

Queere KI als materiell-diskursive Apparate

Ann-Kristin Kühnen

1. Einleitung

Algorithmen und Systeme künstlicher Intelligenz sind in den letzten Jahren immer wieder damit auffällig geworden, ihren Neutralitätsversprechen nicht gerecht zu werden. Vielmehr äußern sie sich als an der Konfiguration gesellschaftlicher (Ungleichheits-)Verhältnisse beteiligt (vgl. Noble 2018; Eubanks 2019; Benjamin 2019; Browne 2015). Dies betrifft auch algorithmische Systeme, die im Gesundheitswesen eingesetzt werden. Nicht zuletzt durch die Corona-Pandemie gibt es einen deutlichen Trend dahin, die Gesundheit einer Bevölkerung mithilfe algorithmischer Systeme zu managen. Potenziale ›ungenutzter‹ Gesundheitsdaten sollen ausgeschöpft, Prozesse effizienter gestaltet und (Krankheits-)Prognosen präziser gestellt werden. Eine Gruppe von Datenwissenschaftler*innen fand jedoch heraus, dass durch den Einsatz eines algorithmischen Systems im US-amerikanischen Gesundheitswesen weiße Patient*innen gegenüber Schwarzen Patient*innen strukturell bevorzugt wurden (vgl. Obermeyer et al.: 2019). Das KI-System wurde eingesetzt, um Risikopatient*innen zu ermitteln und diesen einen Platz in kostenintensiven Pflegeprogrammen zuzuweisen. Dabei hatten weiße Personen eine signifikant höhere Wahrscheinlichkeit, den Pflegeprogrammen zugeordnet zu werden, als Schwarze Personen mit gleichen Vorerkrankungen (vgl. ebd.: 447).

Ein sogenannter *racial bias* wird in datenwissenschaftlichen Forschungen wie der Studie von Ziad Obermeyer et al. häufig als vordergründig technisches Problem verhandelt. Dabei wird *bias* in einem statistischen Sinne als Abweichung von einer Norm oder einem wahren Wert verstanden (vgl. Danks/London 2017: 4692). Datenwissenschaftliche und informationstechnische Studien haben einige Arbeit geleistet, um verschiedene Formen von algorithmischem *bias* herauszustellen (vgl. ebd.; Lee et al. 2019; Panch et al. 2019). Prominent wird als mögliche Quelle eines *bias* immer wieder darauf verwiesen, dass Trainingsdatensätzen, die der Funktionsweise des maschinellen Lernens zugrunde liegen, häufig nur spezifische Bevölkerungsgruppen, bspw. weiße oder männliche Personen, repräsentieren (vgl. Danks/London 2017: 4692). Gemein ist den Auseinandersetzungen dabei der Versuch, einem *bias* durch technische Anpassungen, sogenannten ›fixes‹, entgegen-

zuwirken, indem bspw. Trainingsdatensätze bereinigt oder ergänzt werden. Dabei ist die Annahme vorherrschend, dass es technisch machbar ist, KI-Systeme so zu konfigurieren, dass diesen möglich wird, neutral und objektiv zu agieren.

Dass dieser Neutralitätsanspruch zurückgewiesen werden muss und *bias* vielmehr immer auf spezifische »sozio-kulturelle[...], politisch-ökonomische[...] und strukturell-materielle[...] Bedingungen« (Prietl 2021: 25) verweist, eröffnet sich durch die Perspektive der Feministischen Science-and-Technology-Studies (STS). Mit Donna Haraway (vgl. 1995 [1985]) und anderen Theoretiker*innen der Feministischen STS ausgedrückt, müssen Systeme der künstlichen Intelligenz immer schon als situiert und politisch verstanden werden. Sie agieren nicht entkoppelt von ihrer spezifischen technischen Verfasstheit und gesellschaftlichen Einbettung und sind wie alle Techniken der Wissensproduktion untrennbar mit Machtfragen verknüpft (vgl. Foucault 1978). KI-Systeme mit ihren zugrundeliegenden mathematischen Annahmen, Algorithmen und Daten sind in diesem Verständnis nicht nur technische Instrumente, sondern soziotechnische Gefüge. Sie unterliegen spezifischen Diskursen und Materialisierungsprozessen und sind in historisch gewordenen sozialen und ökonomischen Verhältnissen eingebettet. Dabei erzeugt jede technologische Wissensproduktion unweigerlich Ein- und Ausschlüsse und ist damit beschränkt, kontingenat und partikular (vgl. Haraway 1995a) und zugleich selbst Ausgangspunkt für Grenzverschiebungen und Destabilisierungen (vgl. Weber/Bath 2003: 18). KI-Technologien sind in diesem Verständnis nicht »unschuldig« (Haraway 1995b).¹ Eine durch die Feministischen STS inspirierte Perspektive verweist auf die komplexen Verschränkungen zwischen Sozialem und Technischem und Spannungsfelder zwischen Fluidität und Festsetzung, die insbesondere für eine queertheoretische Betrachtung von KI höchst bedeutsam sind.

Um diese Verschränkungen in den Blick zu nehmen, schlage ich in diesem Beitrag die technofeministische Denkfigur des Apparates vor, die in den Feministischen STS zahlreiche Weiterentwicklungen erfahren hat. Das Verständnis des Apparates als materiell-diskursives Gefüge (vgl. Haraway 1995a; Barad 2007, 2012), in dem sich Technologisches, Politisches, Historisches und Ökonomisches miteinander verschränkt (vgl. Barla 2019), übertrage ich hierfür auf KI-Systeme. Mithilfe der Denkfigur des Apparates soll das in der Studie von Obermeyer et al. angeführte KI-System zur Gesundheitsrisiko-Prognose sowie die Studie selbst queer

¹ Ich verwende in diesem Beitrag die Begriffe Technologie und Technik synonym. Nina Degele stellt zwar heraus, dass ›Technik‹ in der Soziologie dreistrahlig als Artefakt, Formen des Handelns und Formen von Wissen verstanden werden kann und der Begriff der Technologie in seiner ursprünglichen Bedeutung ›die Lehre und das systematisierte Wissen von Technik‹ (2002:20) meint, diese Unterscheidung aber u.a. aufgrund der Verwissenschaftlichung von Technik zunehmend belanglos wird. Zumal wird im angloamerikanischen Raum der Begriff ›technology‹ ebenso im Sinne von Technik und Technologie verwendet.

gelesen werden. Queering verstehe ich dabei als Praktik einer kritischen Analyse, mit der das KI-System auf die Verwobenheit von Macht, Wissen und Materialität befragt werden soll. Dabei wird insbesondere die historische, ökonomische und biopolitische Einbettung des KI-Systems in den Blick genommen. Des Weiteren sollen durch ein Queering mithilfe der Denkfigur des Apparates herrschaftsförmige Dualismen von Technischem und Sozialem, von Materialität und Diskurs durchkreuzt werden. Statt auf Letztbegründungen wie technologische ›fixes‹ zu fokussieren, werden die mannigfaltigen Verschränkungen des KI-Systems zum Ausgangspunkt gemacht, um die Verschiebung und Destabilisierung bestehender Macht- und Ungleichheitsverhältnisse zu denken. Die Denkfigur des Apparates wird hier als »speculative tool« (ebd.: 145) relevant.

Nach einer kurzen Einführung in die theoretischen (Weiter-)Entwicklungen der Denkfigur des Apparates möchte ich mithilfe der Apparat-Figur das KI-System zur Gesundheitsrisiko-Prognose als materiell-diskursives Gefüge in den Blick nehmen. Dabei werde ich herausarbeiten, dass das KI-System nicht losgelöst von der Geschichte des Rassismus im US-Gesundheitswesen betrachtet werden kann und eine »colorblind ideology« (Benjamin 2019: 9) – d.h. eine Form von Rassismus, die sich nicht mehr explizit auf die Kategorie Race bezieht, aber dennoch strukturell wirkmächtig ist – reproduziert. Diese Rassismen setzen sich unter den ökonomischen Einsatzzielen der KI fort. Die ökonomische Ausrichtung auf das Ziel der Kostenreduktion wird dabei vorwiegend zur Gefahr für Schwarze Patient*innen und ihre Körper. Vor diesem Hintergrund argumentiere ich, dass das KI-System als biopolitische Regulierungstechnologie verstanden werden kann. Der Beitrag schließt mit der Frage nach möglichen Verschiebungen und Destabilisierungen in und durch die Apparate, die vermögen Ungleichheiten abzuschwächen.

2. Die Denkfigur des Apparates

Von Haraway als feministisches Instrument eingeführt, um technowissenschaftliche Herstellungspraktiken von Organismen zu erfassen, hat der Begriff des Apparates in den Feminist STS produktive Weiterentwicklungen erfahren. In Anlehnung an die Ausarbeitungen der Literaturwissenschaftlerin Katie King entwickelt Haraway in ihrem viel rezipierten Aufsatz *Situated Knowledge* die Denkfigur »Apparatus of bodily Production« (Haraway 1995a).² Apparate der körperlichen Produktion beschreiben nach Haraway materiell-semiotische Arrangements von

2 Hier ist anzumerken, dass Haraway in Bezug auf die Begriffsbezeichnung zwar explizit auf Katie Kings Arbeiten zurückgreift, maßgebliche Ideen, die mit dem Konzept des Apparates verbunden sind, aber auch auf Haraways Auseinandersetzung mit Schwarzen Denker*innen wie Buchi Emechta oder bell hooks zurückzuführen sind (vgl. Haraway 1991: 109f.; FN 2,5).

(menschlichen und mehr-als-menschlichen) Entitäten und Praktiken, durch die Organismen hervorgebracht werden (vgl. Barla 2019: 103). Dabei ist für Haraway zentral, dass die Wissensobjekte selbst an den konstituierenden Praktiken beteiligt sind. Dieses materiell-semiotische Verständnis, das bei Haraway theoretisch und methodologisch noch ungenau bleibt (vgl. ebd.: 124) entwickelt die queerfeministische Wissenschaftler*in und Physiker*in Karen Barad in ihrem Konzept des Apparates weiter (vgl. Barad 2007; 2012). Barad fokussiert dabei auf die »Prozesse der Materialisation« (Barla 2014: 155) und holt die Denkfigur des Apparates zugleich aus ihrem biowissenschaftlichen Entstehungskontext heraus. Barad versteht Apparate als »materiell-diskursive Praktiken«, als Gefüge, in denen Materialität und Diskurs miteinander intra-agieren und durch diese Intra-Aktionen wiederum materiell-diskursive Phänomene hervorbringen (vgl. Barad 2017: 607). Dabei sind Apparate selbst »grenzziehende Praktiken« (ebd.: 599). Durch spezifische Intra-Aktionen, sogenannte »agentielle Schnitte«, erzeugen Apparate »Bestimmtheit in prinzipiell ontisch unbestimmten ›Phänomenen‹« (Lemke 2017: 566).³ Apparate setzen damit »das in Kraft, was relevant ist und was vom Relevantsein ausgeschlossen wird« (Barad 2017: 602). In Barads Erweiterung von Michel Foucaults Macht- und Diskursverständnis ist Materie ein »konstitutive[r] Bestandteil von Machtbeziehungen« (Lemke 2017: 565). Die Frage, welche Bedeutungen und Körper durch die Apparate inkludiert beziehungsweise geschaffen werden und wie diese in Beziehung gesetzt werden (vgl. Schadler 2017: 180), steht damit auch in engem Zusammenhang mit der Herausbildung von Hierarchien und Ungleichheiten.

In Anknüpfung an Haraways und Barads Arbeiten vollzieht Josef Barla in seiner Dissertation *Techno-Apparatus of Bodily Production* eine weitere Rekonfiguration der Denkfigur des Apparates, die er an zwei »worldly examples« (Barla 2019: 15 zit. Haraway 2000: 46) entfaltet. Am Beispiel der Technologie des Spirometers und genetischer und biometrischer Grenztechnologien (vgl. ebd.: 151ff.), schlägt Barla vor, spezifische Technologien nicht als rein technische Instrumente zu verstehen, sondern immer schon in ihrer materiell-diskursiven Verstrickung von technischen, ökonomischen, historischen, sozialen und politischen Entitäten, Beziehungen und Praktiken zu begreifen (vgl. ebd.: 163).⁴ Technologien als Apparate zu fassen, bedeutet für Barla, diese im Sinne von »Orten«, an denen biologische, technische und politi-

3 Dieser Neologismus, der sich vom Begriff der Interaktion abheben soll, beschreibt einen Prozess, indem die Relata einer Beziehung dieser nicht vorgängig sind, sondern vielmehr durch sie, also durch die Intra-Aktion, entstehen. Barad vollzieht damit eine grundlegende Verschiebung einer Denkbewegung weg von vorgegebenen Entitäten hin zur Frage nach »Praktiken, Tätigkeiten und Handlungen« (Barad 2017: 578).

4 Barla grenzt sich hier von einer Auffassung ab, in der Technologie weder als autonom noch als reine Instrumente oder neutrale Werkzeug verstanden werden, die in unterschiedlicher Weise und zu unterschiedlichen Zwecken eingesetzt werden (vgl. Barla 2019: 71).

sche Kräfte miteinander intra-agieren und dabei ein jeweils spezifisches Phänomen materialisieren« (Barla 2014: 154), zu verstehen.

Apparate zu denken ist dabei, wie es Barla im Anschluss an Barad ausdrückt, »a matter of cutting things together and apart« (Barad 2007: 394). Während »cutting together« auf die Verschränkungen hinweist, fokussiert »cutting apart« die Prozesse der Differenzierung innerhalb eines Phänomens (vgl. Barla 2019: 152). Die Identifikation von Apparaten unterliegt Konstruktions- und Abstraktionsprozessen. Denn wie Barad und im Anschluss Barla betonen, sind die Grenzen der Apparate nicht vorgegeben, sondern vielmehr »open-ended practice« (ebd.: 157) und bestimmte Apparate immer schon in Intra-Aktion mit anderen Apparaten (vgl. Barad 2007: 203). Vor diesem Verständnis ist es niemals möglich, die Grenzen des Apparates im Vorphein zu bestimmen, da ihnen »homogene Letztelemente« fehlen, »die eine klare Abgrenzung zwischen ‚innen‘ und ‚außen‘ zulassen würden« (Barla 2014: 11). Eine Auflistung aller Ursachen und Kräfte, die in einem Apparat wirken, oder nicht wirken, ist nicht möglich (vgl. Barla 2019: 147). Barla betont jedoch, dass dies nicht zu einer Willkürlichkeit der Grenzen des Apparates führt und diesbezüglich alle konstituierenden Intra-aktionen die gleichen Bedeutungen haben (vgl. ebd.: 157). Vor diesem Hintergrund bedarf eine Betrachtung immer einer Festlegung »was von Bedeutung ist, wie es von Bedeutung ist und für wen es von Bedeutung ist« (Barla 2014: 11).

Für Barla stellt der »Techno-Apparatus of Bodily Production« neben einer Denkfigur (»figure«) auch ein spekulatives Werkzeug (»speculative tool«) dar (vgl. Barla 2019: 145). Barla nutzt die Figur des Apparates, um Erzählungen von Technologien und Körpern zu ermöglichen, die über die Narrative von Verlust und Beherrschung und damit technikdeterministische Perspektiven hinausgehen, ohne jedoch die konkreten Konsequenzen spezifischer Verschränkungen und Materialisierungsprozesse zu ignorieren (vgl. ebd.: 152). Die Frage von Macht und Werden, die die Apparate durchdringen, ist für Barla zentral (vgl. Barla 2015: 11). Diese Ausführungen Barlas lassen sich mit Anregungen von Cornelia Schadler erweitern. Nach Schadler kann mithilfe des Apparates eine kritische Analyse erfolgen, die auf die Identifikation von Apparaten, »die spezifische Ungleichheiten und transversale Körper inkludieren« (Schadler 2017: 180) oder negieren, abzielt. Der Apparat-Begriff stellt zugleich aber auch ein Werkzeug dar, um die Gefüge von menschlichen und nicht-menschlichen Entitäten herauszuarbeiten oder zu entwickeln, »die diese Ungleichheiten abschwächen« (ebd.: 181) oder gar transformieren. Dabei ist der Blick nicht allein auf Menschen als Handlungsträger*innen ausgerichtet, sondern fragt auch nach anderen diskursiven und materiellen Aktant*innen (vgl. ebd.: 182).

3. KI-Systeme als Apparate. Überlegungen am Beispiel algorithmischer Gesundheitsrisiko-Prognosen

In diesem Beitrag möchte ich den Weiterentwicklungen der Denkfigur des Apparates durch Haraway (vgl. 1995), Barad (vgl. 2007; 2012), Barla (vgl. 2014, 2019) und Schadler (vgl. 2017) folgen, um KI-Systeme queertheoretisch zu analysieren. KI-Systeme als Apparate zu denken, bedeutet diese als soziotechnische und materiell-diskursive Gefüge in den Blick zu nehmen. KI-Systeme werden in diesem Sinne als »Orte« begriffen, an denen sich Historisches, Politisches, Ökonomisches und Technologisches miteinander verschränkt und Rekonfigurationen von Wissen und materiellen Körpern auftreten, die wiederum weitreichende soziale, ökonomische und politische Konsequenzen haben. Verdeutlichen möchte ich diese Herangehensweise am KI-System zur Gesundheitsrisiko-Prognose, das in der Studie von Obermeyer et al. verhandelt wird. Wie eingangs schon kurz angeführt, arbeitet dieses KI-System mit Algorithmen, die im US-amerikanischen Gesundheitswesen landesweit eingesetzt werden, um Risikopatient*innen für sogenannte »high-risk-care-management«-Programme ausfindig zu machen (vgl. Obermeyer et al. 2019: 447). Hierfür weist das KI-System Patient*innen Risiko-Scores zu. Obermeyer et al. stellen in ihrer datenwissenschaftlichen Analyse heraus, dass Schwarze Patient*innen deutlich seltener für die Pflegeprogramme vorgeschlagen wurden, als weiße Personen mit den gleichen Vorerkrankungen und das KI-System damit einen *racial bias* erzeugt.⁵

In Anknüpfung an Barla schlage ich vor, das KI-System nicht als rein technisches Instrument zu fassen, dessen primäres Ziel es ist, Risiko-Scores von Patient*innen zu errechnen, sondern als soziotechnischen Apparat, an dem sich Macht, Wissen und Materialität verschränken. Wie in den Ausführungen zur Denkfigur deutlich wurde, sind die Grenzen von Apparaten nicht vorbestimmt, sondern konstituieren sich in einem stetigen Prozess und in Abhängigkeit zu spezifischen Methoden und Perspektiven. Mit den ausgewählten Theorien, Studien und Materialien, die ich zur Analyse des KI-Systems verwende, kann ich Materialisierung- und Differenzierungsprozesse in den Blick nehmen. Diese Herangehensweise vermag auf verschiedene Aspekte des Apparates zu verweisen und Intra-Aktionen in komplexen soziotechnischen Gefügen zu identifizieren, bleibt dabei aber immer partikular und entzieht sich der Suche nach Letztbegründungen. Mithilfe der Denkfigur des Apparates soll das KI-System und die dazugehörige Studie in diesem Sinne queer gelesen werden. Im Zentrum meiner Analyse des KI-Systems steht die Frage nach den materiell-diskursiven Verschränkungen und »grenzziehende Praktiken« (Barad 2017).

5 Der Fokus auf die datenwissenschaftliche Studie von Obermeyer et al. stellt die (Re-)Produktion von Rassismen heraus. Wie weitere Arbeiten zu KI-Systemen im Gesundheitswesen aufzeigen, kann ein solcher *bias* nicht nur rassistische Personen, sondern auch neurodivergente, trans, inter* und queere Personen betreffen (vgl. Keyes 2020, Keyes et al. 2020).

599), die den *racial bias* hervorbringen. Im Folgenden werde ich ausgehend von Obermeyers et al. Studie die technische Verfasstheit des KI-Systems in den Blick nehmen und daraufhin Intra-Aktionen mit historischen, ökonomischen und biopolitischen Aspekten herausarbeiten.

3.1. Bias als technisches Problem?

Wie kann der *bias* im KI-System aus einer datenwissenschaftlichen Perspektive verstanden werden? Wie werden die Gesundheitsrisiko-Prognose und die Pflegeplatz-Vergabe technisch umgesetzt? Mit diesen Fragen setzen sich Obermeyer et al. in ihrer Studie intensiv auseinander und nehmen vor diesem Hintergrund Datensets, Algorithmen, zugrundeliegende Modellannahmen und Trainingsprozesse in den Blick (vgl. Obermeyer 2019: 448f.). Wie Obermeyer et al. herausstellen, gelten Programme wie das »high-risk-care-mangement«-Programm in den USA allgemein als wirksam, um den Gesundheitsstatus von Patient*innen zu verbessern (vgl. ebd.: 477). Da diese Programme sehr kostenintensiv sind, werden für die Vergabe algorithmische Systeme eingesetzt, die eine präzise und effektive Platzzuweisung versprechen (vgl. ebd.). Hierzu legten die Entwickler*innen in ihren Modell-Grundannahmen fest, dass diejenigen Patienten*innen von einem Platz in den Pflegeprogrammen am meisten profitieren könnten, die schon in der Vergangenheit einen hohen Pflegebedarf hatten (vgl. ebd.). Um diese Annahme im technischen System umzusetzen, gingen in die Berechnungen Daten »über die gesamten medizinischen Ausgaben eines Jahres sowie feingranulare Daten zur Inanspruchnahme von Gesundheitsleistungen des Vorjahres« (Orwat 2019: 53) ein. Obermeyer et al. stellen vor diesem Hintergrund fest, dass die algorithmische Vorhersage des zukünftigen Pflegebedarfs eigentlich eine Vorhersage zukünftiger Gesundheitskosten ist (vgl. ebd.). Der Risiko-Score gibt keinen direkten Aufschluss über den zu erwartenden Gesundheitszustand, sondern prognostiziert vielmehr die Kosten zukünftiger Behandlungen. Dieses Ergebnis der Datenwissenschaftler*innen eröffnet wichtige Einblicke in das KI-System als Apparat. So wird der Pflegebedarf nicht auf der Basis von Daten zu körperlichen Vorerkrankungen, sondern anhand vergangener Gesundheitskosten errechnet. Auf den ersten Blick scheint dies keine großen Auswirkungen zu haben, da im Durchschnitt die Gesundheitskosten und der Pflegebedarf von Patient*innen miteinander korrelieren, jedoch geht vor allem bei rassifizierten Personen in den USA der Bedarf von Gesundheitsversorgung und die eigentliche geleistete Versorgung dieser Personen oft erheblich auseinander (vgl. Obermeyer 2019: 450). So kommt es dazu, dass ab einem bestimmten Gesundheitszustand (gemessen an der Anzahl chronischer Krankheiten) Schwarze Personen durchschnittlich weniger Kosten für das Gesundheitssystem erzeugen als weiße Personen mit dem gleichen Gesundheitszustand (vgl. ebd.). In der Praxis bedeutet dies, dass bei einer Schwarzen und bei einer wei-

ßen Patientin, für die das algorithmische System den gleichen Risiko-Score ausgibt, die Wahrscheinlichkeit signifikant erhöht ist, dass die Schwarze Patientin einen deutlich schlechteren Gesundheitszustand hat als die weiße Patientin (vgl. Benjamin 2019b: 421). In der Konsequenz heißt das, dass Schwarze Personen deutlich seltener für die Pflegeprogramme vorgeschlagen werden.

Das KI-System tritt hier als ›Ort‹ zum Vorschein, in dem die Produktion von Wissen auf Grundlage statistischer Modelle und Grundannahmen spezifische Ausschlüsse erzeugt – in diesem Fall einen *racial bias*. In diesem Zuge stellt sich die Frage, wie die unterschiedlich hohen Gesundheitskosten bei Schwarzen und weißen Personen zu erklären sind. Schon in der Studie von Obermeyer et al. deuten sich Anknüpfungspunkte für diese Fragen an. Obermeyer et al. erwähnen zwei Erklärungssätze für die unterschiedliche Höhe der Gesundheitskosten bei weißen und Schwarzen Patient*innen. Erstens stellen sie heraus, dass die Kategorie Race mit Armut korreliert. So waren laut einer Studie des U.S. Zensus Büro im Jahr 2020 in den USA 19,5 Prozent der Schwarzen Bevölkerung von Armut betroffen, dagegen lag die Armutsquote in der Gesamtbevölkerung bei 11,4 Prozent (vgl. Shrider et al. 2020: 15). Personen, die in Armut leben, sind im Gesundheitswesen zahlreichen Hürden und diskriminierenden Mechanismen ausgesetzt (vgl. Obermeyer et al. 2019: 450). Diese Mechanismen greifen häufig selbst dann, wenn eine Krankenversicherung vorliegt (vgl. ebd.), bspw. durch schlechten Zugang zu Verkehrsmitteln oder Arbeits- oder Pflegeverpflichtungen, die es erschweren, Krankenpflege wahrzunehmen, aber auch fehlende Informationen und Behandlungswissen im Allgemeinen (vgl. ebd.). Zweitens führen Obermeyer et al. die unterschiedliche Höhe an Gesundheitskosten darauf zurück, dass Schwarze Patient*innen im Gesundheitssystem direkte Diskriminierung erfahren. So wird in Studien dokumentiert, dass Ärzt*innen oft abweichende Annahmen über Schwarze Patient*innen in Bezug auf Intelligenz, Zugehörigkeit oder auch Schmerztoleranz haben und dass Schwarze Patient*innen dem US-Gesundheitssystem durchschnittlich weniger vertrauen (vgl. ebd.).

Auch wenn Obermeyer et al. diese Studien erwähnen, werden die soziotechnischen Verschränkungen, die sich in diesen Erklärungsansätzen andeuten, durch die Datenwissenschaftler*innen nicht weiterverfolgt. Den Datenwissenschaftler*innen ist im Sinne einer anwendungsbezogenen Analyse daran gelegen Lösungsvorschläge anzubieten, wie der *bias* mithilfe technischer ›fixes‹ behoben werden kann. Als Hauptproblem identifizieren sie vor diesem Hintergrund das Label der Daten. Während die Modell-Infrastrukturen nach Obermeyer et al. beibehalten werden können, müssten die Label der Daten grundlegend verändert werden (vgl. ebd.). Dieses solutionistische Vorgehen entspricht, wie Bianca Prietl (vgl. 2019) in Bezug auf algorithmische Entscheidungssysteme herausgearbeitet hat, einer typischen Herangehensweise datenwissenschaftlicher Analysen. Diese eröffnen wichtige informationstechnische Aufschlüsse, die zwar auf die gesellschaftliche Einbettung algorithmischer Systeme verweisen, selbige aber nicht präzise genug

kontextualisieren. Stattdessen verharren datenwissenschaftliche Analysen, wie es die Rassismus- und STS-Forscherin Ruha Benjamin ausdrückt, auf der Ebene passiver und steriler Beschreibungen (vgl. Benjamin 2019b: 422). Bezugnehmend auf Obermeyers et al. Studie verdeutlicht Benjamin, dass das US-amerikanische Gesundheitswesen, in welchem das algorithmische System eingesetzt wurde, als ein gesellschaftlicher Teilbereich betrachtet werden muss, in dem historisch gewordene, systematische und strukturelle Rassismen wirkungsmächtig werden. Mit Benjamins Perspektive wird das KI-System als Apparat fassbar, indem sich technologische Aspekte mit historisch gewordenen Strukturen verschränken. Im Folgenden möchte ich Benjamins Überlegungen nachgehen und damit Intra-Aktionen innerhalb des Apparates freilegen, die auf komplexe Verschränkungen von Wissen, Macht und Materialität innerhalb des KI-Systems verweisen.

3.2. Rassistische Kontinuitäten im US-amerikanischen Gesundheitswesen

Bis ins Jahr 1964 galten in den USA die Jim-Crow-Gesetze, die die Rassentrennung in den USA gesetzlich absicherten. Das Gesundheitswesen war eines von vielen Bereichen, in denen Schwarze Personen aufgrund der Jim-Crow-Gesetze einen ungleichen Zugang zu lebenswichtigen Ressourcen hatten, unterdrückt und ausgebeutet wurden. Dies zeigte sich bspw. anhand segregierter Krankenhäuser, in denen mancherorts ausschließlich weiße Personen behandelt wurden, aber auch an rassistischen Medizin-Lehrplänen und ungleichen Versicherungsstrukturen (vgl. Benjamin 2019b: 422). Um die Mechanismen hinter den Jim-Crow-Gesetzen und deren Kontinuitäten bis in die Gegenwart genauer zu verstehen, ist es notwendig das Gefüge, das sich hinter der Bezeichnung »Jim Crow« zeigt, näher in den Blick zu nehmen. Der Begriff »Jim Crow« nimmt seinen Ursprung in der Bezeichnung eines Charakters einer Minstrel-Show von 1832.⁶ Mit diesem Ausgangspunkt wurde die Bezeichnung »Jim Crow« mit der Zeit gesellschaftlich immer populärer und wurde ab den 1950er-Jahren letztendlich auch als akademische Kurzbezeichnung für die legalisierte Rassentrennung und Unterdrückung im Süden der USA benutzt (vgl. Benjamin 2019a: 9). »Jim Crow« erwies sich in diesem Bezug als elastischer Begriff, der zugleich eine Ära, eine geografische Region, Gesetze, Institutionen, Zölle sowie einen Verhaltenskodex, der die weiße Vormachtstellung aufrechterhielt, beschrieb (vgl. ebd.). Für Benjamin stellte »Jim Crow« eine Intersektion von »legal codes, social codes, and building codes« (ebd.: 91) dar. Diese Codes ermöglichen es gemeinsam, die Bevölkerung voneinander getrennt und in Ungleichheit zu halten, und ihre

6 In sogenannten Minstrel-Shows stellten weiße Schauspieler zu Unterhaltung eines weißen Publikums Schwarze Charaktere dar. Dies geschah im Rückgriff auf gängige Stereotype und in herabwürdigender und verspottender Art und Weise.

Wirkung setzt sich bis heute fort, wie es einige Theoretiker*innen, darunter Michel-Alexander (vgl. 2012), herausstellen. Alexander analysiert in ihrem Buch *The New Jim Crow* anhand des Gefängnissystems in den USA die gegenwärtige Kontinuität des Jim-Crow-Gefüges. Dabei betont Alexander eine entscheidende Verschiebung in der Wirkungsweise des ›New Jim Crow‹ »from explicit racialization to a color-blind ideology« (ebd.: 9). Race ist heutzutage keine legitime Kategorie mehr für die Ungleichbehandlung von Menschen. In diesem Zuge wird die anhaltende Kontinuität von Rassismus durch die Annahme verdeckt, dass alle Menschen gleich seien. In Anknüpfung an Alexanders Arbeiten etabliert Ruha Benjamin den Begriff ›New Jim Code‹: Darunter versteht Benjamin den Einsatz neuer Technologien, die bestehende Ungleichheiten widerspiegeln und reproduzieren, dabei aber als objektiver und progressiver als Systeme der Unterscheidung einer früheren Ära angepriesen und wahrgenommen werden (vgl. Benjamin 2019a: 5f.).

Benjamin nimmt in den Blick, was passiert, wenn verdeckte Rassismen in Form einer »colorblind ideology« (ebd.: 9) in technische Codes überführt werden. Dabei vertritt sie eine Perspektive, in der Rassismen nicht nur ein Symptom oder die Folge des Einsatzes dieser Technologien sind, sondern mitunter eine Voraussetzung für ihre Herstellung (vgl. ebd.: 44). Benjamin versteht hierbei Rassismus in Verbindung mit anderen Formen der Unterdrückung nicht nur als Ideologie oder Geschichte, sondern »as a set of technologies« (ebd.), die die soziale Struktur maßgeblich mit verändern und spezifische Muster sozialer Beziehungen erzeugen. Rassismus wird damit strukturell und zugleich »Black-boxed as natural, inevitable, automatic« (ebd.: 45, H. i. O.). Benjamin folgend kann nur aus diesem Blickwinkel der Beständigkeit und Anpassungsfähigkeit von Rassismus auf den Grund gegangen werden (vgl. ebd.: 91). Im Zentrum steht dann die Frage, wie algorithmische Systeme in der Aufrechterhaltung, aber auch an der Transformation dieser Rassismen beteiligt sind (vgl. ebd.: 76).

Mit der Denkfigur des Apparates gedacht, ist das KI-System der ›Ort‹, an dem sich Historisches und Technologisches miteinander verschränkt und dabei Rassismen reproduziert werden, die wiederum unter dem Deckmantel einer »colorblind ideology« (ebd.: 9) und den Vorzeichen der Neutralität verschleiert werden. Das KI-System als Apparat setzt damit »in Kraft, was relevant ist und was vom Relevantsein ausgeschlossen wird« (Barad 2017: 602). In dieser Betrachtung verschiebt sich der Fokus der Analyse vom *racial bias* zu systematischen und strukturellen Rassismen. Wie ich in den nächsten beiden Abschnitten zeigen werde, ist Rassismus als »set of technologies« (Benjamin 2019a: 45) immer schon mit anderen Strukturen – insbesondere ökonomischen – verzahnt. Um diese zu verdeutlichen, werde ich ökonomische Aspekte als Teile des Apparates beleuchten und anschließend die Verschränkungen von Technologien, Rassismus und ökonomischen Logiken mit dem Begriff der Biopolitik in Verbindung bringen.

3.3. Ökonomische Einsatzziele

Das KI-System zur Gesundheitsrisiko-Prognose wird als Technologie verhandelt, die die Identifikation von Risikopatient*innen und eine bedürfnisorientierte Verteilung von Pflegeressourcen im Gesundheitswesen ermöglichen soll. Obermeyer et al. stellen im Rahmen ihrer Analyse jedoch heraus, dass Kostenreduktion ein weiteres wichtiges Ziel des Einsatzes darstellt. Das KI-System wird als eine Investition verstanden, mit der die zukünftigen Kosten für das Gesundheitswesen verringert werden können – ein Ansatz, der branchenweit als wünschenswert gilt (vgl. Obermeyer et al. 2019: 451). Hierin spiegelt sich die Ökonomisierung und Neoliberalisierung des Gesundheitswesens wider. Nicht nur in den USA übt die Tendenz der Privatisierung verstärkt ökonomischen Druck auf das Gesundheitswesen aus (vgl. Mathar 2010: 47ff.). Da das Gesundheitswesen nun wirtschaftliches Wachstum erreichen soll, wird die gesundheitliche Patient*innenversorgung immer mehr zu einem Mittel für privatwirtschaftliche Gewinne (vgl. Bauer 2006: 18). Damit geht eine Budgetorientierung einher, die dazu führt, dass medizinische Leistungen rationiert und überwiegend »sozial selektiv erbracht werden« (ebd.: 20). Gesundheit wird vor diesem Hintergrund zur Ware und das Postulat der Kostensenkung und Profitmaximierung übertrumpft in vielen Fällen die gesundheitlichen Bedürfnisse der Patient*innen (vgl. Benjamin 2019a: 30). Wie sich am Beispiel des KI-Systems zeigt, herrscht jedoch sehr häufig die Annahme vor, dass Kostenreduktion und medizinische Ziele wie die bedürfnisorientierte Behandlung von Risikopatient*innen miteinander zu vereinen seien. Mithilfe der Denkfigur des Apparats offenbart sich jedoch, dass der Fokus auf Kostenreduktion als Teil einer ökonomischen Logik zu einer Wissensproduktion führt, die vorrangig rassifizierte Patient*innen an lebensbedrohliche Ränder drängt.

Auch auf der direkten Ebene der Technologien lassen sich problematische Effekte dieser Ökonomisierung erkennen. So stellen Obermeyer et al. und Benjamin heraus, dass durch den urheberrechtlichen Schutz der Algorithmen in den meisten Fällen verunmöglicht wird, die Funktionsweise der algorithmischen Systeme nachzuvollziehen (vgl. Obermeyer et al. 2019: 447; Benjamin 2019b: 421). Dies erschwert kritische Interventionen und Forschungen, um Gründe von *bias* zu identifizieren. Urheberrechte und der Schutz von geistigem Eigentum im Allgemeinen sind dabei nicht als Konsequenz, sondern als eng verzahnt mit den historisch gewordenen Bedingungen einer kapitalistischen Wirtschaftsweise zu verstehen (vgl. Schmidt 2010: 6). Profit wird auch im gegenwärtigen Kapitalismus maßgeblich durch die Bedingungen von Eigentum erzielt.

Ökonomische Strukturen und Logiken äußern sich hier als beteiligt an der Herstellung und Rekonfiguration von Wissen und Bedeutungen durch das KI-System. Sie erzeugen in Intra-Aktion mit den algorithmischen Systemen nicht nur

rassistische Ausschlüsse, sondern bringen als Teil bestimmter Apparate auch spezifische Urheberrechte hervor, die wiederum kritische Analysen und Interventionen einschränken.

3.4. Das KI-System als biopolitische Regulierungstechnologie

Während der Blick auf die Intra-Aktionen von Ökonomischem und Technischem schon einige wichtige Verweise auf machtvolle Ausschlüsse gibt, ist es vor allem der Begriff der Biopolitik, der auf die Intra-Aktion von Ökonomischem, Technologischem und Rassismen verweist. Mit der Entstehung der Biomacht wird laut Foucault, das »alte Recht [des Souveräns], sterben zu *machen* oder leben zu *lassen*, abgelöst von einer Macht, leben zu *machen* oder in den Tod zu *stoßen*« (ebd.: 165; H.i.O.). In einer seiner Vorlesungen im Jahr 1976 fragt Foucault in diesem Bezug: »Wie kann eine Macht töten, wenn es stimmt, daß es im Wesentlichen darum geht, Leben aufzuwerten, seine Dauer zu verlängern, seine Möglichkeiten zu vervielfachen, Unfälle fernzuhalten oder seine Mängel zu kompensieren?« (Foucault 2014 [1976]: 104). Als Antwort bringt Foucault Rassismus in den Fokus. Die Rolle des Rassismus liegt nach Foucault darin, die Trennlinien zwischen dem, was leben soll und dem, was sterben muss, zu erzeugen (vgl. Folkers/Lemke 2014: 13). Dabei wirkt der moderne Rassismus in vielen Fällen nicht im Sinne einer physischen Vernichtung, sondern ist vielmehr verschleiert in den sozialen und politischen Formen »jemand[en] der Gefahr des Todes aus[zu]liefern, für bestimmte Leute das Todesrisiko [zu] erhöhen oder einfach den politischen Tod, die Vertreibung, Zurückweisung« (Foucault 2014: 106).⁷ Mit dem Begriff der Biopolitik eröffnen sich neue Perspektiven auf das KI-System als Apparat. Das KI-System stellt selbst eine Technologie dar, die eingesetzt wird, um Risikopatient*innen präventiv zu behandeln, sie zu umsorgen, ihre Leben zu verlängern – kurz, um eine Technologie mit dem Ziel »leben zu *machen*« (Foucault 1995: 165; H.i.O.). Als biopolitisches Instrument an dem Ökonomisches, Technologisches und Historisches intra-agiert, hat das KI-System aber auch die Konsequenz »in den Tod zu *stoßen*« (Foucault 1995: 165; H.i.O.). Statistische Instrumente wie das algorithmische System werden eingesetzt, um darüber zu entscheiden, welche ökonomischen Ressourcen wem zugestanden werden sollen. Sie sind biopolitische Regulierungstechnologien. Dabei sind es strukturelle durch eine »colorblind ideology«

⁷ Mit seinem Konzept der Nekropolitik setzt Achille Mbembe an dieser Stelle an und verdeutlicht damit den Fortbestand der Macht des Souveräns in der Moderne, den er bei Foucault als zu kurz gedacht kritisiert. Mbembe nimmt die »zeitgenössische Unterwerfung des Lebens unter die Macht des Todes« (Mbembe 2011: 89) in den Blick und stellt dabei heraus, dass gerade in der Moderne große Bevölkerungsgruppen spezifischen Lebensbedingungen unterworfen werden, die sie in den »Status lebendiger Toten versetzen« (vgl. ebd.). Als Beispiele führt er u.a. Konzentrationslager, Plantagen und Apartheid-Systeme an.

(Benjamin 2019a: 9) verschlüsselte Rassismen, die die Grenze zwischen Leben und Sterben erzeugen.

Das KI-System als Apparat produziert nicht nur bestimmtes Wissen in Form von Vorhersagen über das Gesundheitsrisiko, sondern bringt performativ spezifische Körper hervor. Die Verteilung von Gesundheits-Ressourcen geht mit der Konsequenz einher, dass einige Körper intensiver, andere weniger intensiv behandelt werden. Vor dem Hintergrund der biopolitischen Einbettung kann zugespitzt formuliert werden, dass das algorithmische System Körper hervorbringt, die leben können, und Körper, die sterben gelassen werden. Das KI-System als Apparat zu denken, ermöglicht jedoch nicht nur die materiell-diskursiven Praktiken zu beleuchten, durch die Race in algorithmische Systeme eingebettet wird, sondern auch jene Praktiken in den Blick zu nehmen, durch die rassifizierte Körper als Abweichung einer weißen Norm auftreten. Das KI-System als Apparat ist an der Markierung Schwarzer Körper beteiligt. Es ist grenzziehende Praktik, die Differenzen hervorbringt. Doch Körper werden durch Apparate nicht nur rekonfiguriert, sie sind in Form körperlicher Materialität mit an der Konstitution von Apparaten beteiligt. Durch das algorithmische System wird körperliche Materialität – in Form von Erkrankungen, Schmerzen, Ausflüssen und Emotionen – durch die Operationalisierungen in Gesundheitskosten als Indikator des Gesundheitsrisikos fast vollständig rationalisiert. Anstelle von körperlicher Materialität sollen allein die Gesundheitskosten für den Gesundheitszustand der Patient*innen sprechen. Doch in diesen Gesundheitskosten finden sich die Spuren der Rassifizierung von Körpern. So hat Rassifizierung einen entscheidenden Einfluss darauf, wer Behandlungen rechtzeitig und häufig genug wahrnehmen kann, beziehungsweise eine dem Gesundheitszustand angemessene Behandlung überhaupt erst erhält. Die Kategorie Race tritt hier als »temporäre [...] Materialisierung [...] in Körpern« (Namberger 2014: 143) auf, die sich wiederum in den Daten zu den Gesundheitskosten materialisieren. Diese Materialisierungsprozesse werden jedoch durch die Annahme von Gesundheitskosten als neutraler und objektiver Indikator vollständig ausgeblendet und Schwarze Körper damit im Zeichen einer »colorblind ideology« (Benjamin 2019a: 9) unsichtbar gemacht.

4. Ausblick: Alternative Apparate und queere KI?

Am Beispiel eines KI-Systems, das zur Gesundheitsrisiko-Prognose eingesetzt wurde, habe ich verdeutlicht, wie die Denkfigur des Apparates genutzt werden kann, um die Analyse von KI-Systemen zu queeren. Statt auf Letztbegründungen zu fokussieren, eröffnet die Denkfigur des Apparates die Möglichkeit, KI-Systeme als Gefüge zu verstehen und deren mannigfaltigen, materiell-diskursiven Verschränkungen in den Blick zu nehmen. Eine solche multiperspektivische Herangehensweise ist maß-

geblich, um die Reproduktion von Macht- und Herrschaftsverhältnissen mit und durch KI-Systeme analysieren zu können. Der Fokus auf die Intra-Aktionen innerhalb des Apparates eröffnet Einblicke in das Wechselspiel von Fluidität und Festsetzung. Die Wirkmächtigkeit aller materiellen und diskursiven Aktant*innen wird nicht a priori festgelegt, sondern als in und durch die Apparate entstehend konzipiert. Dies gilt auch für Ungleichheitskategorien und Normen, die erst in den Apparaten wirkmächtig werden. KI-Systeme als Apparate (re-)produzieren machtvolle Ein- und Ausschlüsse entlang bestehender Ungleichheitskategorien, erzeugen zugleich aber auch immer Überschüsse und Uneindeutigkeit. Andere Verschränkungen werden dadurch denkbar. Es ist mitunter dieser Aspekt, der die Denkfigur des Apparates als »speculative tool« (Barla 2019: 145) für das Queering von KI relevant macht. Mit der Apparat-Figur eröffnet sich der Raum für Überlegungen, wie Verschiebungen in den Apparaten aussehen könnten, die Ungleichheit abschwächen bzw. Macht- und Herrschaftsverhältnisse zu destabilisieren vermögen. Dabei können nicht nur menschliche Aktant*innen, sondern auch mehr-als-menschliche Aktant*innen in den Blick genommen werden.

Wie würde sich der Apparat um das KI-System zur Gesundheitsrisiko-Prognose verändern, wenn dieses in einem anderen, historisch gewordenen, gesellschaftlichen Kontext eingesetzt würde, bspw. im deutschen Gesundheitssystem? Wie würden sich die Apparate verändern, wenn eine Offenlegung der technischen Verfasstheit von KI-Systemen gesetzlich geregelt wäre oder die KI-Systeme im Sinne von Open-Source-Anwendungen der Öffentlichkeit frei zugänglich wären? Welchen Effekt auf das Gefüge von Macht, Wissen und Materialität hätte es, wenn die Prognose des Pflegebedarfs anstelle der vergangenen Gesundheitskosten auf der Basis von Daten, die auf körperliche Vorerkrankungen der Patient*innen verweisen, hochgerechnet würde?

Diese Überlegungen lassen auf maßgebliche Veränderungen der Apparate schließen, die die KI-Systeme jedoch nicht von ihrer herrschaftsförmigen Einbettung lösen könnten. Umso interessanter ist es eine Destabilisierung der Ungleichheitsverhältnisse in und durch Apparate mit Wendy Chun weiterzudenken. Chun schlägt in ihrem Buch *Discriminating Data* vor, diskriminierende Vorhersagen in KI-Systeme als Beweis vergangener Diskriminierungen und Möglichkeit zur Feststellung gegenwärtiger Ungleichheiten zu verstehen (vgl. 2021: 2). Diese Herangehensweise entlehnt Chun der Verhandlung und Verwendung von Klimamodellen, die die Entwicklung des Klimas berechnen. Diese Modelle verweisen auf die zukünftigen Konsequenzen vergangener und gegenwärtiger Emissionen. In diesem Sinne werden sie zum Ausgangspunkt von politischen Handlungen und sozialökologischem Wandel (vgl. ebd.: 254). Auf KI-Systeme als Apparate übertragen, entsteht durch diese diskursive Verschiebung die Möglichkeit *bias* nicht als unlösbares Problem, sondern als Beweis für historisch gewordene Ungleichheiten, die mit einer kapitalistischen Wirtschaftslogik verzahnt sind, zu verhandeln und sie in

diesem Sinne zu politisieren. Diese Perspektive fasst *bias* als »glitch« (Russell: 2021) – also als eine Art Störung –, die dann zum Ausgangspunkt politischer Intervention wird. KI-Systeme werden als Möglichkeit verstanden, Mechanismen, die Ungleichheitsverhältnisse verursachen, sichtbar zu machen. KI-Systeme werden als politisch begriffen und die Vermutung, sie seien unschuldig, wird zurückgewiesen.

Literaturverzeichnis

- Alexander, Michelle. 2012. *The new Jim Crow. Mass incarceration in the age of colorblindness*. New York: New Press.
- Barad, Karen. 2007. *Meeting the universe halfway. Quantum physics and the entanglement of matter and meaning*. Durham/London: Duke University Press.
- Barad, Karen. 2012. *Agentieller Realismus: Über die Bedeutung materiell-diskursiver Praktiken*. Berlin: Suhrkamp.
- Barad, Karen. 2017. Agentieller Realismus. In *Science and Technology Studies. Klassische Positionen und aktuelle Perspektiven*, Hg. Susanne Bauer, Torsten Heinemann und Thomas Lemke, 574–643. Berlin: Suhrkamp.
- Barla, Josef. 2014. Apparate Körperlicher Produktion und die Materialität der Figuration. In *Critical matter. Diskussionen eines neuen Materialismus*, Hg. Tobias Goll, Daniel Keil und Thomas Telios, 148–164. Münster: Edition Assemblage.
- Barla, Josef. 2019. *The Techno-Apparatus of Bodily Production. A New Materialist Theory of Technology and the Body*. Bielefeld: transcript.
- Bauer, Ulrich. 2006. Die sozialen Kosten der Ökonomisierung von Gesundheit. *Aus Politik und Zeitgeschichte* 8–9: 17–24.
- Benjamin, Ruha. 2019a. *Race after Technology. Abolitionist Tools for the New Jim Code*. Cambridge: Polity Press.
- Benjamin, Ruha. 2019b. Assessing risk, automating racism. *Science* 366: 421–422.
- Browne, Simone. 2015. Dark matters. On the surveillance of blackness. Durham: Duke University Press.
- Chun, Wendy Hui Kyong. 2021. *Discriminating data. Correlation, neighborhoods, and the new politics of recognition*. Cambridge: The MIT Press.
- Danks, David und Alex London. 2017. Algorithmic Bias in Autonomous Systems. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*: 4691–4697.
- Degele, Nina. 2002. *Einführung in die Techniksoziologie*. München: Wilhelm Fink.
- Eubanks, Virginia. 2019. *Automating inequality. How high-tech tools profile, police, and punish the poor*. New York: Picador St. Martin's Press.
- Folkers, Andreas und Thomas Lemke (Hg.). 2014. *Biopolitik. Ein Reader*. Berlin: Suhrkamp.

- Foucault, Michel. 1978. *Dispositive der Macht. Über Sexualität, Wissen und Wahrheit*. Berlin: Merve.
- Foucault, Michel. 1995. *Der Wille zum Wissen. Sexualität und Wahrheit 1*. Berlin: Suhrkamp.
- Foucault, Michel. 2014 [1976]. In *Verteidigung der Gesellschaft. Biopolitik. Ein Reader*, Hg. Andreas Folkers und Thomas Lemke, 88–114. Berlin: Suhrkamp.
- Haraway, Donna. 1991. *Simians, Cyborgs, and Women. The Reinvention of Nature*. New York: Routledge.
- Haraway, Donna. 1995a. Die Wissenschaftsfrage im Feminismus und das Privileg einer partialen Perspektive. In *Die Neuerfindung der Natur. Primaten, Cyborgs und Frauen*, Hg. Donna Haraway und Carmen Hammer, 73–97. Frankfurt a.M.: Campus.
- Haraway, Donna. 1995b. Ein Manifest für Cyborgs. Feminismus im Streit mit den Technowissenschaften. In *Die Neuerfindung der Natur. Primaten, Cyborgs und Frauen*, Hg. Donna Haraway und Carmen Hammer, 33–72. Frankfurt a.M.: Campus.
- Keyes, Os, Burren Peil, Rua M. Williams und Katta Spiel. 2020. Reimagining (Women's) Health: HCI, Gender and Essentialised Embodiment. *ACM Transactions on Computer-Human Interaction* 27(4): 1–42.
- Keyes, Os. 2020. Automating Autism, Disability, discourse, and Artificial Intelligence. *Journal of Sociotechnical Critique* 1(1): 1–31.
- Lee, Nicole, Turner, Paul Resnick und Genie Barton. 2019. *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*. Brookings Institute. Washington, DC, USA.
- Lemke, Thomas. 2017. Neue Materialismen. Einführung. In *Science and Technology Studies. Klassische Positionen und aktuelle Perspektiven*, Hg. Susanne Bauer, Torsten Heinemann und Thomas Lemke, 551–573. Berlin: Suhrkamp.
- Mathar, Thomas. 2010. *Der digitale Patient. Zu den Konsequenzen eines technowissenschaftlichen Gesundheitssystems*. Bielefeld: transcript.
- Mbembe, Achille. 2011. Nekropolitik. In *Biopolitik – in der Debatte*, Hg. Marianne Pierper, Thomas Atzert, Serhat Karakayali und Vassilis Tsianos, 63–96. Wiesbaden: Springer VS.
- Namberger, Verena. 2014. Rassismustheorien und die Materialität des Körpers. In *Critical matter. Diskussionen eines neuen Materialismus*, Hg. Tobias Goll, Daniel Keil und Thomas Telios, 134–150. Münster: Edition Assemblage.
- Noble, Safiya Umoja. 2018. *Algorithms of Oppression. How Search Engines reinforce Racism*. New York: New York University Press.
- Obermeyer, Ziad, Brian Powers, Christine Vogeli und Sendhil Mullainathan. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366: 447–453.

- Orwat, Carsten. 2019. *Diskriminierungsrisiken durch Verwendung von Algorithmen*. Baden-Baden: Nomos.
- Panch, Trishan, Heather Mattie und Rifat Atun. 2019. Artificial intelligence and algorithmic bias. Implications for health systems. *Journal of global health* 9(2): 1–5.
- Prietl, Bianca. 2019. Algorithmische Entscheidungssysteme revisited: Wie Maschinen gesellschaftliche Herrschaftsverhältnisse reproduzieren können. *feministische studien* 37: 303–319.
- Prietl, Bianca. 2021. Warum Ethikstandards nicht alles sind. Zu den herrschaftskonservierenden Effekten aktueller Digitalisierungskritik. *Behemoth* 14(2): 19–30.
- Schadler, Cornelia. 2017. Widerständige Apparate: Was ein anti-dualistischer und anti-dialektischer Materialismus zur Analyse von Differenz und Ungleichheit beitragen kann. In *Material turn: Feministische Perspektiven auf Materialität und Materialismus*, Hg. Christine Löw, Imke Leicht, Nadja Meisterhans und Katharina Volk, 171–186. Opladen: Barbara Budrich.
- Schmidt, Christian. 2010. Die Rolle des geistigen Eigentums im Kapitalismus der Gegenwart. *Kurswechsel* 4: 6–14.

