

Gender als Escape-Room – Künstliche Intelligenz als Katalysator

Waltraud Ernst (waltraud.ernst@jku.at) 

Abstract: Digitalisierung und Künstliche Intelligenz (KI) erhöhen die Dringlichkeit, Gender als bestehende rassistische und klassistische patriarchale Geschlechterordnung zu überwinden. Das heißt, KI wirkt wie ein Katalysator, der die Verwirklichung einer egalitären Geschlechterordnung, die rechtlich ja längst verbrieft ist, noch dringender macht. Dabei macht KI noch offensichtlicher, dass dies – wie in einem Escape-Room – nur in einer gemeinsamen Anstrengung aller Beteiligten möglich wird. Der Aufsatz zeigt die Verschärfung und Verschleierung von Diskriminierung durch KI anhand wichtiger Studien und schlägt einen intersektionalen Ansatz vor, der Geschlechterbinaritäten und stereotype Klassifizierungen transzendiert. Partizipatorische und aktivistische Ansätze werden diskutiert und eine auf Gerechtigkeit basierende Forschung zu Digitalisierung und KI gefordert.

Schlagworte: Digitalisierung, Diskriminierung, Gender, Intersektionalität, Künstliche Intelligenz

Eingereicht: 06. November 2022

Angenommen: 09. Mai 2023

Veröffentlicht: 19. Juni 2024

DOI: <https://doi.org/10.17169/oj.2024.221>

Dieser Beitrag wurde redaktionell betreut von Sara Morais dos Santos Bruss, Maria Häusl, Eva Sängler und Kathrin Ganz.

Gender als Escape-Room – Künstliche Intelligenz als Katalysator

“What social groups are classified, corralled, coerced, and capitalized upon so others are free to tinker, experiment, design, and engineer the future? [...] How may technoscience be appropriated and reimaged for more liberatory ends?” (Benjamin 2019a, 4)

Gender wird in diesem Aufsatz als diskriminierender Ordnungsraum verstanden, aus dem wir nur gemeinsam entkommen können. ‚Künstliche Intelligenz‘ (KI) wirkt dabei wie ein Katalysator in einem Experiment, ein Beschleuniger, der die Notwendigkeit zur Transformation der bestehenden diskriminierenden politischen Ordnung für immer mehr Menschen offensichtlich macht. Denn wie Ruha Benjamin in den oben zitierten Fragen zum Ausdruck bringt, gilt es erstens, zu erkennen, wie Menschen aufgrund ihrer demografischen Zuteilung in bestimmte Gruppen auf Kosten anderer Menschen Privilegien genießen. Zweitens gilt es auszuloten, wie technowissenschaftliche Entwicklungen angeeignet werden können, um aus diesen Verhältnissen zu befreien. Der vorliegende Aufsatz präsentiert und analysiert Antworten auf diese beiden Fragen sowie deren theoretische und praktische Verwobenheit. Denn die Hoffnung, auf der Basis technologischer Forschung eine KI zu entwickeln, welche die Perspektiven-, Situations- und Standortgebundenheit menschlicher Einsicht überwinden könnte, erweist sich als Illusion. Ganz im Gegenteil, so wird hier gezeigt, repräsentiert KI die Weltsicht genau derer, die KI entwickeln: Menschen, die mathematische Theorien entwickeln und Computerprogramme schreiben, Trainingsdatensätze produzieren und für die ‚lernende‘ Maschine aufarbeiten. Erst auf dieser Grundlage „webt“ eine solche Maschine in algorithmischen Systemen die entsprechenden „algebraischen Muster“ (Lovelace 1843, zitiert in Krämer 2015, 85). Ganz ‚automatisch‘ prägen auf diese Weise vorherrschende Stereotypen und faktische Ungleichheiten die neue Technologie (Eubanks 2017). KI, wie sie bisher betrieben wird, riskiert so, Ungerechtigkeit nicht zu überwinden, sondern soziale und materielle (körperliche sowie ökonomische) Ungleichheiten intransparent zu machen, zu zementieren, herzustellen und zu verschärfen (Benjamin 2019b; Noble 2018; O’Neil 2017; Prietl 2019; Nakamura 2014). Daher erweisen sich, so die These dieses Aufsatzes, folgende Fragen bei der Erforschung und Implementierung von KI als zentral: Was wird berechnet? Wer profitiert davon? Wer trägt die (sozialen) Risiken und Kosten? KI wirkt also als Katalysator in dem

Sinne, dass sie die Gemüter derer erhitzt, die um den sozialen Zusammenhalt der Menschheit bangen.

Fragen der Gerechtigkeit wurden bisher meist übersehen, wenn an technischen Innovationen geforscht wurde (Ernst 2021a). Pauschale Appelle an ‚die‘ Ethik oder ‚europäische Werte‘ verfehlen das Ziel, denn mehr oder weniger versteckt hinter diesen Appellen findet bis heute Unterdrückung statt (Ricaurte 2022). Daher ist es notwendiger denn je, zu fragen, in welcher Zukunft wir leben wollen und wie eine demokratische Zukunft erforscht, entwickelt, erstritten und auch technisch umgesetzt werden kann, in der alle Menschen ein gleich „lebbares Leben“ (Butler 2012) haben. Präzisierend lautet meine These, dass KI die Dringlichkeit erhöht, Gender als bestehende rassistische und klassistische patriarchale Geschlechterordnung gemeinsam und auf allen Ebenen zu überwinden. Hier wird also untersucht, inwiefern KI als Katalysator für die Verwirklichung einer egalitären Geschlechterordnung wirkt, die 1948 in der Allgemeinen Erklärung der Menschenrechte durch die Generalversammlung der Vereinten Nationen versprochen worden ist. Dabei wird offensichtlich, dass die Überwindung der bestehenden Geschlechterordnung wie ein Escape-Room funktioniert: Nur in der gemeinsamen Anstrengung aller Beteiligten wird ein Entkommen möglich sein. Das Anliegen und der innovative Beitrag dieses Textes ist es also, durch eine diffraktive Analyse der im Folgenden aus- und zusammengeführten intersektionalen Analysebeispiele Verschiebungsmuster für einen möglichen Ausweg sichtbar werden zu lassen. In Anlehnung an Karen Barad (2014) und Iris van der Tuin (2014) wird Diffraktion hier als Methode verwendet, als ein Ineinanderlesen von Differenzierungsprozessen, das alte Beugungsmuster („diffraction patterns“) verschieben und neue Verschiebungsmuster erzeugen kann. Zusammenhänge von Technik und Geschlecht können so in einem neuen Licht erscheinen.

Im folgenden Abschnitt „Gender und KI überdenken“ wird zunächst Gender als normative Geschlechterordnung theoretisch ausgeführt und KI in ihrer Funktionsweise erläutert. Im anschließenden Abschnitt „Intersektionale Analysen“ wird in intersektionalen Analysen identitätsstiftender KI-Produkte Diskriminierung und deren mögliche Überwindung sichtbar gemacht. Dies wird im darauffolgenden Abschnitt „Binarität überdenken“ mit Fokus auf die binäre Geschlechterdatenerfassung ergänzt. Nach und nach werden einander ergänzende Auswege aus dem Escape-Room Gender in Form von Verschiebungsmustern von Vorstellungen und technischen Auslegungen angeboten. Anschließend werden im Abschnitt „Partizipation und Aktivismus“ auf der praktischen Ebene der dargelegten „ethico-onto-epistem-ology“ (Barad 2007, 185), also der ethisch-reflektierten Wissenspraktiken in Phänomenen des Werdens, Workshops zur

partizipativen und aktivistischen Technikentwicklung beispielhaft vorgestellt. Auf diese Weise wird ein immer engmaschigeres Gewebe von möglichen sowie bereits erkannten oder realisierten radikalen Schritten aus der aktuellen Automatisierung von Diskriminierung gezeichnet; ebenso zeichnen sich Wege aus der Instrumentalisierung demografisch funktionalisierter Differenzierungen zur hierarchischen Ordnung von Menschen ab. Im Ausblick wird der Gang der Analyse zusammengefasst, zugespitzt und kontextualisiert.

Gender und KI überdenken

Geschlecht wird in der aktuellen Genderforschung als ambivalentes, historisch vielfältiges sowie vielfältig mit anderen Kategorisierungen verknüpftes und widersprüchliches Phänomen des Werdens erforscht (Ernst 2021b). Das heißt, anti-rassistische, queer-feministische Geschlechterforschung unter Berücksichtigung von Intersektionalität (Collins 2019) ermöglicht, simplifizierende, hierarchische und binäre Geschlechtervorstellungen zu erkennen und zu überwinden. So kann Vergeschlechtlichung als Prozess erforscht werden, in dem Geschlecht erst im sozio-kulturellen Zusammenhang hergestellt wird, als „doing gender“ und „doing difference“ (Fenstermaker/West 2001) und als „soziale Institution“ (Lorber 1999) Ausdruck findet sowie als „Geschlechterapparat“ und „Geschlechternorm“ (Butler 2004) veränderbar wird. Vergeschlechtlichungen von Personen und Dingen in, von und durch Anwendungen von KI umfassen daher vielschichtige und machtvolle Prozesse, die es zu erforschen gilt – gerade weil KI das, was als Wirklichkeit gilt, nicht nur vermittelt, sondern durch ihre weite Verbreitung und vielfältige Nutzung auch hervorbringt. Die vielfältigen Verknüpfungen, Überlappungs- und Verschiebungsmuster – „entanglements“ und „diffractions“ (Barad 2007) – von Gender und KI werden in diesem Aufsatz untersucht, um Auswege aus der Zementierung von Gender durch KI aufzuzeigen.

Als Katalysator wird in der Chemie ein Stoff bezeichnet, der die Reaktionsgeschwindigkeit einer chemischen Reaktion beeinflusst, ohne dabei selbst verbraucht zu werden. Dies geschieht durch Herauf- oder Herabsetzung der Aktivierungsenergie. Künstliche Intelligenz setzt also die Aktivierungsenergie für das Verhandeln von Geschlecht als nicht (so leicht) kategorisierbares herauf. Der Einsatz von KI für menschliche und soziale Zusammenhänge beschleunigt die Dringlichkeit für die Auseinandersetzung um Geschlecht als nur scheinbar eindeutig binäre Kategorie. ‚Weiblich‘ und ‚männlich‘ werden sichtbar als historisch hergestellte Vereindeutigung zum Zwecke der ungleichen Verteilung von Rechten und Pflichten, Lohn und Arbeit, Privilegien und Chancen. Denn die digitalen

Muster, die gewebt werden, scheinen Eindeutigkeit als Grundlage zu erfordern. Gender wird dabei zum Escape-Room in dem Sinne, dass sich das Verständnis von Geschlecht nur kollektiv weiterentwickeln lässt. Eine gemeinschaftliche Auseinandersetzung ist daher erforderlich, um der Eingeschlossenheit in eine veraltete patriarchale Geschlechterordnung miteinander zu entkommen. Auf diese Weise unterzieht der vorliegende Beitrag die begrifflichen und methodischen Grundlagen der interdisziplinären Forschung zu KI und Gender einer diffraktiven Analyse (Barad 2007; Ernst 2014; Bath et al. 2013). Bei einer diffraktiven Analyse wird Wirklichkeit als ständiger Prozess der Auseinandersetzung und der Veränderung verstanden. Aus den Überlagerungen von KI und Gender werden so Verschiebungsmuster herausgearbeitet, die für Prozesse der radikalen Entprivilegierung im Sinne eines dekolonialen Feminismus (Sandoval 2000) und einer interdisziplinären Zusammenarbeit auf Augenhöhe von Gender- und KI-Forschung (Crutzen 2013) wegweisend sind.

Die theoretische Herausforderung für die KI-Forschung in der Informatik besteht also darin, so meine These, Geschlecht als Phänomen des Werdens in der KI auszulegen, das heißt, der Vielfältigkeit geschlechtlicher Existenz in der KI-Forschung Raum zu geben. Das bedeutet, geschlechtliche Existenz muss in ihrer wandelbaren Verflochtenheit und Uneindeutigkeit zum Forschungsthema in der KI-Forschung gemacht werden. Auf diese Weise könnte KI-Forschung Prozesse der Entprivilegierung der Privilegierten durch Künstliche Intelligenz anstoßen. Anti-rassistische, queer-feministische KI-Forschung könnte dann in und durch Produkte wie Suchmaschinen, Übersetzungsprogramme oder digitale interaktive „software agents“ (Suchman 2003), die wie *Siri* oder *Alexa* Anweisungen empfangen und ausführen, zum Beispiel die privilegierten Besitzer*innen dieser digitalen Bediener*innen dazu anleiten, ihre Befehle zu überdenken oder selbst auszuführen und so zu neuem sozialen Lernen anregen.

Zunächst muss hierfür die Funktionsweise von Algorithmen betrachtet werden. KI basiert auf sogenanntem maschinellem Lernen, dem Management von großen Datenmengen, sogenannten ‚Big Data‘ sowie der Automatisierung von Testmethoden. Die Auswertung der Daten geschieht maschinell, automatisiert, nach einer Verknüpfung von Methoden der Datensammlung und der Mustererkennung. Aus den vergangenheitsbezogenen Daten entwickelt eine KI eine zukunftsgerichtete Wahrscheinlichkeitsrechnung auf der Basis einer konditionalen Zukunftsidee. Das heißt, Zukunftsprognosen werden aufgrund von mathematischer Verknüpfung von Daten, die in der Vergangenheit gesammelt worden sind, in Kategorien sortiert, die von (Menschen mit) Maschinen erfasst werden. Diese Daten werden in der Folge mit einer statistischen Mittelwertrechnung konditional auf die Zukunft projiziert, als ob die Zukunft lebendiger,

menschlicher Wesen aus deren Vergangenheit linear vorauszuberechnen wäre. Isolierung, Quantifizierung und Mechanisierung stellen dabei die basalen methodischen Zugänge dar. Dies erfordert eine genaue Überprüfung, welche Daten überhaupt verfügbar sind und aus welchen Prozessen Aspekte erst isoliert werden müssen, damit sie zählbar werden.

Um Datensätze zu produzieren, treffen IT-Ingenieur*innen Entscheidungen darüber, welche Daten wichtig oder relevant, gesichert oder unveränderlich, sind. Entwickler*innen von Algorithmen treffen auch Entscheidungen darüber, nach welchen Mustern und Mechanismen ausgewertet wird. IT-Ingenieur*innen entscheiden, nach welchen Kriterien sortiert, bewertet und eingeteilt wird (D'Ignazio/Klein 2020). Mehr oder weniger zufällige Korrelationen drohen dabei, zu vermeintlichen Kausalitäten zu gerinnen, wie Paola Lopez ausführte: "Datafication turns a heterogeneous, complicated reality into a supposedly homogeneous microcosm with phenomena that seem classifiable, comparable and controllable." (Lopez 2021, 8). Widersprüchliches menschliches Sein und Werden läuft so Gefahr, in linearen Vorausberechnungen festgelegt zu werden. Veränderliche Unterscheidungen werden beim Erstellen von Datensätzen bisher als fixierte, statische Differenzen ausgelegt. Im Folgenden wird anhand mehrerer konkreter Beispiele intersektionaler Analysen diese algorithmische Herstellung von Wirklichkeit im Sinne des agentuellen Realismus spezifiziert.

Intersektionale Analysen

Die Orientierung an einer Analyse vielfältig verwobener Privilegierungs- und Diskriminierungsprozesse, die nicht beim Ziel ihrer Berechnung und Sanktionierung stehen bleibt, sondern mit dem Erkenntnisinteresse ihrer strukturellen Überwindung und Verunmöglichung arbeitet, macht Ansätze der Intersektionalitätsforschung und der Queer Theory zum notwendigen Reflexionswerkzeug für die KI-Forschung. Die intersektionalen Zusammenhänge von Gender methodisch präzisierend, entwickelt Encarnación Gutiérrez Rodríguez in Auseinandersetzung mit Patricia Hill Collins die Notwendigkeit eines „Denken[s] in Relationen und die Wahrnehmung von Simultaneität“ (Gutiérrez Rodríguez 2011, 88). Sie betont die interaktive Prozesshaftigkeit der Vergeschlechtlichung in Herrschaftsverhältnissen:

„Prozesse der Feminisierung oder Maskulinisierung sind demnach einerseits in Zusammenhang mit historischen Ereignissen wie denen des Kolonialismus, Imperialismus und des modernen/kolonialen Weltsystems in Beziehung zu setzen und andererseits mit gesellschaftlichen

Verhältnissen und institutionalisierten kulturellen Praktiken, in denen hegemoniale Verständnisse in den Alltagsverstand übertragen und von den Subjekten performativ angeeignet und verkörpert werden.“ (Gutiérrez Rodríguez 2011, 89)

Daran anschließend entwickeln Ulrike Lembke und Doris Liebscher in ihrer Kritik am Konzept der Intersektionalität in der Rechtsdogmatik den Vorschlag eines „postkategorialen Antidiskriminierungsrechts“, das „die essentialisierende Zuordnung zu einer oder mehreren hierarchisch angeordneten sozialen Gruppen mit benachteiligender Intention oder Wirkung“ als das die Diskriminierung produzierende und rechtlich zu adressierende Problem erkennt und nicht besondere Persönlichkeitsmerkmale oder Gruppenzugehörigkeit (Lembke/Liebscher 2014, 283).

Im Folgenden werden intersektionale Analysen von Suchmaschinen, Trainingsdatensätzen, Algorithmen und mathematischen Modellen als Grundlage von Algorithmen vorgestellt und erörtert. Dabei wird gezeigt, wie eine algorithmische Herstellung von Wirklichkeit geschieht. Im ersten Beispiel werden die Suchergebnisse kommerzieller Suchmaschinen auf Identitätskonstruktionen hin untersucht. Safiya Umoja Noble (2013; 2018) konnte nachweisen, dass die Internetsuche nach „Black girls“ via Google auf der ersten Seite hauptsächlich Pornografie sowie sexistische und rassistische Hinweise auf Sexarbeiterinnen ergab, sowohl bei den Werbeanzeigen als auch in den unbezahlten Suchergebnissen. Noble intervenierte bei Google und fragt nach der Verantwortung: „If Google software engineers are not responsible for the design of their algorithms, then who is it?“ (Noble 2018, 66) Sie macht deutlich, dass das Unternehmen durchaus auf Kritik an den Ergebnisdarstellungen der Suchmaschine reagierte, das heißt auch reagieren kann. Noble stellt diese Diskriminierung durch die Suchmaschine in den Kontext der systematischen Entmachtung und Ausbeutung Schwarzer Frauen, insbesondere seit Beginn der kolonialen Sklaverei. Sie zeigt weiter, wie auch der nach wie vor virulente Sexismus und Rassismus in den USA Schwarze Frauen bis heute abwertet und stigmatisiert. Gerade dies wird demnach durch die Ergebnisse der Suchmaschine verschärft.

Auf der sozialen Ebene folgt aus dieser diskriminierenden Identitätskonstruktion durch die angeblich ‚automatische‘ erniedrigende Verknüpfung von Schwarzen Frauen mit Sexarbeit und Pornografie in dieser Suchmaschine, dass die Realität bedeutender Persönlichkeiten wie zum Beispiel der Literaturnobelpreisträgerin Toni Morrison oder der Computerpionierin Christine Mann Darden schwieriger denkbar ist, auch als Identifizierungspotenzial für Schwarze Frauen und Mädchen. Denn soziale Identität konstituiert sich nach Noble interaktiv: „Social identity is both a process of individual actors participating in

the creation of identity and also a matter of social categorization that happens at a social-structural level and as a matter of personal definition and external definition.” (Noble 2018, 104) Die digitale diskriminierende Konstituierung von Gruppenidentitäten ist Noble zufolge besonders folgenschwer, da das Internet zum alltäglichen Medium geworden ist, das offizielle staatliche und institutionelle Prozesse beherbergt. Das gibt dem Medium Legitimität. Wenn bei diesen digitalen Infrastrukturen nun kommerzielle Angebote dominieren und in diesen Angeboten soziale Identität entlang bestehender Privilegien und Bedürfnisse der Privilegiertesten geprägt und verkauft wird, muss Noble zufolge von Algorithmen der Unterdrückung („algorithms of oppression“) gesprochen werden.

Dies muss jedoch nicht notwendigerweise so bleiben. Schon die Datensätze, die Suchmaschinen zugrunde gelegt werden, könnten vor ihrem Einsatz einem feministischen Geschlechtertest unterzogen werden. Susan Leavy (2018) zeigt auf, wie bei ‚lernenden‘ Maschinen Algorithmen auf Trainingsdatensätze angesetzt werden, die riesige öffentliche und veröffentlichte Textquellen (wie Enzyklopädien, Tageszeitungen, literarische Werke) darstellen, in denen auf verschiedenste Weise patriarchale Geschlechterideologie der letzten Jahrhunderte eingeschrieben ist. Das führt unter anderem zur viel häufigeren verniedlichenden Verwendungen des Wortes „girls“ zur Bezeichnung erwachsener Frauen verglichen mit „boys“ für Männer. Ebenso gibt es erprobte Techniken zur Analyse von Kontexten, in denen Geschlechterbegriffe eventuell auf stereotype Weise auftauchen: “Extracting of adjectives used to describe women in training data could therefore be incorporated as part of gender proofing the textual data that is used to training machine learning algorithms.” (Leavy 2018, 2) Um der Automatisierung von Geschlechterstereotypen entgegenzuwirken, muss nach Leavy die Fähigkeit, auf Jahrzehnte feministischer Sprachkritik zurückgreifen zu können, Personen gelehrt werden, die Trainingsdatensätze auswählen und Algorithmen überprüfen. So könnten Erkenntnisse der Genderforschung mit der KI-Forschung in der Informatik neu verknüpft und wegweisende Verschiebungen als Überwindung des Status quos algorithmischer Diskriminierung hervorgerufen werden.

Die nun folgende intersektionale Analyse bezieht sich auf die staatlich automatisierte Bewertung von Bürger*innen. Wie ein Katalysator provozierte der sogenannte AMS-Algorithmus einen breiten Diskurs über die Bedeutung der Gleichwertigkeit aller Menschen und sozialen Zusammenhalt, also einen Ausweg aus der demografisch gesteuerten Abwertung letztlich großer Teile der Bevölkerung. Doris Allhutter untersuchte in einem interdisziplinären Team einen 2019 probeweise beim Österreichischen Arbeitsmarktservice (AMS) ein-

gesetzten Algorithmus (Allhutter et al. 2020). Erklärtes Ziel des AMS war die Effizienzsteigerung der Mittel. Hierfür fand durch den Algorithmus eine Kategorisierung der gemeldeten Erwerbslosen auf Basis demografischer Daten und Erwerbsarbeitsgeschichte – nach Punkten – statt: Punktabzug gab es aufgrund von Weiblichkeit, aufgrund von Mutterschaft (nicht Vaterschaft), aufgrund von Migration, aufgrund von Alter sowie aufgrund von kurzfristigen Anstellungen. Auf der Basis der Punktezahl ordnete das System ‚automatisch‘ die Erwerbslosen in drei Gruppen: Leicht Vermittelbare, die keine Maßnahmen angeboten bekommen sollten, schwerer Vermittelbare, die verschiedene Förderungen, Umschulungen und Weiterbildungen angeboten bekommen sollten, und sehr schwer Vermittelbare, die auf psychosoziale Beratung anderer Institutionen verwiesen werden sollten. Das heißt, durch mehrfachen Punktabzug nur aufgrund demografischer Klassifikation und deren Bewertung entstand die Gefahr des Ausschlusses aus der Vermittlungsförderung und damit intersektionale Diskriminierung.

Durch eine parlamentarische Anfrage kam der Aufbau des Algorithmus an die Öffentlichkeit und provozierte eine breite Diskussion in verschiedenen Medien. Am Ende wurde für viele offensichtlich, dass der AMS-Algorithmus dazu tendiert, Bürger*innen als Kund*innen zu betrachten, das Problem Erwerbslosigkeit zu individualisieren, die Makro-Ebene auszuschließen, Arbeitslosengeld-Empfänger*innen machtlos zu machen sowie aufgrund der Verknüpfung von Geschlecht, Alter, Herkunft und anderem intersektional zu diskriminieren. Bis heute wird der Fall „AMS-Algorithmus“ zwischen dem österreichischen Bundesverwaltungsgericht und dem Verwaltungsgerichtshof in Wien hin- und hergeschoben, wo die Diskriminierungsvorwürfe geprüft werden (Pflügl 2024). Nach Paola Lopez diskriminiert der AMS-Algorithmus auf der technischen, der sozio-technischen sowie auf der gesellschaftlichen Ebene, da er soziale Hierarchien technisch erst herstellt sowie schlicht abbildet. „The AMS algorithm depicts societal bias which means that it shows the aggregated individual biases of institutions and decision-makers in the realms of the labour market that are structured via intersectional inequalities.“ (Lopez 2021, 19) Lopez zufolge könnte eine solche Erkenntnis über gesellschaftliche Diskriminierung durch einen Algorithmus einen emanzipatorischen Moment eröffnen, indem Personen, die in mehrerer Hinsicht Diskriminierung erfahren, besonders gefördert würden, dies sei beim AMS-Algorithmus jedoch weder der Fall noch intendiert.

Der Einsatz von KI zur Geschlechterklassifizierung von Gesichtern führt ebenfalls zu vielschichtigen Problemen. Joy Buolamwini und Timnit Gebru untersuchten drei kommerzielle Geschlechterklassifizierungsalgorithmen für die Gesichtserkennung: „We found that all classifiers performed best for light-

er individuals and males overall. The classifiers performed worst for darker females.” (Buolamwini/Gebru 2018, 12) Die KI-Expert*innen, die durch diese Studie international für breites Aufsehen gesorgt haben, zeigen, dass erst die intersektionale Analyse der Algorithmen auf Methoden hinweist, wie die Zusammensetzung von Datensets, Auswahl der Eigenschaften und neuronalen Netzwerkarchitekturen verbessert werden kann. Das heißt, die sexistische und rassistische soziale Ordnung, die wie hier bei der Gesichtserkennungssoftware in, von und durch Anwendungen von KI automatisiert wird, kann in interdisziplinären Analysen durch Expert*innen sichtbar gemacht werden. Besondere Brisanz entsteht dadurch, dass KI das, was als Wirklichkeit gilt, nicht nur vermittelt, abbildet oder wiedergibt, sondern durch ihre weite Verbreitung und vielfältige Nutzung auch verschärft.

Algorithmen basieren auf mathematischen Modellen. Das folgende Analysebeispiel zeigt, dass es durchaus möglich ist, vertrauenswürdige mathematische Modelle von anderen zu unterscheiden. Cathy O’Neil (2017) legt ausführlich dar, was vertrauenswürdige Modelle von destruktiven Modellen als Grundlage von Algorithmen unterscheidet. Sie zeigt, inwiefern die auf mathematischen Verfahren beruhenden Anwendungen von Algorithmen auf Entscheidungen basieren, die von fehlbaren Menschen getroffen wurden – manchmal sogar mit den besten Absichten. O’Neil weist nach, wie menschliche Vorurteile, Missverständnisse und Voreingenommenheit in Softwaresysteme einfließen, die in immer höherem Maße unseren Alltag bestimmen: „Solche mathematischen Modelle sind undurchschaubar wie Götter, und ihre Funktionsweise ist nicht zu erkennen, außer für ihre Hohepriester: die Mathematiker und Computerspezialisten.“ (O’Neil 2017, 11f.) Die Autorin erläutert die Funktionsweise eines vertrauenswürdigen Modells anhand eines Baseball-Wettsystems: „[E]s erhält ständige Rückmeldungen von dem jeweiligen Vorgang aus der realen Welt, den es zu verstehen oder vorherzusagen versucht. Die Umstände ändern sich ständig, und deswegen muss auch das Modell ständig angepasst werden.“ (O’Neil 2017, 29f.) Dieses Modell stellt sie einem in den USA weit verbreiteten „Rückfallquotenmodell“ gegenüber, das Richter*innen Prognosen über die Wahrscheinlichkeit liefert, ob eine bereits verurteilte Person noch einmal straffällig wird. Daran bemisst sich das Strafmaß. Dieses Modell basiert auf den Daten aus einem Fragebogen, den Personen ausfüllen müssen, sobald sie ins Gefängnis kommen. O’Neil zeigt, wie der auf diesem Fragebogen aufgebaute Algorithmus intersektionale Diskriminierung reflektiert und verschärft, zum Beispiel wenn betroffene Personen in Stadtteilen aufgewachsen sind, die besonders häufig von der Polizei kontrolliert werden. So entsteht eine toxische Feedbackschleife.

Daher stellt O'Neil als Informatikerin und Aktivistin sieben Forderungen nach der Regulierung von Algorithmen. Diese beinhalten erstens, die Auswirkungen algorithmischer Entscheidungssysteme zu messen und ihre Algorithmen zu prüfen. Zweitens gilt es, immer deren genaue Funktion zu recherchieren. Drittens ist es notwendig, Annahmen, die dem jeweiligen Modell zugrunde liegen, zu rekonstruieren. Viertens muss von einem unabhängigen Gremium eingestuft werden, wie gerecht diese Annahmen sind. Fünftens fordert O'Neil, ungerechte Systeme auszumustern sowie positive Feedbackschleifen einzurichten, die Informationen liefern, aufgrund derer das Modell verbessert werden kann. Sechstens fordert sie, Algorithmen flächendeckend zu prüfen, ähnlich einer Betriebsprüfung. Siebtens sei es erforderlich, Algorithmen in Verbraucher*innenschutzgesetzen detailliert zu regulieren (O'Neil 2017, 281ff.). Insofern Algorithmen als Bestandteile digitaler Infrastrukturen umfassende Anwendungsgebiete zur Regelung und Regulierung des Sozialen darstellen, ist es ein wichtiger demokratischer Schritt, staatliche Aufsicht über die Produktionsprozesse und Funktionsweisen von Algorithmen einzufordern. Was als ‚Erfolg‘ algorithmischer Entscheidungssysteme gemessen wird, wenn sie eingesetzt werden, um Menschen zu bewerten, muss also neu reflektiert werden, auch weil diese meist von privaten, auf wirtschaftlichen Profit ausgerichteten Unternehmen entwickelt werden. Tatsächlich wird seit einigen Jahren an der Regulierung von KI gearbeitet. Das Europäische Parlament verabschiedete am 13. März 2024 nach über fünfjähriger Beratung das sogenannte KI-Gesetz (AI Act), um Risiken von KI zu regulieren.

Binarität überdenken

Die vorliegenden Analysen zeigen, wie grundlegende Begriffe und Methoden digitaler Infrastrukturen und der KI-Forschung zu überdenken sind. Hinsichtlich von Unterscheidungsweisen muss gefragt werden: Wie werden Differenzen hergestellt – durch simplifizierende Kategorien? Wie werden Differenzierungen gedacht – als statisch, unveränderlich? Werden Differenzierungen eindimensional gedacht – oder vielfältig und überschneidend?

Durch die Kritik der Gender Studies an der Naturalisierung binärer hierarchischer Geschlechterordnung kann der geschlechtlich gelebte und definierte Körper in seiner prozesshaften und widersprüchlichen Dimension auch in die KI-Entwicklung einbezogen werden. Der Körper kann so als mehr und mehr selbstverständlicher Bestandteil einer sozialen, psychischen, intellektuellen und materiellen Auseinandersetzung denkbar und in seinem unablässigen Gelebtwerden und Geordnetwerden und Regiertwerden relevant werden. Die

Kritik der Queer Studies an einer als natürlich gedachten binären und heteronormativen Geschlechterordnung, in der Männlichkeit privilegiert und Weiblichkeit abgewertet wird (Butler 1991), kann auch für die KI-Forschung wegweisend werden (Klipphahn-Karge/Koster/Morais dos Santos Bruss 2023). Mit ihrer Studie zu Materialisierungsprozessen aus der Perspektive der Quantenfeldtheorie „Transmaterialities“ (Barad 2015) begründete Karen Barad weiterführend die prinzipielle Unbestimmtheit von Materie, Zeit und Energie. Aus den elektromagnetischen Phänomenen zwischen Elektronen und Photonen schließt Barad auf eine prinzipielle ontologische Unbestimmtheit, ein (Un-)Geschehenmachen von Identität, was die Grundlagen des (Nicht-)Seins erschüttert. Damit begründet Barad auf der epistemologischen Ebene eine immerwährende Erneuerbarkeit der Vorstellungsräume von Geschlechtererfahrung und damit Geschlechteridentität. So werden die prinzipielle Offenheit, Veränderbarkeit, Verbundenheit und Spezifität jeder körperlichen (Geschlechter-)Erfahrung von der Ausnahme zur Normalität und Natürlichkeit.

Bisher wird Geschlecht jedoch in digitalen Formularen, die häufig die Datengrundlage für Algorithmen darstellen, meist binär erfasst. Inwiefern dadurch weder die rechtliche noch die soziale Wirklichkeit gelebten Lebens erfasst wird, wird in diesem Abschnitt erörtert. Dabei werden folgende Fragen nach dem Geschlecht präzisiert: Werden Daten von allen Geschlechtern als gleich wichtig erachtet? Werden die ‚lernenden‘ Maschinen von der Vielfalt und Ambivalenz geschlechtlicher Körper sowie geschlechtlicher und sexueller Lebensweisen unterrichtet? Was oder wer wird als Standard, als Norm und Normalität den Rechenautomaten vermittelt?

Die autoethnografische Untersuchung Katta Spiels (Spiel 2021) von vergeschlechtlichten digitalen Infrastrukturen (digitalen Formularen, die auszufüllen sind, um sich zum Beispiel ein Flugticket zu kaufen, sich bei einer Universität oder für ein Forschungsprojekt zu bewerben, eine Krankenversicherung abzuschließen, eine Zeitung zu abonnieren) veranschaulicht, dass diese die inzwischen verfassungsrechtlich garantierte geschlechtliche Vielfalt nicht abbilden. Spiel zeigt auf, wie diese digitalen Infrastrukturen nicht-binären Personen die Teilhabe mit korrektem Gendereintrag verunmöglichen. In einem zweiten Schritt legt Spiel dar, dass Betroffene sozio-technischen Mikroaggressionen nicht nur durch die technischen Infrastrukturen ausgesetzt werden, sondern oft auch durch sogenannte Ansprechpartner*innen, die Korrekturwünsche entgegennehmen.

Als Ergebnis hält Spiel fest: Digitale Formulare beziehungsweise Infrastrukturen verlangen oft ohne einsehbaren Grund einen (binären!) Geschlechtereintrag (Spiel 2021, 488). Dies sollte überdacht werden, denn der Geschlechterein-

trag stellt eine höchst intime Information dar. Es muss den Beteiligten bewusst werden, dass der Geschlechtereintrag aufgrund der Komplexität, Fluidität und Veränderbarkeit geschlechtlicher Existenz auf der gelebten, der gesetzlich regulierten und der soziotechnisch ermöglichten Ebene variiert.

Demzufolge muss Geschlecht als geschützte und schützenswerte Information betrachtet werden. Spiel gibt weiterführende Empfehlungen für größere Gendersensibilität technologischer Infrastrukturen: Das Einfordern geschlechtlicher Daten soll auf Situationen minimiert werden, in denen diese tatsächlich notwendig sind (zum Beispiel für die Erforschung von Geschlechterdiskriminierung). Die Bereitschaft zum Verbessern der digitalen Schnittstellen (Formulare) soll sowohl bei Betreiber*innen als auch bei Entwickler*innen und Designer*innen sowie beim Servicepersonal durch Gender-Schulungen erhöht werden (Spiel 2021, 488). Lehrbücher, insbesondere für Informatikstudierende, sollen so aktualisiert werden, dass Geschlechterbinarität nicht mehr als Normalfall gelehrt wird, der adäquat durch einen Boolean-Datentyp (mit nur 0 oder 1 als möglichen Werten) ausgelegt werden kann (ebd.).

Das folgende Beispiel macht deutlich, wie Unternehmen gezielt mit den vielfältigen digitalen Möglichkeiten, geschlechtliche Marker zu setzen, spielen und handeln. Rena Bivens zeigt auf, wie Facebook von 2004 bis 2008 keinen Geschlechtereintrag verlangte, dann einen binären und schließlich 2014 die Optionen für Geschlechteridentität auf 58 vervielfachte (Bivens 2017). Bivens zeigt, wie Geschlechterbinarität nach wie vor Facebooks Designstrategie dominiert, obwohl das Unternehmen suggeriert, sowohl den Nutzer*innen als auch den Werbekund*innen mit ihren nicht-binären beziehungsweise binären Ansprüchen zu entsprechen:

“While the persistence of the binary on the sign-up page in 2014 despite the new custom gender options is puzzling, it highlights the continuing tension between the software’s production of binary and non-binary conditions for existence. From a queer theory lens, we see the software’s production of gendered subjects in distinct spaces and the co-existence of multiple gendered subjects as individual users are morphed by binary and non-binary affordances. This tension facilitates Facebook’s relationships with both users and advertising clients, but at the software’s roots, Facebook continues to comply with society’s hegemonic norms.” (Bivens 2017, 887)

Diese Zwiespältigkeit und Ambivalenz zwischen dem Einfordern von Geschlechterbinarität und dem Zulassen von Geschlechtervielfalt auf einer so weit verbreiteten digitalen Kommunikationsplattform widerspiegelt abgesehen von der Profitlogik des Unternehmens auch die Brisanz des Themas auf einer weiteren

gesellschaftlichen Ebene. Hier wird deutlich, dass ungleiche Verschiebungsmuster von gelebtem und reguliertem Geschlecht sich in digitalen Werkzeugen ausdrücken und im Falle einer Steuerung durch KI intransparent und automatisiert mit menschlichen Akteur*innen interagieren.

Vor diesem Hintergrund ist die weitreichende Digitalisierung umfassender Datenbanken mit unzähligen simplifizierenden und oft veralteten Daten über Menschen, die zu verschiedenen Zeiten, mit unterschiedlichsten Zielen, unterschiedlichsten Methoden und aus unterschiedlichsten Anlässen gesammelt und geordnet worden sind, als ein asynchrones Unternehmen zu bezeichnen. Verschärft wird dies durch die Verbindung von digitalen Instrumenten mit KI, wenn also auf Basis dieser mangelhaften Datensätze in Prozessen des ‚maschinellen Lernens‘ automatisierte Entscheidungs-, Diagnose- und Prognosesysteme über Menschen hergestellt werden. In ihrem aktuellen Buch „Data Feminism“ erläutern Catherine D’Ignazio und Lauren F. Klein, inwiefern Daten zu erheben und Zahlen zu präsentieren in unserer auf zählbares Wissen fokussierten Gesellschaft unverzichtbar ist, um Diskriminierung aufzuzeigen oder Schritte der Emanzipation nachvollziehbar zu machen (D’Ignazio/Klein 2020). Denn zählbare Daten und Zahlenverhältnisse erzeugen auf faszinierend effektive Weise Faktizität. Weder Daten noch Zahlen sprechen den Autor*innen zufolge jedoch für sich selbst. Sie ergeben nur durch sorgfältige Kontextualisierung und Interpretation einen Sinn. Daher muss immer wieder abgewogen werden, ob, wann und wie nach Geschlecht gefragt wird: “The ethical complexity of whether to count gender, when to count gender, and how to count gender illuminates the complexity of acts of classification against the backdrop of structural oppression.” (D’Ignazio/Klein 2020, 111) Die strukturelle Unterdrückung stellt demnach die Realität dar, vor deren Hintergrund Datafizierung untersucht werden muss: “An intersectional feminist approach to counting insists that we examine and, if necessary, rethink the assumptions and beliefs behind our classification infrastructure, as well as consistently probe who is doing the counting and whose interests are served.” (D’Ignazio/Klein 2020, 123) Dies ist insbesondere im Zusammenhang aktueller Förderung von Digitalisierung, Datafizierung und KI relevant, wo nicht selten suggeriert wird, dass Daten einfach als anonymisierte und entkontextualisierte Rohdaten gesammelt, gespendet und erhoben werden können.

Welche Daten mit welchen Klassifizierungen als Repräsentant*innen der realen Welt durch maschinelles Lernen in große mathematische Modellierungsprojekte zur Vorhersage zukünftiger Realitäten einfließen, muss also einer ständigen intersektionalen Kontrolle und Reflexion unterzogen werden. Daher fungiert KI als Katalysator für eine dringende wirklichkeitsgerechte, intersektionale Auseinandersetzung um Geschlecht. Denn die sozialen Zukunftsmuster, die auf-

grund der hier analysierten Vorgehensweisen sichtbar werden, sind für viele Menschen unerträglich.

Partizipation und Aktivismus

Wie kann also die Erkenntnis von Ambivalenz, geschlechtlicher Vielfalt und menschlicher Ambiguität in unvorhersehbaren Verbindungen und dem Eingebettetsein in intersektionale Zusammenhänge von Privilegierung und Diskriminierung gerade jetzt die Digitalisierung und Automatisierung von Stereotypen in der KI verhindern? Kann KI-Forschung in der Informatik geschlechtliche Vielfalt und menschliche Ambiguität methodisch aufgreifen? Kann diese Auseinandersetzung, wie bei der kollektiven Suche nach einem Ausweg aus dem Escape-Room, egalitäre Alternativen zur jahrhundertealten sozialpolitischen Instrumentalisierung demografisch funktionalisierter Differenzen für eine hierarchische Ordnung von Menschen hervorbringen?

Partizipation und Aktivismus über die etablierten Systeme der Wissensproduktion hinaus haben schon mehrfach ihre erkenntnisleitende Kraft entfaltet. So zeigt Chela Sandoval in ihrer wegweisenden Studie „New Sciences: Cyborg Feminism and the Methodology of the Oppressed“ (Sandoval 1999), wie differenzielles Bewusstsein als permanente Neuaufschlüsselung von Raum und Grenzen, von horizontalen und vertikalen Anordnungen oppositioneller Kräfte gedacht werden kann:

“Under postmodern disobediences the self blurs around the edges, shifts ‘in order to ensure survival’, transforms according to the requisites of power, all the while, under the guiding force of the methodology of the oppressed carrying with it the integrity of a self-conscious awareness of the transformations desired, and above all, a sense of the impending ethical and political changes that those transformations will enact.” (Sandoval 1999, 260)

Sandoval stellt Methoden zur Dekolonisierung sozialer Vorstellungskraft vor, die Herrschaftstechnologien pseudo-eindeutiger demografischer Klassifizierung übersteigen und Wege für Koalitionen und Solidaritäten jenseits demografischer Identitäten der europäischen Moderne eröffnen (Ernst 2020). Vielleicht könnte dieses Aufgreifen von politischem Aktivismus auch an partizipatives Forschen früherer feministischer Ansätze anknüpfen (Mies 1984)?

Das Anliegen, kollektive Bewusstseinsprozesse über die radikale institutionelle Entgrenzung von Erkenntnisprozessen anzuregen, wird im Folgenden beispielhaft veranschaulicht. In einem Projekt, das soziale Fragen der Diskriminierung mit technologischer Innovation und künstlerischen Prozessen verband,

fragte das Institut für Frauen- und Geschlechterforschung der Johannes Kepler Universität Linz gezielt, „how to become a high-tech anti-discrimination activist collective“ (IFG@Ars Electronica Festival 2020), und lud Lisa Nakamura und Safiya Umoja Noble ein, sich in einem Vortrag dieser Frage zu widmen. Zusätzlich fanden Workshops statt, die die Frage auf partizipative und aktivistische Weise präzisierten und methodisch umsetzten. Doris Allhutter widmete sich in ihrem Workshop dem Anliegen „When I Encountered Discriminating IT-Systems and Did not Want to Take it Anymore“. Deconstructing Affective Entanglements in Society-Technology Relations“ mit der Methode des ‚Mind Scriptings‘ (Allhutter 2021), bei der unbewusste Selbstverständlichkeiten von Technologie-Entwickler*innen und -Benutzer*innen kollektiv dekonstruiert werden. Astrid Mager und Hong Phuc Dang vermittelten in ihrem Workshop „How to Create Your Own AI Device with SUSI.AI. An Open Source Platform for Conversational Web“ historisches, theoretisches und praktisches Wissen, wie Suchmaschinen und Service-Assistenten wie *Alexa*, *Siri* oder *Google Home* funktionieren und wie gemeinsam an Gerechtigkeit orientierte Instrumente entstehen (Mager 2016). Adriana Torres Topaga, Martyna Lorenc und Andrea Maria Handler entwickelten in ihrem Workshop „Phantom Data in Our Bodies and Imagination“ mit den Teilnehmenden ein Verständnis davon, wie unsere Vorstellungen ihr eigenes Datenset haben (unsere Erfahrungen, Erinnerungen, Gefühle, Annahmen und Erkenntnisse). Mit Körperübungen und performativen Spielen wurde erlebbar, was unseren Vorstellungsraum einnimmt und was ihm entgeht. Nushin Isabelle Yazdani erarbeitete gemeinsam mit Karla Zavala und Adriaan Odendaal in ihrem Workshop „[D/R]econstructing AI. Dreams of Visionary Fiction and Zine-Making“ mit den Teilnehmenden auf der Grundlage der Einsicht in die Normativität algorithmischer Entscheidungssysteme spekulative Ansätze für nicht-diskriminierende Systeme in einem digitalen Zine (Yazdani/Zavala/Odendaal 2020). Auf diese Weise konnte durch die genannten methodischen Ansätze ein gemeinsamer interdisziplinärer und intersektionaler Erkenntnisprozess von Entwickler*innen, Forscher*innen und Benutzer*innen entstehen, um systematischer Diskriminierung mit digitalen Werkzeugen und algorithmischen Entscheidungssystemen – wie aus einem Escape-Room – zu entkommen.

Partizipatorische Methoden, die in der feministischen Technikforschung entwickelt wurden, können also auch im Bereich der Digitalisierung und der KI-Forschung eingesetzt werden (Ernst 2019). Dabei ist die Frage zu klären, ob und gegebenenfalls wie KI mit der Widersprüchlichkeit, Veränderlichkeit und Ambivalenz personaler Identität und Interaktion denkbar wird. Des Weiteren scheint eine breitere öffentliche Diskussion darüber notwendig, welche sozialen

Prozesse und Fragestellungen überhaupt digitalisiert werden sollen und wo der Einsatz von KI nicht weiter gefördert werden soll.

Ausblick

Anhand von aktuellen Analysebeispielen wurde in diesem Beitrag veranschaulicht, wie entscheidend es aus dekolonialer und queer-feministischer Sicht ist, welche Daten in welchem Zusammenhang zur Herstellung von welchen Datensätzen erhoben werden und wer dadurch auf welche Weise erfasst, erhoben, realisiert und damit exponiert wird. Daher erweist sich die Untersuchung und die Transparenz der Herstellungsprozesse von Datensätzen, digitalen Infrastrukturen und algorithmischen Entscheidungssystemen als zentral. Auf diese Weise konnten hier Verschiebungsmuster sowie ein immer engmaschigeres Gewebe von radikalen Schritten als Ausweg aus der aktuellen Automatisierung von Diskriminierung sowie aus der Instrumentalisierung demografisch funktionalisierter Differenzen zur hierarchischen Ordnung von Menschen gezeichnet werden.

Die Herstellung eines informierten, breiten und öffentlichen Diskussionsraums, der Verschiebungsmuster dekolonialer und queerer Feminismen mit kritischer KI-Expertise und Digitalisierungsforschung zusammenbringt, erscheint also am Ende als Ausweg aus dem Escape-Room automatisierter Unterdrückung. Daher wird abschließend noch einmal die Frage nach der ethischen und politischen Grundlage gestellt, woran sich eine kollektive Anstrengung, aus dem Escape-Room veralteter patriarchaler und kolonialer kapitalistischer Geschlechterherrschaft zu entkommen, orientieren kann. Judith Butler gibt in ihrer Dankesrede bei der Verleihung des Theodor-W.-Adorno-Preises der Stadt Frankfurt am Main 2012 einen hilfreichen Hinweis in der Erkenntnis radikaler menschlicher Interdependenz und Verletzlichkeit:

„Unsere gemeinsame Gefährdung durch Prekarität ist nur ein Grund unserer potenziellen Gleichheit und unserer wechselseitigen Verpflichtung zur gemeinsamen Schaffung der Bedingungen für ein lebbares Leben. Indem wir uns eingestehen, dass wir einander brauchen, bekennen wir uns zugleich zu grundlegenden Prinzipien der sozialen und demokratischen Bedingungen dessen, was wir als ‚das gute Leben‘ bezeichnen.“ (Butler 2012, 20)

Diese Möglichkeit der ethischen Orientierung an einem guten Leben, die Adorno unter den gegebenen Verhältnissen von Privilegierung und Unterdrückung bis hin zur Vernichtung verneinte, macht Butler hier also in der Verbindung von Ethik mit Politik zur kollektiven Bedingung eines lebbares Lebens für alle Einzelnen in Interdependenz.

Um einen solchen Ansatz für neue digitale Voraussage- und Risikobewertungstechnologien nutzbar zu machen, ist eine Erweiterung unserer Vorstellungswelten im Sinne von Sandoval (1999; 2000) notwendig, damit wir bestehende Herrschafts- und Unterdrückungsstrukturen klar erkennen und darüber hinausgehen. Wie Benjamin aufzeigt, gibt es dafür schon Beispiele, wie zum Beispiel ein „White-Collar Early Warning System“, das einen Alarm auslöst, wenn man in die Nähe von Orten gerät, an denen die Wahrscheinlichkeit, dass finanzwirtschaftliche Verbrechen passieren, erhöht ist (Benjamin 2019b, 196). Ein weiteres Beispiel besteht in einer Gesichtserkennungssoftware, die wahrscheinliche finanzwirtschaftliche Straftäter*innen markiert. Für das Trainingsdatenset zum Design des Algorithmus wurden 7.000 Fotos aus dem LinkedIn-Netzwerk heruntergeladen, die zu Personen gehörten, die bestimmte professionelle Kriterien erfüllten und daher überwiegend männliche, weiße Gesichter zeigten (Benjamin 2019b, 196f.). Zusammengefasst geht es Benjamin ähnlich wie Butler darum, kollektives Freisein und Gedeihen zu materialisieren: “The task, then, is to challenge not only forms of discriminatory design in our inner and outer lives, but to work with others to imagine and create alternatives to the *techno quo* – business as usual when it comes to technoscience – as part of a larger struggle to materialize collective freedoms and flourishing.” (Benjamin 2019a, 12) Hierfür ist es Benjamin zufolge notwendig, sich einer befreienden Vorstellungskraft zu bedienen, die auf einer „black radical tradition“ (ebd.) und auf Einsichten und Strategien aufbaut, die Gerechtigkeit zur Grundlage haben.

Anschließend an Barads agentiellen Realismus wurde also hier die Produktion von Wissen als Teil eines Phänomens betrachtet, in dem das, worüber Wissen erzeugt wird, erst in einer ganz bestimmten Weise hervorgebracht wird. Entsprechend dieser „ethico-onto-epistem-ology“ (Barad 2007) wird hier auf der Meta-Ebene argumentiert, dass die Geschlechterordnung in genau der Weise, in der sie be- oder erforscht wird, (auch immer wieder neu) realisiert wird. Dies führt zu einer konsequenten Kontextualisierung von Erkenntnis, als „entanglements“ in sozialen Prozessen. Queer-feministisches und anti-rassistisches oder transgressives, emanzipatorisches Wissen kann demnach entsprechend kontextualisierte Realitäten aufzeigen und plausibel machen.

Literatur

- Allhutter, Doris (2021): Memory Traces in Society-Technology Relations. How to Produce Cracks in Infrastructural Power. In: Hamm, Robert (Hg.): Reader Collective Memory-Work. Sligo, Ireland: BeltraBooks, 426–452. <https://collectivememorywork.net/wp/wp-content/uploads/2021/10/Reader-Collective-Memory-Work-ebook.pdf>
- Allhutter, Doris/Cech, Florian/Fischer, Fabian/Grill, Gabriel/Mager, Astrid (2020): Algorithmic Profiling of Job Seekers in Austria. How Austerity Politics Are Made Effective. In: *Frontiers in Big Data* 3, 1–17. doi: [10.3389/fdata.2020.00005](https://doi.org/10.3389/fdata.2020.00005)
- Barad, Karen (2015): Transmaterialities. Trans*/Matter/Realities and Queer Political Imaginings. In: *GLQ. A Journal of Lesbian and Gay Studies* 21 (2–3), 387–422. doi: [10.1215/10642684-2843239](https://doi.org/10.1215/10642684-2843239)
- Barad, Karen (2014): Diffracting Diffraction. Cutting Together-Apart. In: *Parallax* 20 (3), 168–187. doi: [10.1080/13534645.2014.927623](https://doi.org/10.1080/13534645.2014.927623)
- Barad, Karen (2007): Meeting the Universe Halfway. Quantum Physics and the Entanglement of Matter and Meaning. Durham, London: Duke University Press. doi: [10.2307/j.ctv12101zq](https://doi.org/10.2307/j.ctv12101zq)
- Bath, Corinna/Meißner, Hanna/Trinkhaus, Stephan/Völker, Susanne (Hg.) (2013): *Geschlechter Interferenzen. Wissensformen – Subjektivierungsweisen – Materialisierungen*. Berlin: LIT Verlag.
- Benjamin, Ruha (Hg.) (2019a): *Captivating Technology. Race, Carceral Technoscience, and Liberatory Imagination in Everyday Life*. Durham, London: Duke University Press. doi: [10.1215/9781478004493](https://doi.org/10.1215/9781478004493)
- Benjamin, Ruha (2019b): *Race After Technology. Abolitionist Tools for the New Jim Code*. Cambridge/Massachusetts: Polity Press.
- Bivens, Rena (2017): The Gender Binary Will not be Deprogrammed: Ten Years of Coding Gender on Facebook. In: *New Media & Society* 19 (6), 880–898. doi: [10.1177/1461444815621527](https://doi.org/10.1177/1461444815621527)
- Buolamwini, Joy/Gebru, Timnit (2018): Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: *Proceedings of Machine Learning Research* 81, 1–15.
- Butler, Judith (2012): Kann man ein gutes Leben im schlechten führen? Dankesrede bei der Verleihung des Adorno-Preises in der Frankfurter Paulskirche am 11. September 2012. In: *Frankfurter Rundschau* 2012, 1–20.
- Butler, Judith (2004): *Undoing Gender*. New York, London: Routledge. doi: [10.4324/9780203499627](https://doi.org/10.4324/9780203499627)

- Butler, Judith (1991): *Das Unbehagen der Geschlechter*. Übersetzung von Katharina Menke. Frankfurt am Main: Suhrkamp.
- Collins, Patricia Hill (2019): *Intersectionality as Critical Social Theory*. Durham and London: Duke University Press.
- Crutzen, Cecile K. M. (2013): Nicht-menschlich ist auch Gender. In: *Informatik-Spektrum* 36 (3), 309–318. doi: [10.1007/s00287-013-0697-9](https://doi.org/10.1007/s00287-013-0697-9)
- D'Ignazio, Catherine/Klein, Lauren F. (2020): *Data Feminism*. Cambridge, Massachusetts, London, England: The MIT Press.
- Ernst, Waltraud (2021a): Feministische Technikphilosophie. In: Grunwald, Armin/Hillerbrand, Rafaela (Hg.): *Handbuch Technikethik*. 2. Auflage. Berlin, Heidelberg: J. B. Metzler Verlag, 114–118. doi: [10.1007/978-3-476-04901-8_21](https://doi.org/10.1007/978-3-476-04901-8_21)
- Ernst, Waltraud (2021b): Phänomene des Werdens. Intersektionalität, Queer, Postcolonial, Diversity und Disability Studies als Orientierungen für die Medienforschung. In: Dorer, Johanna/Geiger, Brigitte/Hipfl, Brigitte/Ratkovic, Viktorija (Hg.): *Handbuch Medien und Geschlecht. Perspektiven und Befunde der feministischen Kommunikations- und Medienforschung*. Wiesbaden: Springer VS, 57–72. doi: [10.1007/978-3-658-20712-0_6-1](https://doi.org/10.1007/978-3-658-20712-0_6-1)
- Ernst, Waltraud (2020): Feministische Netzpolitik und Netzaktivismus. In: Thomas, Tanja/Wischermann, Ulla (Hg.): *Feministische Theorie und Kritische Medienkulturanalyse. Ausgangspunkte und Perspektiven*. Bielefeld: transcript, 522–537.
- Ernst, Waltraud (2019): Technikverhältnisse. Methoden feministischer Technikforschung. In: Beate Kortendiek, Birgit Riegraf, Katja Sabisch (Hg.): *Handbuch Interdisziplinäre Geschlechterforschung*. Wiesbaden: Springer, 447–455. doi: [10.1007/978-3-658-12496-0_41](https://doi.org/10.1007/978-3-658-12496-0_41)
- Ernst, Waltraud (2014): Diffraction Patterns? Shifting Gender Norms in Biology and Technology. In: Ernst, Waltraud/Horwath, Ilona (Hg.): *Gender in Science and Technology. Interdisciplinary Approaches*. Bielefeld: transcript, 147–163. doi: [10.1515/transcript.9783839424346.147](https://doi.org/10.1515/transcript.9783839424346.147)
- Eubanks, Virginia (2017): *Automating Inequality. How High-tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.
- Fenstermaker, Sarah/West, Candace (2001): "Doing Difference" Revisited. Probleme, Aussichten und der Dialog in der Geschlechterforschung. In: Heintz, Bettina (Hg.): *Geschlechtersoziologie*. Wiesbaden: Westdeutscher Verlag, 236–249.

- Gutiérrez Rodríguez, Encarnación (2011): Intersektionalität. Oder: Wie nicht über Rassismus sprechen? In: Hess, Sabine/Langreiter, Nikola/Timm, Elisabeth (Hg.): Intersektionalität revisited. Empirische, theoretische und methodische Erkundungen. Bielefeld: transcript, 77–100. doi: [10.14361/transcript.9783839414378.77](https://doi.org/10.14361/transcript.9783839414378.77)
- IFG@Ars Electronica Festival(2020):How to Become a High-tech Anti-discrimination Activist Collective. <https://www.jku.at/institut-fuer-frauen-und-geschlechterforschung/veranstaltungen-archiv/tagungen-des-ifg/ifgars-electronica-festival-2020-how-to-become-a-high-tech-anti-discrimination-activist-collective> (06.06.2024).
- Klipphahn-Karge, Michael/Koster, Ann-Kathrin/Morais dos Santos Bruss, Sara (Hg.) (2023): Queere KI. Zum Coming-out smarterer Maschinen. Bielefeld: transcript Verlag. doi: [10.14361/9783839461891](https://doi.org/10.14361/9783839461891)
- Krämer, Sybille (2015): Wieso gilt Ada Lovelace als die ‚erste Programmiererin‘ und was bedeutet überhaupt ‚programmieren‘? In: Krämer, Sybille (Hg.): Ada Lovelace. Die Pionierin der Computertechnik und ihre Nachfolgerinnen. Paderborn: Wilhelm Fink, 75–89. doi: [10.30965/9783846759868_007](https://doi.org/10.30965/9783846759868_007)
- Leavy, Susan (2018): Gender Bias in Artificial Intelligence. The Need for Diversity and Gender Theory in Machine Learning. In: 2018 ACM/IEEE 1st International Workshop on Gender Equality in Software Engineering. Piscataway, NJ: IEEE. doi: [10.1145/3195570.3195580](https://doi.org/10.1145/3195570.3195580)
- Lembke, Ulrike/Liebscher, Doris (2014): Postkategoriales Antidiskriminierungsrecht? Oder: Wie kommen Konzepte der Intersektionalität in die Rechtsdogmatik? In: Meier, Isabella/Apostolovski, Veronika/Starl, Klaus/Schmidlechner, Karin Maria/Philipp, Simone (Hg.): Intersektionelle Benachteiligung und Diskriminierung. Soziale Realitäten und Rechtspraxis. Baden-Baden, Wien, Zürich, St. Gallen: Nomos; Facultas.wuv; Dike, 261–290.
- Lopez, Paola (2021): Bias Does not Equal Bias. A Socio-technical Typology of Bias in Data-based Algorithmic Systems. In: Internet Policy Review 10 (4). doi: [10.14763/2021.4.1598](https://doi.org/10.14763/2021.4.1598)
- Lorber, Judith (1999): Gender Paradoxien. Opladen: Leske + Budrich.
- Mager, Astrid (2016): Search Engine Imaginary. Visions and Values in the Co-production of Search Technology and Europe. In: Social Studies of Science, 1–23. doi: [10.3390/isis-summit-vienna-2015-T3.3009](https://doi.org/10.3390/isis-summit-vienna-2015-T3.3009)
- Mies, Maria (1984): Methodische Postulate zur Frauenforschung – dargestellt am Beispiel der Gewalt gegen Frauen. In: Beiträge zur feministischen Theorie und Praxis 7 (11), 7–25.

- Nakamura, Lisa (2014): Gender and Race Online. In: Graham, Mark/Dutton, William H. (Hg.): *Society and the Internet*. Oxford University Press, 81–95. doi: [10.1093/acprof:oso/9780199661992.003.0006](https://doi.org/10.1093/acprof:oso/9780199661992.003.0006)
- Noble, Safiya Umoja (2018): *Algorithms of Oppression. How Search Engines Reinforce Racism*. New York: New York University Press. doi: [10.2307/j.ctt1pwt9w5](https://doi.org/10.2307/j.ctt1pwt9w5)
- Noble, Safiya Umoja (2013): Google Search. Hyper-visibility as a Means of Rendering Black Women and Girls Invisible. In: *InVisible Culture* 19, 1–23. doi: [10.47761/494a02f6.50883fff](https://doi.org/10.47761/494a02f6.50883fff)
- O’Neil, Cathy (2017): *Angriff der Algorithmen. Wie sie Wahlen manipulieren, Berufschancen zerstören und unsere Gesundheit gefährden*. München: Carl Hanser Verlag. doi: [10.3139/9783446257788](https://doi.org/10.3139/9783446257788)
- Pflügl, Jakob (2024): Entscheidet die KI über Jobs? Höchstgericht lässt AMS-Algorithmus erneut prüfen. In: *Der Standard* vom 10.02.2024. <https://www.derstandard.at/story/3000000206845/entscheidet-die-ki-ueber-jobs-hoechstgericht-laesst-ams-algorithmus-erneut-pruefen> (06.06.2024).
- Prietl, Bianca (2019): Algorithmische Entscheidungssysteme revisited: Wie Maschinen gesellschaftliche Herrschaftsverhältnisse reproduzieren können. In: *Feministische Studien* 37 (2), 303–319. doi: [10.1515/fs-2019-0029](https://doi.org/10.1515/fs-2019-0029)
- Ricaurte, Paola (2022): Ethics for the Majority World: AI and the Question of Violence at Scale. In: *Media, Culture & Society* 44 (4), 726–745. doi: [10.1177/01634437221099612](https://doi.org/10.1177/01634437221099612)
- Sandoval, Chela (2000): *Methodology of the Oppressed*. Minneapolis: University of Minnesota Press.
- Sandoval, Chela (1999): *New Sciences: Cyborg Feminism and the Methodology of the Oppressed*. In: Wolmark, Jenny (Hg.): *Cybersexualities. A Reader on Feminist Theory, Cyborgs and Cyberspace*. Edinburgh: Edinburgh University Press, 247–263. doi: [10.1515/9781474473668-018](https://doi.org/10.1515/9781474473668-018)
- Spiel, Katta (2021): “Why are they all obsessed with Gender?”. (Non)binary Navigations Through Technological Infrastructures. In: *Designing Interactive Systems Conference 2021*. New York, NY, USA: ACM, 478–494. doi: [10.1145/3461778.3462033](https://doi.org/10.1145/3461778.3462033)
- Suchmann, Lucy (2003): Figuring ‚service‘ in discourses of ICT: The case of software agents. In: Weber, Jutta/Bath, Corinna (Hg.): *Turbulente Körper, soziale Maschinen. Feministische Studien zur Technowissenschaftskultur*. Opladen: Leske + Budrich, 65–74.
- Van der Tuin, Iris (2014): Diffraction as a Methodology for Feminist Onto-Epistemology. On Encountering Chantal Chaawaf and Posthuman Interpellation. In: *Parallax* 20 (3), 231–244. doi: [10.1080/13534645.2014.927631](https://doi.org/10.1080/13534645.2014.927631)

Yazdani, Nushin Isabelle/Zavala, Karla/Odendaal, Adriaan (2020): [D/R]econstructing AI. Dreams of Visionary Fiction. <https://algorithmsoflatecapitalism.tumblr.com/zines> (06.06.2024).