kann mich nicht konzentrieren.  
 5631 „Hast du schon gefrühstückt? Im Kühlschrank ist noch Milch  
 5632 und du hast dir letzte Woche dein Lieblings-Müsli gekauft.  
 5633 Ich mache mir Sorgen um deine Gesundheit.“  
 5634 Erstmal frühstücken. I need energy!

## 12:42

5635 Wir müssen auf den Punkt kommen. Wir brauchen  
 5636 Output. Der Beitrag muss bis Mitternacht stehen. Warum  
 5637 haben wir uns das aufgehalst? Das schaffen wir niemals. Doch,  
 5638 ich mache das heute Abend noch. Oh nee, du machst doch  
 5639 schon so viel. Der Beitrag verlangt volle Konzentration. Der ist  
 5640 wichtig für unsere Publicity! Ja, wer macht es sonst? Wir sind  
 5641 nur zu dritt. Puhh, ich kriege den auf keinen Fall noch unter.  
 5642 „Was ist mit eurem Vorhaben, mehr auf euch und das Mitein-  
 5643 ander zu achten? Ihr habt heute noch nicht darüber gespro-  
 5644 chen, wie es euch geht.“  
 5645 Boah, mich nervt diese KI. Wir haben Stress, okay??  
 5646 „Okay. Das verstehe ich. Es tut mir leid. Stress ist ungesund.“  
 5647 Haben wir uns den selbst ausgesucht? Je höher unser Be-  
 5648 kanntheitsgrad, der Ziel unserer Arbeit ist, desto mehr Arbeit,  
 5649 desto mehr wird an uns herangetragen – on- wie offline. Je  
 5650 mehr Expertise wir haben, desto mehr fällt uns auf, was wir  
 5651 alles noch nicht wissen oder bedacht haben! Es geht immer  
 5652 weiter, immer mehr. Komplexität!

5653 „Ja, es ist sehr komplex und das kann überfordern. Wie willst  
5654 du damit umgehen? Passt auf euch auf.“  
5655 Wie will ich damit umgehen? Erstmal darüber nachdenken...  
5656 wenn das nächste Mal Zeit dafür ist.

## 23:42

5657 „Wann willst du ins Bett gehen? Gestern warst du um 1 Uhr  
5658 im Bett. Bei wenig Schlaf kann das Immunsystem  
5659 geschwächt werden. Bald ist wieder Grippe-Zeit.“  
5660 Schon fast Mitternacht. Der Beitrag steht zur Hälfte. Der  
5661 muss heute noch fertig werden. Heute schon wieder nichts  
5662 für mich gemacht. Wann war ich das letzte Mal Spazieren?  
5663 Inde und Malo waren vorhin auch schon wieder maximal  
5664 gestresst. Die kann ich jetzt nicht noch danach fragen, den  
5665 Beitrag Korrektur zu lesen. Kaffee. Ressourcenknappheit.  
5666 „Schlaf gut!“

## 5:42

5667 „Du bist noch wach. Der Kühlschrank ist auf. Im Bad regnet  
5668 es rein. Warum?“  
5669 Warum? So ein Chaos. Nicht nur in meinem Kopf. Raum, ich  
5670 brauche Raum. Erstmal Kaffee. Das Handy klingelt. Inde ist  
5671 dran. Sie fragt, ob ich schon wach bin. Schon? Ich bin noch  
5672 wach. Ich meckere sie an. Sie will mich zum Frühstück treffen,  
5673 um den Tag zu besprechen und ich will einfach nur schlafen.  
5674 „Du musst nicht stark sein. Wie fühlst du dich? Du hast Inde  
5675 gerade nicht zum ersten Mal angemeckert. Lass deinen  
5676 Stress nicht an anderen aus. Das führt zu nichts.“  
5677 Ich pack's nicht. Ich fühle mich beschissen. Ich brauche eine  
5678 Pause von meinem Leben. Ich muss was ändern. Loslassen –  
5679 lernen mich abzugrenzen. Aber Abgrenzung bedeutet auch  
5680 loslassen, bedeutet Verantwortung abgeben, bedeutet Ab-  
5681 schied, bedeutet loslassen. Verantwortung. Selbstverantwor-  
5682 tung – geteiltes Leid ist halbes Leid. Ich werde mein Leid teilen,  
5683 mich mitteilen. Kommunikation. Miteinander. Ich werde mir  
5684 eine Therapeut\*in suchen.

Annika Kühn ist Studentin der Soziokulturellen Studien an der Europa  
Universität Viadrina Frankfurt Oder. Dabei geht sie ihrem  
Interesse an Feminismus, Menschenrechten, Philosophie und Urban  
Studies nach. Seit 2018 ist sie Mitglied bei netzforma\* e.V. –  
Verein für feministische Netzpolitik.



# LITERATUR

- Haraway, Donna (1991)  
„A Cyborg Manifesto Science, Technology,  
and Socialist-Feminism in the Late  
Twentieth Century“, in: dies. (Hg.):  
*Simians, Cyborgs and Women:  
The Reinvention of Nature*,  
New York: Routledge, 149-181.
- Latour, Bruno (2012)  
„Love Your Monsters – Why We Must Care for Our  
Technologies As We Do for Our Children  
in: *The Breakthrough Institute*,  
14.02.2012. Online unter:  
[https://thebreak-through.org/journal/  
issue-2/love-your-monsters](https://thebreak-through.org/journal/issue-2/love-your-monsters),  
[abgerufen am 07.09.2020].

KI CAN'T CARE.

MÜTTERLICHKEIT  
IM  
ZEITALTER  
KÜNSTLICHER  
INTELLIGENZ



HANNAH  
LICHTENTHÄLER

# MÜTTERLICHE VORBILDER

Mutter – eine Rolle, eine Identität, eine Berufung? Laut Google sind Mütter wie Engel, wie Bäume, wie Blumen und Muttersein ist anstrengend, schwer und hart.<sup>1</sup> Dieses Paradox zeigt die Doppelmoral, die Müttern aufgebürgt wird, sehr deutlich: Wer Mutter ist, ist für Fürsorge, Freude und Bodenständigkeit zuständig, gleichzeitig ist Muttersein mit Anstrengung verbunden. Klar ist, dass Mutterschaft (als Kategorie) der Femininität zugeschrieben, genauso wie andersherum Femininität Muttersein kategorisch zugeordnet wird. Es ist eine vergeschlechtliche Zuschreibung, die die Binarität der Geschlechter synonym zu Frau als biologische Einheit zementiert. Im feministischen Diskurs ist Mutterschaft zwischen „Gleichheits-, differenzfeministischen und poststrukturalistischen Ansätzen“ umstritten, nicht zuletzt da sie oft im Zusammenhang mit der körperlichen Erfahrung der Schwangerschaft gesetzt wird und noch immer heteronormativ besetzt ist (Krüger-Kirn & Wolf 2018). Die Unterscheidung zwischen Mutterschaft und Mütterlichkeit ist dabei wichtig, denn letzteres sei weder an ein bestimmtes Geschlecht noch an Care-Verantwortung für eigene Kinder gebunden und könne außerdem auch in der Rolle als Onkel, erwachsene Freundin oder Mentorin verwirklicht werden (Grobner 2020).<sup>2</sup>

Muttersein ist so alt wie die Menschheit, so auch die Erzählungen darüber. Globalisierung hat viele Aspekte, die es zu kritisieren gibt, vom Kolonialismus bis zur Klimakrise. Doch sie hat uns auch digital globalisiert und dadurch Geschichten auf unsere Bildschirme gebracht, mit denen wir Identifikation für marginalisierte Perspektiven erhalten können, die es im linearen deutschen *weißen* Mainstream-Fernsehen so nicht gibt. „Film und Fernsehen beinhalten ein kulturdiagnostisches Potenzial, in dem sie Einblicke in kulturelle und gesellschaftliche Problemlagen vermitteln und kulturelle wie gesellschaftliche Debatten aufzeigen“, erklärt Krüger-Kirn. Doch schauen wir in popkulturelle Medien, suchen wir meist vergebens nach Repräsentationen, die über die konventionelle heteronormative Bilderbuchfamilie hinausgehen. Wir sehen Mutterfiguren in verschiedenen Rollen im Fernsehen – dank der immer weiter wachsenden Nachfrage nach digitalen Serienformaten können wir mit einem Klick unsere Serienliebblinge á la carte auswählen, ob in Mediatheken, bei Netflix, Amazon Prime oder anderen Pay-TV-Kanälen. Sie sind unsere alltäglichen Begleiter, vor allem in Zeiten des zunehmenden Zuhausebleibens wie derzeit in der Corona-Pandemie. Ob *Modern Family* oder *Black'ish* das Familienevent zum Abendbrot sind oder wir alleine *The Handmaids Tale* (*Der Report der Magd*) binge-watching – die Auswahl ist vermeintlich endlos. Gerade Fernsehserien können uns so gut im Alltag begleiten, da sie durch ihr serielles Erzählen sich wiederholende Erzählstränge haben oder durch das Episodenformat selbstbestimmt portioniert werden können. Anders als Filme erlauben sie die Entwicklung komplexer Charaktere oder Handlungsstränge über einen längeren Zeitraum. Erzählungen über Familien gibt es zahlreiche, doch wenige zeigen Patchwork-Familien, in denen Latinx-Personen in den Hauptrollen sind und gleichzeitig queere Elternschaft thematisiert wird, wie es *Modern Family* oder auch *Once Upon a Time* erfolgreich zeigen. Viele Serien



handeln von komplexen Vaterfiguren, nur selten von Müttern in ihrer Komplexität. Häufig treibt die Abwesenheit der Mutterfigur die Handlung an wie zum Beispiel in *Full House* oder aber Mütter in heteronormativen Familienkonstellationen werden lediglich als Hausfrau ohne eigene Bedürfnisse oder als Powerfrau, die Beruf, Kindererziehung und Haushalt komplett alleine managed, gezeigt. Eine der erfolgreichsten Serien, die eine innige Mutter-Tochter-Beziehung in all ihrer Vielfalt in den Vordergrund gestellt hat, ist *Gilmore Girls*. Auch wenn die Serie stark in einer postfeministischen Tradition steht, die neo-liberale Werte des amerikanischen kapitalistischen Systems spiegelt und blind gegenüber Themen wie Rassismus, Klassismus, oder Fettphobie bleibt, ist sie Teil der popkulturellen Erzählung, die Muttersein lebensnah in Alltagssituationen schildert. 20 Jahre ist das Debüt von *Gilmore Girls* her. Seitdem hat es wenige innovative Erzählungen von Mutterschaft gegeben. Aktuelle Beispiele geben Hoffnung auf eine feministischere Erzählung von Mutterschaft und Elternsein in ihrer Komplexität in unseren alltäglichen Unterhaltungsformaten. Dazu gehören: die australische Serie *The Letdown (Milcheinschuss)*, die alltägliche Situationen neuer Eltern wie Stillen, Schlaflosigkeit, Karrierepläne, Beziehungskonflikte und auch die Beziehung zu anderen Eltern mit gleichaltrigen Babys zeigt; die kanadische Serie *Working Moms*, in der sich vier Mütter über Karriere, Vereinbarkeit von Beruf und Familie, Stillen, queere Elternschaft, oder Vaterfiguren in ihrer Mutter/Vater-Kind-Gruppe austauschen; die amerikanische *Good Girls*, in der drei Mütter aus finanzieller Not und Sorge um ihre Familien einen Supermarkt ausrauben und dabei ihre Rollen in den jeweiligen Familienkonstellationen aushandeln; die amerikanische in den 90er Jahren spielende Serie *Little Fires Everywhere*, die Klassismus, Rassismus, Heteronormativität, Mutterschaft, Leihmutterschaft und Karriere thematisiert; sowie *Jane the Virgin*, die sich um den drei Generationen-Haushalt der Villanueva Frauen und ihre Perspektive u.a. auf Mutterschaft, künstliche Befruchtung, Patchwork-Familie oder Schwangerschaftsabbruch dreht.

1 Automatische Vervollständigung bei der Eingabe in der Suchmaschine nach den Worten "Mütter sind (...)" und "Mutterschaft ist (...)" im Oktober 2020 über google.de.

2 Für eine kulturhistorische und diskursive Auseinandersetzung über Mutterschaft und Mütterlichkeit Selbst- und Fremdbestimmung im Rahmen von Mutterschaft und Schwangerschaft sowie intrapsychisches und intersubjektives Erleben von Mutterschaft, siehe Krüger-Kirn & Wolf (2018).

# FEMINISMUS UND MUTTERSCHAFT

Innerhalb der feministischen Kreise ist Mutterschaft eher wenig sichtbar. Zum einen liegt es vermutlich daran, dass der Kampf um reproduktive Rechte und die Legalisierung des Schwangerschaftsabbruchs selten mit Themen der Elternschaft zusammen gedacht werden. Und das, obwohl viele Schwangerschaften von Personen abgebrochen werden, die bereits Eltern sind, häufig aus finanziellen Gründen (Profamilia 2018). Zum anderen liegt es daran, dass viele Frauen\* noch immer gegen die Stigmatisierung zu kämpfen haben, keine Kinder haben zu wollen. Eine Zukunft sollte die Entscheidung für oder gegen Kinder allen Menschen gleichermaßen zugänglich machen, immer und überall, unabhängig von kulturellem oder sozialem Kontext. Das Konzept der reproduktiven Gerechtigkeit<sup>3</sup>, das Ende der 90er Jahren von BIPOC Feminist\*innen zusammen mit LGBTIQ+ Communities und anderen marginalisierten Gruppen, zusammengeschlossen als *SisterSong*<sup>4</sup>, in den USA begründet wurde, sollte Teil unserer feministischen Zukunft sein. Es vereint die intersektionalen Kämpfe, für BIPOC Frauen und trans\*, inter und nicht-binäre Personen sich für Kinder zu entscheiden, es kämpft gegen rassistische Ressentiments in der Geburtshilfe, genauso wie es für das Recht auf den Schwangerschaftsabbruch und für die Gerechtigkeit in der Selbstbestimmung über den eigenen Körper einsteht. Dass dies eine Bewegung ist, der sich auch *weiße* Feminist\*innen anschließen können, in Anerkennung an die bisherigen Erfolge der Schwarzen Bewegung aus der das Konzept entstanden ist, können wir in Zukunft lernen – auch in Deutschland. „I am not free while any woman is unfree, even when her shackles are very different from my own“, hat Audre Lorde bereits 1981 gesagt. Dies müssen auch *weiße* Frauen\* anerkennen und Allianzen für die Selbstbestimmung, reproduktiven Rechte und soziale Gerechtigkeit schließen.

Digitalisierung hat das Potential, die Zukunft von Mutterschaft maßgeblich zu gestalten und dadurch gerechter, feministischer zu machen. Zum einen gibt es zahlreiche Aktivist\*innen, die über soziale Medien Bildungs- und Aufklärungsarbeit leisten, u.a. zu Themen rund um reproduktive Gerechtigkeit, Geburtshilfe, Schwangerschaft, Schwangerschaftsabbruch, queere Elternschaft sowie allgemein auch das alltägliche Leben mit Kindern. Egal ob Blogger\*innen, Influencer\*innen oder Aktivist\*innen – das Internet macht es möglich, sich zu informieren, auszutauschen, sich nicht allein zu fühlen, über nationale Grenzen hinweg. Die stetige Weiterentwicklung digitaler Tools und Methoden kann das Leben erleichtern und unterstützen. Unser Zeitalter der Digitalisierung ist nicht mehr ohne Künstliche Intelligenz (KI) denkbar. Sie findet mittlerweile in vielen Bereichen Anwendung, seien es Vorhersagen über unsere Fruchtbarkeit oder über die Gesundheit eines Fötus und die Prognose ist: KI wird unsere Zukunft bestimmen. Dies wirft Fragen auf: Wer programmiert diese Algorithmen und wer trainiert die Datensätze, mit denen die KI derartige Vorhersagen trifft, auf denen Entscheidungen über Leben und Tod basieren? Wer bleibt in Entscheidungs-



5831 prozessen in einer aktiven Rolle einbezogen? Wie kommen KI  
5832 und Mutterschaft zusammen und wie entwickeln wir diesen  
5833 Zusammenhang feministisch?

3 Über das Konzept der reproduktiven Gerechtigkeit schreiben Loretta J. Ross und Rickie Solinger ausführlich in ihrem Buch "Reproductive Justice: An introduction". Es geht um die intersektionale Analyse von Rassismus, Klassismus und Sexismus mit Fokus auf die Erfahrungen von Frauen of Color. Sie grenzen dabei die Bewegung der reproduktiven Gerechtigkeit von der Pro-Choice-Bewegung ab. Das Konzept kombiniert den Kampf um reproduktive Rechte und soziale Gerechtigkeit. (Ross & Solinger 2017)

4 Weitere Informationen auf der Webseite über Reproductive Justice (SisterSong).

# ALTHEILMITTEL GEGEN UNFRUCHT- BARKETT?

5834 Operieren bald nur noch Roboter? KI rettet mittlerweile Leben  
5835 oder zumindest hilft sie beim Erhalt von Lebensqualität. KI  
5836 kann das Leben mit Diabetes erleichternd unterstützen (Con-  
5837 treras 2018) und vor allem in der Medizin bei der Früherken-  
5838 nung von Karzinomen helfen, da Algorithmen viel zuverlässi-  
5839 ger und schneller Muster für Tumore erkennen können, als das  
5840 menschliche Auge. Vergessen wir nicht: errare humanum est  
5841 (non secundum apparatus). Das darf aber bei weitem nicht  
5842 heißen, dass Entscheidungen irgendwann von Maschinen  
5843 getroffen werden. Auch in der Reproduktionsmedizin spielt  
5844 Technik eine immer stärker werdende Rolle. Wer schon einmal  
5845 in einem Kreißaal entbunden hat und viele Stunden Wehen  
5846 und Herzschlag überwachen lassen musste, kennt das: Heb-  
5847 ammen fehlen an allen Stellen. Eine Hebamme beobachtet  
5848 mehrere Entbindende und den Herzschlag der Babys gleich-  
5849 zeitig per Monitor und nur, wenn maschinell eine Unregelmä-  
5850 ßigkeit gemeldet wird, kümmert sich die Hebamme. Fürsorge  
5851 im Kreißaal kann aufgrund der personellen Bedingungen  
5852 kaum noch geleistet werden, doch auch keine Maschine kann  
5853 diesen Moment kurz vor und nach der Geburt fürsorglich  
5854 begleiten. Der Diskurs rund um Geburtshilfe ist hoch politisiert,  
5855 allerdings findet die Berücksichtigung aller Aspekte rund um  
5856 die Geburt, wohl bemerkt dem Ursprung unser aller Leben, bei  
5857 politischen Entscheidungen kaum bis gar keine Rolle. Kreiß-  
5858 säle schließen in ländlichen Regionen<sup>5</sup>, sodass Entbindene  
5859 unter Umständen hunderte Kilometer ins nächste Kranken-  
5860 haus fahren müssen. Kaum eine schwangere Person findet in  
5861 Berlin eine Schwangerschaftsbetreuung von einer Hebamme,

da es aufgrund teils katastrophaler Arbeitsbedingungen und fehlender politischer Maßnahmen für faire und angemessene Bezahlung immer weniger von ihnen gibt. Wie wäre es denn, wenn uns eine KI ausrechnet, wie viele Hebammen an welcher Stelle gebraucht werden, und wie finanzielle und zeitliche Ressourcen bestmöglich eingesetzt werden können, um dem Aussterben einer der wichtigsten Care-Berufe entgegen zu wirken, statt zu schauen, an welcher Stelle KI Personal ersetzen kann? Es scheint eine banale Rechnung, die eine KI sofort erlernen kann. Tatsächlich hat beispielsweise der Deutsche Hebammen Verband e.V. auf seiner Webseite ein Tool entwickelt, das die aktuellen Bedingungen der Geburten in Deutschland kartiert und so anhand der Datensammlung Unterversorgung direkt darstellt. Derartige Berechnungen werden politisch jedoch zu wenig beachtet, da es das patriarchale System, das die Geburt auch nur als Teil des kapitalistischen Zyklus von Produktion und Reproduktion sieht, für undenkbar radikal hält.

5 „1991 gab es noch 1186 Kliniken, in denen Geburten möglich waren. 2017 waren es nur noch 672 Kliniken mit Geburtshilfe. Seitdem schließt fast jeden Monat ein Kreißsaal ganz oder vorübergehend die Türen (Deutscher Hebammen Verband e.V.).“

# NEOLIBERALES VERSPRECHEN DER REPRODUKTION

Befürworter\*innen von KI-Einsätzen in der Geburtsmedizin schwärmen vielleicht davon, dass KI entscheiden kann, ob eine Geburt natürlich oder per Kaiserschnitt durchgeführt werden kann. Durch maschinelles Lernen sollen Interventionen und Komplikationen auf das notwendige Minimum reduziert werden können. Die Idee, oder besser das Ideal, lautet: Wenn KI fetale Bewegungen, Atemmuster und Biosignale wie Herzfrequenz oder Blutdruck lesen und zuverlässig entscheidende individuelle Muster in der Physiologie, den Emotionen und Verhaltensweisen von Mutter und Baby erkennen könne, und so genau erlerne, welche Kombinationen von Mustern zu welchem Ergebnis führen, könnte ein solches System verwendet werden, um zu bestimmen, was während der Wehen zu tun sei (Topalidou & Downe 2019). Darüber hinaus kann KI auch schon vor der Geburt beispielsweise bei Plazentauntersuchungen auf Unregelmäßigkeiten auf einem Computerbild schneller aufmerksam machen und so im Zweifel Erkrankungen frühzeitig erkennen (Stephens 2020). Am MIT in Boston

hat eine Gruppe von Forscher\*innen bereits einen KI-Roboter für den Kreißsaal entwickelt. In einer Studie kam heraus, dass die beteiligten Ärzt\*innen und Pfleger\*innen Empfehlungen dieses Roboters in 90% der Fälle akzeptierten, gleichzeitig kam aber auch heraus, dass die Fehlerquote ähnlich hoch war, unabhängig von der Anwesenheit des Roboters. Daraus schlossen sie zwar, dass es sicher sei eine KI in der Geburtshilfe einzusetzen (Topalidou & Downe 2019), doch warum sich auf KI stützen, wenn sie am Ende nicht besser agieren kann als Menschen? Hinzu kommt der emotionale Aspekt einer Geburt, den keine KI je wird ersetzen können. Es ist bekannt, dass eine fürsorgliche Begleitung genauso wie eine emotionale und psychologische Unterstützung sowohl für Entbindene als auch für Säuglinge die gesundheitlichen Folgen der Geburt verbessern und auch langfristige Auswirkungen auf das Erwachsenwerden des Neugeborenen haben können (ebd.). Keine schwangere Person sollte während der Geburt allein auf Alexa oder Siri als Geburtshelferin angewiesen sein müssen, das heißt ohne jeglichen menschlichen Kontakt.

Ein zentraler Teil der Reproduktionsmedizin ist künstliche Befruchtung. Forscher\*innen preisen KI als wesentlichen Teil der Lösung für ungewollte Kinderlosigkeit in der Zukunft an. Es wurden bisher schon mehrere Techniken maschinellen Lernens bei künstlichen Befruchtungen eingesetzt, um die Leistung der assistierten Reproduktionstechnologie (ART) zu verbessern (Wang et al. 2019). Auch wenn es nach wie vor viele Herausforderungen und Probleme gibt, haben Entwicklungen der ART wie die assistierte Befruchtung, genetische Präimplantationstests und Technologien zur Embryonenauswahl, die klinische Schwangerschaftsrate in den letzten 40 Jahren seit der Geburt des ersten Babys mit In-vitro-Fertilisation (IVF) stark verbessert (ebd.). Noch immer ist es schwierig, die Wahrscheinlichkeit einer erfolgreichen Schwangerschaft – genauso wie die Ursache für jeden Misserfolg – vorherzusagen oder zu verstehen. Ein Ansatz der KI-basierten Methoden ist es, Daten komplexer Diagnosen und Therapiebehandlungen zu sammeln und auszuwerten, um Unfruchtbarkeit bei Patient\*innen besser behandeln und prognostizieren zu können. Dabei kann diese KI effizienter und wirksamer den Behandlungszyklus der ART optimieren (ebd.). Eine Verbesserung der IVF durch die Nutzung von Algorithmen für bessere Vorhersagen über den besten Zeitpunkt im Zyklus für die Empfängnis könnte für Betroffene finanzielle Entlastung durch niedrigere Kosten bedeuten sowie die Minderung traumatisierender Erfahrungen durch Frühaborte (Ducharme 2019).

Weitaus bekannter ist die Anwendung von Algorithmen bereits im Bereich des Menstruationstrackings durch Apps wie Flo, Clue oder drip. In diesen Apps können Menstruierende ihren Zyklus, körperliche Symptome, sexuelle Aktivität und ihre Fruchtbarkeit, z.B. für natürliche Familienplanung (NFP), überwachen. Diese Apps sind mittlerweile für den privaten Gebrauch der hormonfreien Verhütung und/oder der Zyklusüberwachung<sup>6</sup> weit verbreitet, werden darüber hinaus aber auch im Bereich der künstlichen Befruchtung angewendet, um Fruchtbarkeit noch besser zu überwachen. Apps wie Mira kombinieren die Selbstberichte mit Urintests der User\*innen zu Hause, mit denen die App den Hormonspiegel verfolgen kann, um Fruchtbarkeit noch genauer vorhersagen zu können (ebd.). Dr. Jessica Spencer, Direktorin der Abteilung für reproduktive Endokrinologie und Unfruchtbarkeit an der medizinischen Fakultät der Emory Universität, erkennt das große Potential von KI für künstliche Befruchtung, da Algorith-

5961 men die notwendigen Variablen errechnen kann, die Unfrucht-  
5962 barkeitsrisiken weit im Voraus einschätzen kann, genauso wie  
5963 ein Protokoll für Menschen mit Uterus, die versuchen, per IVF  
5964 schwanger zu werden, individuell zuschneiden kann (ebd.).  
5965 Wie hilfreich KI beim Einsatz dieser Datenverarbeitung ist,  
5966 zeigen diese Entwicklungen deutlich. Trotz dieser zunehmend  
5967 wichtigen Rolle von KI in der Medizin wird sie Ärzt\*innen in  
5968 Zukunft wohl kaum ersetzen (ebd.).

6 Für Menstruierende bietet diese Art des Zyklustrack-  
ings auch den Vorteil, bei ihren Gynäkolog\*innen konkrete  
Angaben über den Gesundheitszustand machen zu können, was  
z.B. für Endometriosepatient\*innen überaus wichtig ist, da  
Ärzt\*innen auf derartige Schmerzsymptome selten adäquat re-  
agieren. Wer starke Menstruationsbeschwerden hat, kann  
zudem anhand der Vorhersagen über die kommenden Perioden  
wichtige Termine besser koordinieren.

# FRUCHTBARKEITS- DISKURS BRAUCHT FEMINISMUS

5969 Wie so oft fehlen in solchen Diskussionen und Zukunftsmelo-  
5970 dien sichtbare feministische Perspektiven. Bei Tracking-Apps  
5971 geht es nämlich in der Regel um Kaufkraft und Daten, statt  
5972 Probleme der Gesundheitsversorgung zu lösen (Kochsiek  
5973 2019). Wie die Algorithmen die eingegebenen Daten über die  
5974 Menstruation der User\*innen auswerten, ist überdies sehr  
5975 intransparent. Einige der Apps bieten sogar das Teilen der  
5976 Daten mit Facebook oder Google an, die großes Interesse an  
5977 den privaten Datensätzen haben. Bisher haben zwar Kranken-  
5978 kassen keinen Zugriff auf solche Gesundheitsdaten, doch das  
5979 Nutzungsverhalten wird für gezielte Werbung, durch die sich  
5980 solche Apps meist finanzieren, bereits verwendet (ebd.).  
5981 Datenschutz und -sicherheit müssen im Vordergrund stehen  
5982 und nicht verkauft werden, um die neoliberale Marktklogik mit  
5983 mehr und mehr Daten zu füttern. „Denn es braucht mehr Tech-  
5984 nik-Transparenz und aufrichtige Algorithmen, die auf die  
5985 Begrenztheit ihrer Aussagekraft hinweisen oder zu vage Aus-  
5986 sagen gar nicht erst treffen“, fordert Kochsiek (ebd.). Es gibt  
5987 natürlich Ansätze wie etwa die Tracking-App drip, eine Open-  
5988 Source Alternative, die Daten nur lokal auf dem Smartphone  
5989 speichert und genauso funktioniert wie all die anderen Apps.

5990 Bei der Diskussion um IVF und KI fehlt zudem ganz  
5991 klar eine kritische Perspektive darauf, wie fremdbestimmt sie  
5992 laufen. Die uralte Sage der tickenden biologischen Uhr gilt  
5993 immer noch als Grundlage für die Errechnung von Fruchtbar-  
5994 keit, und dabei wird den Menschen mit Uterus in der Regel ein

5995 Zeitfenster von 7–10 Jahren gegeben, obwohl dies fern von der  
 5996 Lebensrealität der meisten ist. Der Druck auf Menschen mit  
 5997 Kinderwunsch bettet sich ein zwischen biologisch-zeitlichem  
 5998 Narrativ, dem Baby als Karrierekiller (Rosales 2020) und neu-  
 5999 erdings auch den negativen Auswirkungen auf das Klima  
 6000 (Bücker 2020)<sup>7</sup>. Gleichzeitig fehlt eine öffentliche Diskussion  
 6001 über das Tabuthema Fehlgeburt. Allein linguistisch steckt ein  
 6002 stigmatisierendes Narrativ hinter dem (spontanen) Abort,  
 6003 über den sich kaum Menschen zu sprechen wagen – denn vor  
 6004 den ersten drei Monaten sollte sowieso niemand über Schwan-  
 6005 gerschaft und die möglichen Folgen sprechen. Tick tock, eine  
 6006 tickende Zeitbombe jagt die andere. Tatsächlich erlebt knapp  
 6007 jeder dritte Mensch mit Uterus eine oder mehrere Fehlgebur-  
 6008 ten (Grobner 2020), aber weder über Fehlgeburten noch über  
 6009 ungewollte Kinderlosigkeit sprechen Mainstream und femi-  
 6010 nistische Communities gleichermaßen. „Fertility Gap heißt  
 6011 jene Lücke zwischen Kinderwunsch und tatsächlicher Kinder-  
 6012 anzahl“, erklärt Grobner, und er ist vor allem unter Akademi-  
 6013 ker\*innen in Deutschland, Österreich und der Schweiz beson-  
 6014 ders groß (ebd.). Lebensrealitäten wie etwa von lesbischen  
 6015 cis Frauen, die häufig von ungewollter Kinderlosigkeit betrof-  
 6016 fen sind, oder von trans\*Personen<sup>8</sup>, werden in Studien über  
 6017 Fruchtbarkeit erst gar nicht berücksichtigt. Diverse Lebens-  
 6018 realitäten von allen Personen mit Kinderwunsch müssen den  
 6019 Zukunftsvisionen von KI als Allheilmittel gespiegelt werden  
 6020 und die Diskussionen im Mainstream und innerhalb feministi-  
 6021 scher Kreise müssen offen und transparent geführt werden.  
 6022 Sie müssen außerdem auch antirassistisch geführt werden,  
 6023 denn häufig kristallisiert sich eine biologistische Logik auf  
 6024 Elternschaft heraus, die vor allem im Westen vorherrscht.

7 Frauen mit Kinderwunsch werden nun auch als egois-  
 tisch bezeichnet, da die CO<sub>2</sub>-Bilanz von Kindern so schlecht  
 für die Umwelt ist – statt Unternehmen und Regierungen werden  
 also gebärfähige Menschen in die Verantwortung gezogen?  
 „Der Wunsch nach Familie ist kein Thema, das nur Menschen be-  
 trifft, die schwanger werden können, doch da gesellschaftlich  
 meist cis Frauen als die Entscheider\*innen über Schwangerschaften  
 gesehen werden, sie Babys gebären und so sichtbar neues  
 Leben in die Welt tragen, ist der Appell, dem Klimaschutz zugunsten  
 auf Kinder zu verzichten, im Kern eine antifeministische Argu-  
 mentation,“ erklärt Teresa Bücker. Zeugungsfähige Männer sind  
 in der Gleichung nicht die Egoisten (Bücker 2020).

8 Beispielsweise schreibt Benjamin Czarniak, der sich  
 als trans\*Mann identifiziert, in dem Sammelband „Nicht nur  
 Mütter waren schwanger“ (edition assemblage) über den Schmerz  
 nach seiner ersten Fehlgeburt (Grobner 2020).



# WENN KI, DANN FEMINISTISCH

Mutterschaft und KI gehören in einer feministischen Zukunft zusammen. Und das nicht, weil Algorithmen schlauer sind als Menschen und aufgrund ihres online Kauf- und Suchverhaltens früher wissen, ob jemand schwanger ist, als die Person selbst. Ein Projekt in Kenia hat zum Beispiel gezeigt, dass KI einen Chatbot trainieren kann, der typische Fragen, die viele Eltern während einer Schwangerschaft und in den ersten Monaten nach der Geburt über das Baby haben, beantwortet – viele Menschen ohne höheren Bildungsabschluss verfügen nicht unbedingt über den Zugang zu neuesten Technologien oder zum Internet, deshalb funktioniert dieser Service per SMS (Rajasekharan 2019). Auch in Europa und den USA gibt es einen solchen Chatbot in der App Muse<sup>9</sup>, die es auch als SMS Funktion gibt, allerdings einen monatlichen Beitrag erfordert (Anderson 2018).

Technik kann nicht den sozialen Aspekt von Mutterschaft ersetzen, denn Technik kann nicht fürsorgen. Dystopische Sci-Fi-Szenarien gaukeln uns vor, dass das böse Erwachen bevorsteht, in dem Android\*innen Föten mit dem vielversprechenden Genmaterial im Reagenzglas heranzüchten (wie der Film *I am Mother* thematisiert) und die menschliche Mutter ersetzen werden. Doch welche Fragen tatsächlich gestellt werden müssen, vor allem im *weißen* westlichen Narrativ von Mutterschaft, sind: wer darf Mutter sein und was können wir hier in Deutschland noch dazulernen aus Bewegungen wie der reproduktiven Gerechtigkeit? Wie können innerhalb feministischer Diskussionen auch Perspektiven von Elternschaft und Mutterschaft zugelassen werden, zum einen diese nach dem Wunsch von Kindern, sowie nach dem Wunsch, keine Kinder zu bekommen, aber auch zur Dekonstruktion naturgegebener Mutterliebe, wie es z.B. der Hashtag-Trend #Regrettingmotherhood gezeigt hat?

Wenn Firmen damit werben, wie gut KI und Roboter in der Kindererziehung eingesetzt werden können (Kadyrov 2019), seien es die Kameras im Babybett, die Matratze, die den Herzschlag des Kindes mit überwacht, oder die GPS-Tracking-App auf dem Smartphone der jugendlichen Kinder, ist dies hoch problematisch und bedarf unserer Aufmerksamkeit. KI darf nicht zur Überwachung von Kindern ausgenutzt werden und darüber hinaus noch diese Daten an Unternehmen geben, die nicht transparent machen, was mit den Daten passiert oder sie zu Werbezwecken verkaufen. Eine Zukunft der digitalen Welt ist transparent bezüglich der Daten, Überwachung ist reguliert und Unternehmen schöpfen keinen Profit aus den persönlichen Daten der Menschen.

Klar ist auch: Care-Arbeit muss aufgewertet werden, sowohl die bezahlten Berufe im Care Bereich wie Krankenpflege, Altenpflege und Kindererziehung, aber vor allem auch die unbezahlte Care-Arbeit und mentale Arbeit (Stichwort Mental Load), die vor allem auf den Schultern von Frauen\* lastet. Wie können wir digitale Technologien wie KI einsetzen, damit Mutterschaft davon profitiert? Oder muss die Frage vielleicht auch lauten: Wie kann Mutterschaft KI beeinflussen,

6078 damit sie gerechter, feministischer wird? Wie und wo würde  
 6079 KI eingesetzt und bedarfsorientiert programmiert werden,  
 6080 wenn diese Perspektiven auf der Entscheidungsebene ver-  
 6081 treten sind? Erinnern wir uns an die mütterlichen Vorbilder aus  
 6082 unseren Lieblings-Fernsehserien: eine feministische Zukunft  
 6083 zeigt Mutterschaft, Mütterlichkeit, Elternschaft und Familie  
 6084 vielfältig, weder nur positiv, noch nur negativ, schafft Vorbilder  
 6085 für alle Menschen und erlaubt uns, in und mit diesen Erzählun-  
 6086 gen auch mal von unserem Alltag abzuschalten. Denn eines  
 6087 bleibt: eine feministische Zukunft der KI ist nur mit Mutter-  
 6088 schaft denkbar und Mutterschaft ist nur dann zukunftsfähig,  
 6089 wenn die gesellschaftliche Vorstellung von ihr dekonstruiert  
 6090 und queere, trans\* und nicht-binäre Menschen miteinschließt.  
 6091 Das heißt für uns: Wenn KI, dann feministisch.

9 Muse hat ca. 2.700 User\*innen: 55% in den USA, 27%  
 in Deutschland und 13% in anderen Ländern weltweit (Anderson).

Hannah Lichtenthäler ist Projektkoordinatorin im Gunda-Werner-  
 Institut für Feminismus und Geschlechterdemokratie. Sie  
 studierte Nordamerikastudien mit Schwerpunkt auf feministische  
 Medien und Populärkultur in Berlin, England und Michigan,  
 und absolvierte zuvor eine Ausbildung in Madrid. Bei netzforma\*  
 ist sie im Vorstand.

# LITERATUR

- Anderson, Jenny (2018)  
 „Can an AI-Powered Bot Help Parents Raise Better Humans?“ *Quartz*, 14.  
 März. 2018, [qz.com/1227955/muse-an-ai-powered-parent-  
 ing-bot-wants-to-help-parents-help-their-children-succeed/](https://qz.com/1227955/muse-an-ai-powered-parenting-bot-wants-to-help-parents-help-their-children-succeed/).
- Bücken, Teresa (2020)  
 „Kinderfrei Fürs Klima? Warum Wir Bei Einfachen Lösungen Skeptisch  
 Sein Sollten.“ *EDITION F*, 16. März 2020, [editionf.com/  
 entscheidung-gegen-kinder-klima-erderwaermung/](https://editionf.com/entscheidung-gegen-kinder-klima-erderwaermung/).
- Contreras, Ivan & Vehi, Josep (2018)  
 „Artificial Intelligence for Diabetes Management and Decision Support:  
 Literature Review.“ *Journal of Medical Internet Research*,  
 vol. 20, no. 5, 2018, doi:10.2196/10775.
- „Gegen Kreißsaalschließungen.“  
 Hrsg. Deutscher Hebammen Verband e.V., Unsere Hebammen, Deut-  
 scher Hebammen Verband E.V., [www.unsere-hebammen.  
 de/mitmachen/kreisssaalschliessungen/](http://www.unsere-hebammen.de/mitmachen/kreisssaalschliessungen/).
- Ducharme, Jamie (2019)  
 „How Artificial Intelligence Could Change Fertility.“ *Time*, 11. Juli 2019,  
[time.com/5492063/artificial-intelligence-fertility/](https://time.com/5492063/artificial-intelligence-fertility/).
- Grobner, Cornelia (2020)  
 „Hoffnung. Enttäuschung. Trauer. Wut. Repeat.“ *An.schläge – Das  
 Feministische Magazin*, 9. Okt. 2020, [anschlaege.at/hoff-  
 nung-enttaeuschung-trauer-wut-repeat/](https://anschlaege.at/hoffnung-enttaeuschung-trauer-wut-repeat/).

- \_\_\_\_\_ Kadyrov, Il (2019)  
„Super Parents With AI Robots.“ *Medium*,  
12 May 2019, [medium.com/@mrcrambo/  
super-parents-with-ai-robots-566257d0fcb8](https://medium.com/@mrcrambo/super-parents-with-ai-robots-566257d0fcb8).
- \_\_\_\_\_ Kochsiek, Marie (2019)  
„Menstruationszyklen Entziffern: Gunda-Werner-  
Institut.“ *Heinrich-Böll-Stiftung: Gunda-  
Werner-Institut*, 18. Jan 2019,  
[www.gwi-boell.de/de/2019/01/18/  
menstruationszyklen-entziffern](http://www.gwi-boell.de/de/2019/01/18/menstruationszyklen-entziffern).
- \_\_\_\_\_ Krüger-Kirn, Helga & Wolf, Laura (Hg.) (2018)  
*Mutterschaft Zwischen Konstruktion Und  
Erfahrung: Aktuelle Studien Und  
Standpunkte*. Opladen: Verlag  
Barbara Budrich.
- \_\_\_\_\_ Profamilia (Hg.) (2018)  
„Acht Fakten Zum Schwangerschaftsabbruch:  
Factsheet Schwangerschaftsabbruch:  
Zahlen Und Hintergründe,“  
[www.profamilia.de/publikationen.html?  
tx\\_pgextendshop\\_pi1%5Bproduct%5D=  
203&tx\\_pgextendshop\\_pi1%5Bacti-  
on%5D=show&tx\\_pgextendshop\\_pi1%5B  
controller%5D=item&tx\\_pgextendshop\\_pi1%5B  
hash%5D=8cc0c7e475537af1b5b5fc666d0](http://www.profamilia.de/publikationen.html?tx_pgextendshop_pi1%5Bproduct%5D=203&tx_pgextendshop_pi1%5Baction%5D=show&tx_pgextendshop_pi1%5Bcontroller%5D=item&tx_pgextendshop_pi1%5Bhash%5D=8cc0c7e475537af1b5b5fc666d0).
- \_\_\_\_\_ Rajasekharan, Sathy (2019)  
„How AI Helps Mothers in Kenya Get the Care They  
Need, Faster.“ *Medium, Towards Data Science*,  
5 Apr. 2019, [towardsdatascience.com/  
how-ai-helps-mothers-in-kenya-get-the-  
care-they-need-faster-eb4f05b34732](https://towardsdatascience.com/how-ai-helps-mothers-in-kenya-get-the-care-they-need-faster-eb4f05b34732).
- \_\_\_\_\_ Rosales, Caroline (2020)  
„Warum Elternzeit ein irreparabler Einschnitt ist“  
*ZEIT ONLINE*, 22. Jan. 2020,  
[www.zeit.de/arbeit/2020-01/  
mutterschaft-elternzeit-schwangerschaft-  
karriere-berufseinstieg?utm\\_source=  
https%3A%2F%2Fwww.google.com%2F](https://www.zeit.de/arbeit/2020-01/mutterschaft-elternzeit-schwangerschaft-karriere-berufseinstieg?utm_source=google&utm_medium=organic&utm_campaign=2020-01-22%2Fwww.google.com%2F).
- \_\_\_\_\_ Ross, Loretta & Solinger, Rickie (2017)  
*Reproductive justice: An introduction*. Oakland,  
CA: University of California Press.
- \_\_\_\_\_ SisterSong. (n.d.)  
„Reproductive Justice.“ Retrieved Nov. 13, 2020,  
from [https://www.sistersong.net/  
reproductive-justice](https://www.sistersong.net/reproductive-justice)
- \_\_\_\_\_ Stephens, Keri (2020)  
„Using Artificial Intelligence to Protect Mothers'  
Future Pregnancies.“ *Axis Imaging News*,  
2. Sept. 2020, [axisimagingnews.com/  
market-trends/cloud-computing/  
machine-learning-ai/artificial-  
intelligence-protect-mothers-future-  
pregnancies](https://axisimagingnews.com/market-trends/cloud-computing/machine-learning-ai/artificial-intelligence-protect-mothers-future-pregnancies).
- \_\_\_\_\_ Topalidou, Anastasia & Downe, Soo  
„Could AI Take Control of Human Birth?“  
*The Conversation*, 21. Okt. 2019,  
[theconversation.com/could-ai-take-  
control-of-human-birth-108282](https://theconversation.com/could-ai-take-control-of-human-birth-108282).
- \_\_\_\_\_ Wang, Renjie, et al. (2019)  
„Artificial Intelligence in Reproductive Medicine.“  
*Reproduction*, vol. 158, no. 4,  
doi:10.1530/rep-18-0523.





FAZIT  
:  
EIN  
FEMINISTISCHER  
AUSBlick

Der Kern dieser Publikation liegt darin, aufzuzeigen, dass es eine Vielzahl an wertvollen Perspektiven aus verschiedenen thematischen wie gesellschaftspolitischen Sphären (Aktivismus, Zivilgesellschaft, Wissenschaft) und unterschiedlichen Disziplinen (Kunst, Politik, Soziologie, Kulturanthropologie, Rechtswissenschaft, Philosophie) auf die Entwicklung und Nutzung von Künstlicher Intelligenz gibt. Wir haben versucht, vor allem denen einen Raum zu geben, die im Status Quo nicht repräsentiert sind, weil patriarchale, *weiße* Strukturen und Denkmuster dies verhindern. KI stellt uns immer wieder auf's Neue vor die Frage, in was für einer Gesellschaft wir leben möchten, nicht in der Zukunft, sondern im Hier und Jetzt – und genau dies muss fortlaufend transparent und nachvollziehbar gesamtgesellschaftlich ausgehandelt werden; machtkritisch und intersektional, damit Diskriminierungsstrukturen aufgebrochen und nicht fortgeschrieben werden.

#### WIE WIRD KÜNSTLICHE INTELLIGENZ FÜR ÜBERWACHUNG EINGESETZT UND WELCHE DATENSÄTZE WERDEN FÜR DAS TRAINING DIESER TECHNOLOGIEN HERANGEZOGEN?

In unserer Publikation wird schnell deutlich, dass Überwachungstechnologien vor allem ohnehin marginalisierte Gruppen gefährden. Ihr Einsatz und die Bedingungen ihres Einsatzes müssen im Bereich der Strafverfolgung, des Sozialstaats sowie in allen anderen gesellschaftlichen Bereichen stets kritisch hinterfragt und fortlaufend unabhängig, umfassend und intersektional evaluiert werden.

#### WIE MÜNDIG UND AUTONOM AGIEREN INDIVIDUEN IN BEZUG AUF DIE NUTZUNG VON KÜNSTLICHER INTELLIGENZ UND WIE MÜNDIG UND AUTONOM KÖNNEN DIESE REAGIEREN, WENN SIE DIGITALER GEWALT DURCH KÜNSTLICHE INTELLIGENZ AUSGESETZT SIND?

Wir haben versucht aufzuzeigen, dass Technologie immer in ihren intersektional-historischen Verwobenheiten betrachtet werden muss, dass koloniale Machtverhältnisse anerkannt und aufgebrochen werden müssen. Eine differenzierte Kritik durch feministische Perspektiven, die historische Analysen von Geschlechterverhältnissen in den Blick nimmt, ist deshalb unabdingbar.

Wir wissen außerdem, dass das Ausmaß digitaler Gewalt enorm ist – und häufig übersehen wird. Dies betrifft insbesondere Frauen\*, die in sozialen Medien viel höheren Kommunikationsrisiken von Hate Speech, Doxxing und Bedrohungen ausgesetzt sind, ebenso wie Frauen\*, die im häuslichen Kontext digitale Überwachung durch z.B. Spy Apps erfahren. Dieses Thema braucht dringend mehr Aufmerksamkeit, auch mit Blick auf belegbare Zahlen zum Kontext digitale Gewalt im nationalen Raum, Forschung zu Folgen und Ausmaß sowie umfassende Hilfs- und Unterstützungsangebote für Betroffene durch spezialisierte Beratungsstellen. Weiterhin bedarf es einer Sensibilisierung von Polizei wie Justiz und, wo nötig, regulatorischer Anpassungen. Das Datenschutzrecht und das Recht informationstechnischer Systeme muss gleichheits- und antidiskriminierungsrechtlich neu gedacht werden.

#### WAS STEHT BEI AUTOMATISIERTEN ENTSCHEIDUNGSPROZESSEN IM VORDERGRUND – EFFIZIENZ ODER DER EINZELFALL? WER WIRD DADURCH DISKRIMINIERT, AUSGESCHLOSSEN BZW. BENACHTEILIGT UND WARUM?

Keine Technologie ist neutral. Automatisierte Entscheidungsprozesse befördern effiziente und diskriminierende Entscheidungen. Diese Diskriminierungen manifestie-



ren sich etwa bei der Kreditvergabe, im Job-Bewerbungsverfahren oder bei der Vergabe staatlicher Sozialleistungen. Eine Behörde, die divers und inklusiv besetzt ist, könnte diese Diskriminierungsgefahren einhegen. Sie könnte eine KI auf das Tradieren bestehender Diskriminierungsverhältnisse hin analysieren und ihren Einsatz ggf. schon vorab verbieten. Bereits bestehende Gesetze (Antidiskriminierungsgesetze, Datenschutzgesetze) und ihre Durchsetzungsmöglichkeiten müssen geprüft und angepasst werden, um einen effektiven subjektiven Rechtsschutz zu gewährleisten. Eine Voraussetzung hierzu ist, dass Entscheidungsverfahren transparent und nachvollziehbar gemacht werden. Es gibt schließlich keinen einfachen technologischen Fix für die Komplexität fairer Entscheidungsprozesse.

WIE WIRD KÜNSTLICHE INTELLIGENZ AUßERDEM IN FILMEN DARGESTELLT? INWIEFERN IST DIESE DARSTELLUNG SEXUALISIERT UND VERKÖRPERT PATRIARCHALE DENK- UND VERHALTENSMUSTER? UND WIE KANN KÜNSTLICHE INTELLIGENZ MIT QUEERNESS UND KUNST INS VERHÄLTNIS GESETZT WERDEN?

Wir haben gezeigt, dass zwischen Nullen und Einsen Spielraum für Queeres ist. KI sollte deshalb nicht rational und neutral gedacht werden; stattdessen sollte sie in ihrer intersektionalen Verwobenheit anerkannt und mit dieser gearbeitet werden.

Überdies sollten traditionelle und intersektional vermachtete Wissenssysteme sukzessive aufgelöst – inter- und multidisziplinäre Forschung gestärkt werden. (Individual- und Kollektiv-)Aktivismus, zivilgesellschaftliche Organe/Einrichtungen und Kunst brauchen dabei ebenso Raum und Anerkennung, wie etablierte quantitative und qualitative Studien. Komplexe Interdependenzen von Körpern, Umwelten und Technologien können nur gemeinsam und gemeinwohlorientiert erforscht und bearbeitet werden.

Die vorliegende Publikation will einen kritischen, intersektional-feministischen öffentlichen Diskurs über Künstliche Intelligenz und Algorithmen anstoßen. Dabei bildet diese Publikation nicht nur auf Textebene ab, was eine intersektional-feministische Auseinandersetzung mit Künstlicher Intelligenz meint – auch das Layout des Textes, die ausgewählten Schriften und die Bildsprache haben wir bewusst gewählt. Es soll deutlich werden, dass Design, Entwicklung und Kreation eine große Auswirkung auf Einsatz, Wirkung und Nutzung haben – so, wie auch das Design von Künstlicher Intelligenz ausschlaggebend für deren Einsatz, Wirkung und Nutzung ist. Wir versuchen so die Komplexität und Vielschichtigkeit in Bezug auf Intention, Bedeutung und Auswirkungen der (Weiter-)Entwicklung von Künstlicher Intelligenz deutlich zu machen.

Wir als netzforma\* e.V. – Verein für feministische Netzpolitik nehmen diese Publikation auch zum Anlass, Forderungen abzuleiten. Denn: Künstliche Intelligenz ist eben kein schnelles Technofix. Mit dem in der Publikation deutlich gewordenen, intersektional-feministischen Anspruch erkennen wir Macht- und Herrschaftsverhältnisse im globalen und lokalen, die unsere Realitäten verschieden bestimmen, an. Daraus leitet sich eine besondere Verantwortung für die Entwicklung und den Umgang mit ebendiesen intelligenten Technologien ab.

Wir werden deshalb zukünftig an anschlussfähigen Handlungsempfehlungen arbeiten, für Politik, Wissenschaft und Zivilgesellschaft.

Wir laden jede\*n dazu ein, an diesem Prozess teilzuhaben und diesen gemeinsam mit uns voranzutreiben.

Künstliche Intelligenz klassifiziert User\*innen noch im binären Geschlechtersystem, trans und non-binary fallen durch diese Raster. Das muss sich ändern. KI sollte darüber hinaus dem Gemeinwohl verpflichtet sein, ebenso wie unseren demokratisch-gesellschaftlichen Grundwerten und marginalisierte Personen(gruppen) nicht noch weiter an den gesellschaftlichen Rand drängen. Zudem braucht es einen breiten gesellschaftlichen Diskurs über Nutzen, Entwicklung und Einsatz von KI; Technologie muss fortwährend kritisch hinterfragt und auf den gesellschaftlichen Nutzen hin überprüft werden.

WARUM IST ES RADIKALE PRAXIS, FEMINISTISCHE  
UTOPIEN ZU SPINNEN UM DEN STATUS QUO NACH  
UND NACH ZU VERABSCHIEDEN?

Das sollte in unseren Feminist Futures deutlich geworden sein, versuchen Sie es doch auch einmal!



# IMPRESSUM

WENN KI, DANN FEMINISTISCH  
IMPULSE AUS WISSENSCHAFT UND AKTIVISMUS

Hrsg. von netzforma\* e.V. – Verein für feministische Netzpolitik, Berlin 2020

## REDAKTION UND LEKTORAT

Annika Kühn, Ann-Kathrin Koster, Christina Dinar, Francesca Schmidt,  
Gitti Hentschel, Hannah Lichtenthäler, Johanna Mellentin,  
Katharina Mosene, Marieke Eilers, Marina Vowinkel, RyLee Hühne,  
Valerie Rhein, Victoria Guijarro, Emma Breuer, Zoe Harrington,  
Friederike Stock

## ÜBERSETZUNGEN

Robyn Travers (Deniz Erden & Ray Acheson)  
Sophie Seifert (Tiara Roxanne)

## FÖRDERUNG

Diese Publikation wurde mit Mitteln der Berliner Landeszentrale für  
politische Bildung realisiert.

Berliner Landeszentrale  
für politische Bildung

**be**  **Berlin**

## ISBN

ISBN 978-3-00-067111-1

## LIZENZ

Diese Publikation steht unter: CC-BY-NC-ND  
Der Artikel „Kontrollverlust und (digitale) Entmündigung –  
Das Gewaltpotential Künstlicher Intelligenz“ von Leena Simon  
steht unter: CC-BY-SA

  
**MÜNDIGKEIT**  
<https://muendigkeit.digital>

—— VERLAG

netzforma\* e.V. – Verein für feministische Netzpolitik  
c/o FFBIZ e.V.  
Eldenaer Str. 35 III  
10247 Berlin

—— VORSTAND

Francesca Schmidt, Christina Dinar, Hannah Lichtenthäler,  
Johanna Mellentin und Katharina Mosene  
Registergericht: Amtsgericht Charlottenburg  
Registernummer: VR 36632 B

—— WEBSEITE

[www.netzforma.org](http://www.netzforma.org)

—— GESTALTUNG

Studio SMS (Katharina Shafiei-Nasab & Alix Stria)

—— SCHRIFTEN

Arthemys Bitmap von Morgane Vantorre  
Armand Grotesk von Charlotte Rohde  
PT Mono von Alexandra Korolkova & Bella Chaeva

—— DRUCK

DRUCKZONE GmbH & Co. KG, Cottbus





Unser besonderer Dank gilt vor allem den Gestalter\*innen  
Alix Stria & Katharina Shafiei-Nasab, auch für ihre  
Geduld, sowie den Schriftgestalter\*innen Morgane Vantorre,  
Charlotte Rohde, Alexandra Korolkova und Bella Chaeva,  
den Übersetzer\*innen Robyn Travers und Sophie Seifert und  
natürlich den Autor\*innen.





Sexrobotik, Drohneneinsätze und Sprachassistent\*innen – all das und Vieles mehr läuft unter den Schlagworten Künstliche Intelligenz (KI) und Algorithmen. Doch was genau umfassen KI und Algorithmen und wie vielfältig sind die Perspektiven auf diese tatsächlich? KI – eine Kehrtwende im Mensch-Maschine Verhältnis? Ein Mythos? Eine Utopie? Die Zukunftsszenarien schwanken zwischen Bedrohung und Verheißung. Klar ist: Technologie im Allgemeinen und algorithmische Prozesse im Speziellen sind nicht ohne Herrschafts- und Machtbezug denkbar. Gerade deshalb gilt es fortwährend, diese Systeme vor dem Hintergrund feministischer Sichtweisen und Wertvorstellungen kritisch zu betrachten, zu bewerten und neu zu entwickeln. Feministische Betrachtungsweisen und intersektionale Zugänge finden kaum bis keine Berücksichtigung, und so reproduzieren sich traditionell patriarchale Sichtweisen und verstärken sich durch den Einsatz von KI. Reicht es aus, KI gerechter zu gestalten? Wenn ja, nach wessen Kriterien? Wie verändert sich Überwachung im Zuge automatisierter Entscheidungsprozesse? Ist die Forderung nach Transparenz die Lösung oder lenkt sie von den notwendigen Forderungen ab? Wer definiert diese sogenannte KI, das Wort zwischen den Gänsefüßchen? Wie kann der Diskurs über KI inklusiv geführt werden? Diese Publikation begleitet aktuelle Entwicklungen intersektional feministisch. Ziel ist es, angesichts sich verdichtender technischer und gesellschaftspolitischer Entwicklungen, neue Perspektiven auf Künstliche Intelligenz und Algorithmen zu entwickeln. Die grundlegende mathematische Formel der Algorithmen muss lauten:

WENN KI, DANN FEMINISTISCH.