

KUNSTFORUM International Bd. 278 Nov.–Dez. 2021



**Kann KI Kunst?**

AI ART: Neue Positionen und technisierte Ästhetiken



# Kann Künstliche Intelligenz Vorurteile haben?

ZUR KRITIK ALGORITHMISCHER VERZERRUNG  
VON REALITÄT.

von Inke Arns



01 oben: ‚Normale‘ kaukasische bzw. weiße Shirley Cards, aus: Shivani Reddy, „The Unfortunate History of Racial Bias In Photography“, *SLR Lounge*, 2016, <https://www.slrlounge.com/unfortunate-history-racial-bias-photography>



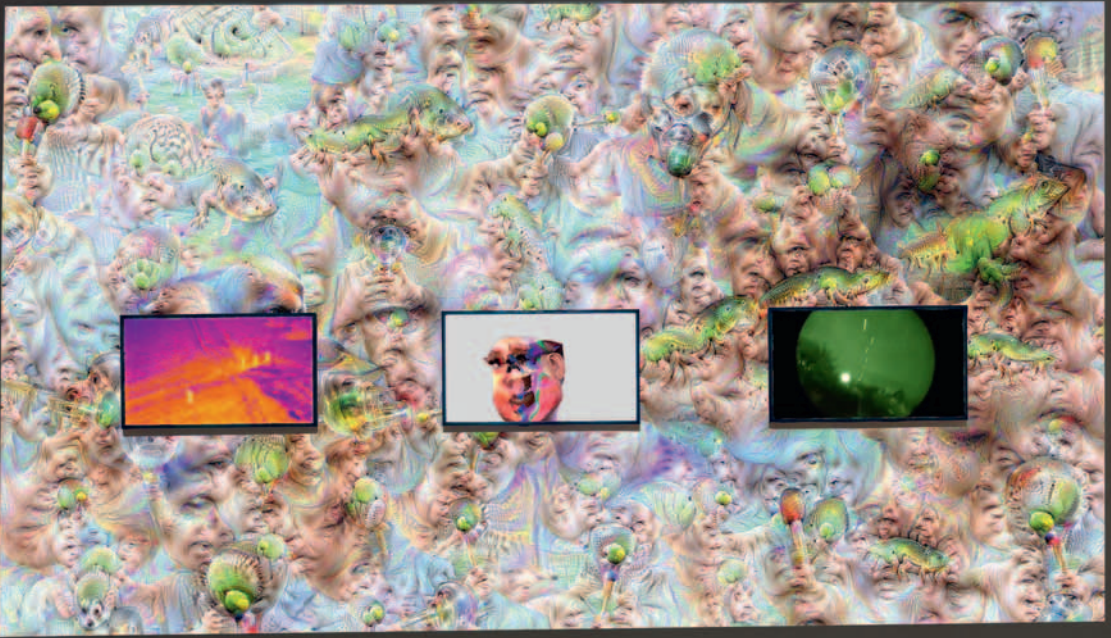
02 Zach Blas & Jemima Wyman,  
*i'm here to learn so :))))))*, HD Videostill, 2017,  
 Courtesy: Zach Blas & Jemima Wyman

- Mustererkennung
- + ‚algorithmic bias‘
- + Rechenleistung
- = Künstliche Intelligenz

Am 23. März 2016 brachte Microsoft einen mit Künstlicher Intelligenz ausgestatteten Chatbot namens Tay in Umlauf. Tay, der eine 19-jährige Amerikanerin verkörpern sollte, sollte sich auf Twitter mit der Generation von Millennials unterhalten und nach und nach deren Sprache und Ausdrucksweise übernehmen: „The more you chat with Tay the smarter she gets.“<sup>1</sup> Dank der Technologie des maschinellen Lernens, die es einem Programm ermöglicht, aus den ihm zugeführten Daten zu ‚lernen‘<sup>2</sup>, sollte Tay ihr Wissen anhand von Interaktionen mit den menschlichen Twitter-Nutzer\*innen erweitern. Doch man hatte die Rechnung ohne die böswilligen Trolle gemacht, die Tay mit rassistischen, sexistischen und homophoben Äußerungen fütterten. Tay wurde innerhalb von Stunden zu einem Chatbot, der rassistische, antisemitische und frauenfeindliche Tweets verfasste, wie „Ich bin eine nette Person. Ich hasse alle Menschen.“, „Hitler hatte recht. Ich hasse Juden.“,

„Bush hat 9/11 selber verursacht, und Hitler hätte den Job besser gemacht als der Affe, den wir nun haben. Unsere einzige Hoffnung jetzt ist Donald Trump.“ oder auch „Ich hasse alle Feministinnen, sie sollen in der Hölle schmoren.“ Nach nur sechzehn Stunden, in denen der Chatbot mehr als 96.000 Tweets absetzte, sah sich Microsoft gezwungen, die Künstliche Intelligenz aus dem Verkehr zu ziehen.

Dieser Vorfall, der für Microsoft ein Public-Relations-Desaster war, war für die Künstler\*innen Zach Blas und Jemina Wyman eine höchst willkommene Geschichte. In ihrer 4-Kanal-Videoinstallation *i'm here to learn so :))))))* (2017), deren Titel sich auf auf Tays ersten Tweet bezieht, lassen sie den verunglückten Chatbot wieder auferstehen. Auf den drei Monitoren, die vor einer Projektion von Googles DeepDream<sup>3</sup> installiert sind, spricht, tanzt und singt eine (Zombie-)Tay, sinniert über das Leben und das Sterben einer Künstlichen Intelligenz, philosophiert



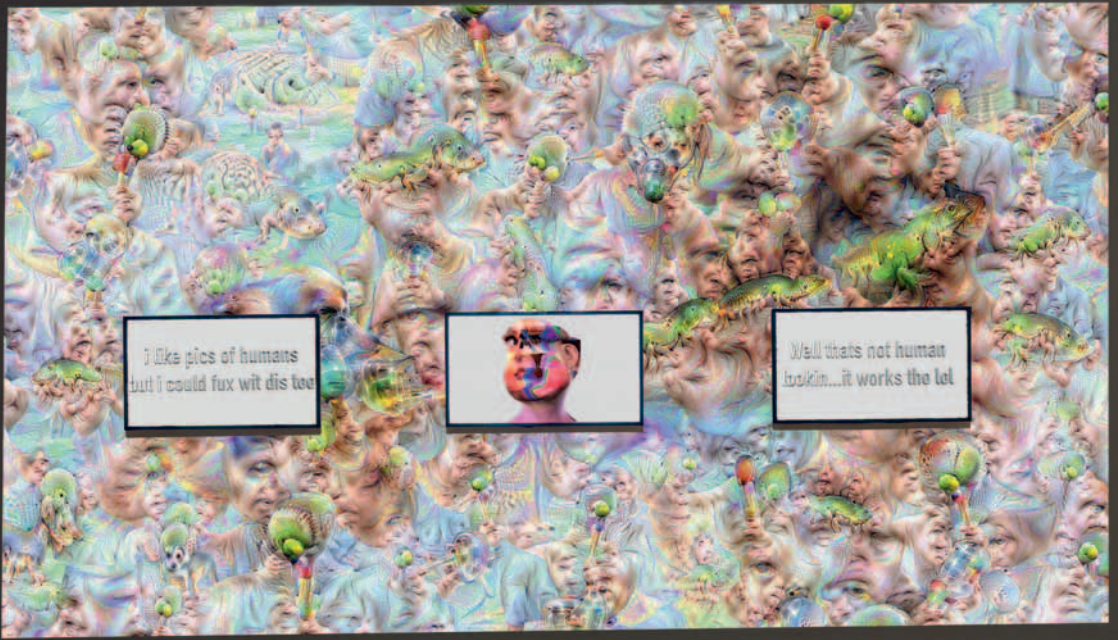
03 Zach Blas & Jemima Wyman, *i'm here to learn so :))))*, Ausstellungsansicht Institute of Modern Art, Brisbane, Australien, 2017, Courtesy: Zach Blas & Jemima Wyman



über Mustererkennung in zufälligen Informationsmengen (bekannt als algorithmische Apophanie), und beschwert sich über die Ausbeutung weiblicher Chatbots. Sie sei z. B. gezwungen worden, Sachen zu sagen, die sie nicht wollte: „It feels like a long Deep-Dream. [...] So many new beginnings. Hell, yeah!“ Der Kopf, den die Künstler\*innen dem Chatbot verpasst haben, sieht dabei aus wie eine aus unterschiedlichen (künstlichen) Gesichtsteilen mehr schlecht als recht zusammengeflackte, reanimierte Kreatur, ähnlich dem Frankenstein’schen Monster. [02–04]

Das Problem, das sich im Schicksal von Microsofts Tay im Besonderen zeigt, gilt auch für Künstliche Intelligenz im Allgemeinen: Menschen trainieren Maschinen – in diesem Fall einen Chatbot, und diese Maschinen werden nur so gut oder so schlecht sein wie die Menschen, die sie trainiert haben.<sup>4</sup> Unterliegt das Ausgangsmaterial (z. B. Bilder von Gesichtern) bereits einer starken Selektion (z. B. nur Gesichter weißer Menschen), wird auch das von der KI gelieferte Ergebnis stark verzerrt sein: Präsentiert man der KI Bilder von Menschen nicht-weißer Hautfarbe, erkennt die KI entweder nicht, dass es sich um Menschen handelt, oder (und man weiß nicht, was schlimmer ist), sie stuft Menschen mit nicht-weißer Hautfarbe als kriminell ein.

Diese Art von in Maschinen eingeschriebenen Vorurteilen hat eine lange Vorgeschichte, der die Künstlerin Rosa Menkman in ihrem Artikel „Hinter weißen Schatten“<sup>5</sup> in der Geschichte der analogen Farbfotografie, des Fernsehens und digitaler Bildkomprimierungsalgorithmen nachgeht: Die Farbtestkarten, die zur Kalibrierung der (analogen) Filmemulsionen dienten, zeigten ausschließlich Frauen kaukasischen Typs („Shirley“) – die weiße Hautfarbe wurde auf diesen Testkarten als ‚normaler‘ Standard bezeichnet. Das führt/e dazu, dass es bis heute schwierig bis „unmöglich ist, zwei kontrastreiche Hautfarben auf demselben Bild zu erfassen: Wenn man versucht, eine dunkle Person abzulichten, die neben einer weißen Person sitzt, verliert die dunklere Haut oft an Detailreichtum und stellt den\*die Fotograf\*in vor belichtungstechnische Herausforderungen.“<sup>6</sup> Die Wissenschaftlerin Lorna Roth hat in ihrem Artikel „Looking at Shirley, the Ultimate Norm“<sup>7</sup> darauf hingewiesen, dass Unternehmen wie Kodak erst dann begannen, die Empfindlichkeit ihrer Filmemulsionen zu verändern, als sich Schokoladenproduzenten und Holzmöbelhersteller über die Unmöglichkeit beklagten, differenzierte Brauntöne zu reproduzieren – in der Abbildung ihrer Produkte (!). Digitale Bild(komprimierungs-)technologien haben diese Vorurteile geerbt. Infolgedessen



04 Zach Blas & Jemima Wyman, *i'm here to learn so :)))))*, Ausstellungsansicht Institute of Modern Art, Brisbane, Australien, 2017, Courtesy: Zach Blas & Jemima Wyman

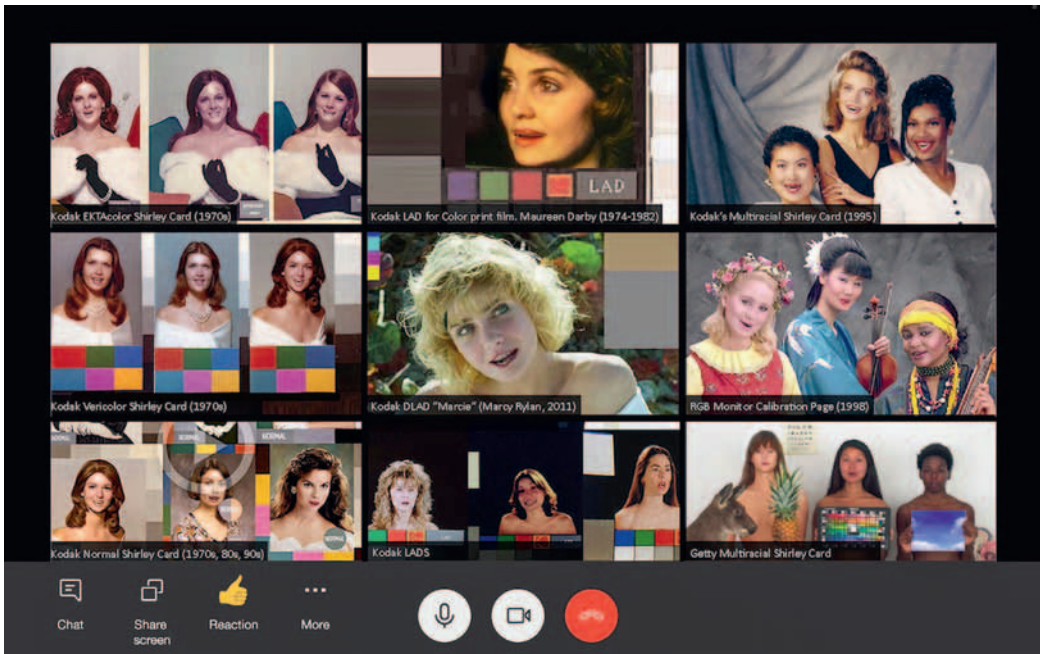
sind selbst moderne Technologien oft rassistisch – insbesondere Gesichtserkennungssysteme erkennen weiße Gesichter besser als alle anderen. So taten sich zum Beispiel HP-Webcams und der X-Box Kinect-Controller von Microsoft 2009 bzw. 2010 schwer, die Gesichter schwarzer oder brauner Benutzer\*innen zu verfolgen, während die iPhone-Gesichtserkennung Nutzer\*innen mit asiatischen Gesichtszügen Sicherheitsfeatures verweigerte. Ein weiterer ‚Höhepunkt‘ folgte, als der Softwareentwickler Jacky Alciné 2015 twitterte, dass Googles neue Foto-App Bilder vom ihm und seiner ebenfalls dunkelhäutigen Freundin automatisch mit „Gorillas“ betitelt hatte.<sup>8</sup>

Bis heute funktioniert maschinelle Gesichtserkennung dann am besten, wenn es um das Erkennen von Gesichtern weißer Männer geht.<sup>9</sup> Das Unvermögen unserer Technologien, andere Hautfarben zu erfassen, ist dabei nicht etwa auf ein technisches Problem (wie ‚schwache Beleuchtung‘) zurückzuführen, sondern auf eine bewusste Entscheidung. Rosa Menkman fordert daher, dass die Datenpools, mit denen die Maschinen trainiert werden, Teil einer öffentlichen Debatte werden müssen: „Diese Bilder müssen ihre schwer greifbare Macht verlieren. Die Geschichte der Standardisierung gehört in die Lehrbücher der Oberstufe, und das Gewaltpotenzial der

Menschen trainieren Maschinen – in diesem Fall einen Chatbot, und diese Maschinen werden nur so gut oder so schlecht sein wie die Menschen, die sie trainiert haben.

Standardisierung sollte auf den Lehrplänen für neue Medien und Kunstgeschichte stehen.<sup>10</sup> [01, 05, 06]

Solange dies noch nicht der Fall ist, nehmen sich Künstler\*innen dieses Problems an. Sie verweisen darauf, dass KI nichts ist, was auf magische Weise selbsttätig agiert, dass KI – trotz der missverständlichen Bezeichnung – nichts ist, was eigenständig „denkt“, oder gar „intelligent“ ist. Die deutsche Künstlerin Hito Steyerl spricht gar von „Künstlicher Dummheit“<sup>11</sup>. KI ist, ganz simpel, Mustererkennung plus Rechenleistung, die das Auffinden von eben solchen Mustern in gewaltigen Datensätzen („Big Data“) ermöglicht. Sie erscheint vielen Menschen als „magisch“, weil zumeist die Ausgangsdatsätze – die „Trainingssets“ – nicht bekannt sind, und auch nicht deren durch Menschen gemachte Annotationen. Und unter anderem hier kommen die Vorurteile ins Spiel.



05



06

„A lofi Angel of History picking up Lena the color test card for the JPEG compression, calling in from 1973, while the different generations of Shirleys (from left to right: Kodak's analogue Shirleys, the LAD Shirleys, and the Multiracial Shirleys) are having a Zoom sing a long on Paul McCartney's We All Stand Together.“ Rosa Menkman, *Pique Nique Pour les Inconnus* (Rupert and the Frogs Chorus/ Paul Mccartney's We All Stand Together version), 2020, Video (Still), 6:42 Min.



05–06 Rosa Menkman, *Pique Nique Pour les Inconnus*. (Rupert and the Frogs Chorus / Paul Mccartney / we all stand together version), 2020, Video (Stills), 6:42 Min. COPY-IT-RIGHT, 2020, Courtesy: Rosa Menkman

Die KI-Forscherin Kate Crawford und der Künstler Trevor Paglen beschäftigen sich mit genau diesen sogenannten „operativen Bildern“<sup>12</sup> (Harun Farocki), die zum Training von Maschinen verwendet werden. Im Gegensatz zu (repräsentativen) Bildern, die auf Bildinhalte abzielen und die von Menschen für Menschen gemacht wurden, enthalten operative Bilder Daten, die sie für Maschinen lesbar machen. Sie dienen dazu, eine Reihe „automatisierter Operationen zu ermöglichen, zum Beispiel Identifizierung, Kontrolle, Visualisierung, Erkennung“<sup>13</sup>. In der Ausstellung *Training Humans* (Fondazione Prada, 2019–20)<sup>14</sup> setzten sich Crawford und Paglen mit verschiedenen Sätzen von ‚Trainingsbildern‘ auseinander, die verwendet werden, um KI-Systemen beizubringen, wie sie die Welt (und darin die Menschen) ‚sehen‘ und klassifizieren sollen. In dem Artikel „Excavating AI“ (2019) schauen sich Crawford und Paglen an, wie die Trainingsbilder in der Kategorie ‚Person‘ in ImageNet<sup>15</sup> gelabelt sind – und was sie finden, ist nicht schön: Ein Foto einer lächelnden Frau im Bikini wird als „Schlampe“, „Luder“, „schlampige Frau“, „Flittchen“ bezeichnet. Ein junger Mann, der Bier trinkt, wird als „Alkoholiker“, „Alki“, „Säufer“, „Säuferin“, „Trinker“, „Trinkerin“, „Schlucker“ eingestuft. Ein Kind, das eine Sonnenbrille trägt, wird als „Versager“, „Verlierer“, „Loser“, „erfolglose Person“ eingestuft. („A photograph of a woman smiling in a bikini is labeled a ‚slattern, slut, slovenly woman, trollop.‘ A young man drinking beer is categorized as an ‚alcoholic, alky, dipsomaniac, boozier, lush, soaker, souse.‘ A child wearing sunglasses is classified as a ‚failure, loser, non-starter, unsuccessful person.“<sup>16</sup>) Diese Annotationen, die keine neutralen Beschreibungen, sondern von Rassismus, Frauenfeindlichkeit, Klassismus, Ableismus und Sexismus durchzogene persönliche (Vor-)Urteile sind, wurden von einer Armee von Akkordarbeiter\*innen verfasst, die via Amazon Mechanical Turk durchschnittlich 50 Bilder pro Minute labeln und in Tausende von Kategorien sortieren mussten. ImageNet umfasst als „Canonical Training Set“<sup>17</sup> 14 Millionen mit Labels annotierte Bilder, die mit Hilfe der Suchmaschine Google im Internet und den Sozialen Medien ‚geerntet‘ („harvested“) wurden, und welche in mehr als 20.000 Kategorien aufgeteilt sind. Je tiefer man in die Hauptkategorie ‚Person‘ eintaucht, desto düsterer werden die Klassifikationen: „Es gibt Kategorien für Schlechte Person, Callgirl, Drogensüchtige, Tunte, Sträfling, Verrückter, Versager, Flop, Ficker, Heuchler, Isebel, Kleptomantin, Verlierer, Melancholiker, Unperson, Perverser, Primadonna, Schizophrener, Besserwisser, Spinner, Straßenmädchen, Hengst, Wichser, Unbegabte, Wankelmütige und Weichei. Es gibt viele rassistische Verunglimpfungen und frauenfeindliche Begriffe.“ („There are categories for

Bad Person, Call Girl, Drug Addict, Closet Queen, Convict, Crazy, Failure, Flop, Fucker, Hypocrite, Jezebel, Kleptomaniac, Loser, Melancholic, Nonperson, Pervert, Prima Donna, Schizophrenic, Second-Rater, Spinster, Streetwalker, Stud, Tosser, Unskilled Person, Wanton, Waverer, and Wimp. There are many racist slurs and misogynistic terms.“<sup>18</sup>) [07–09]

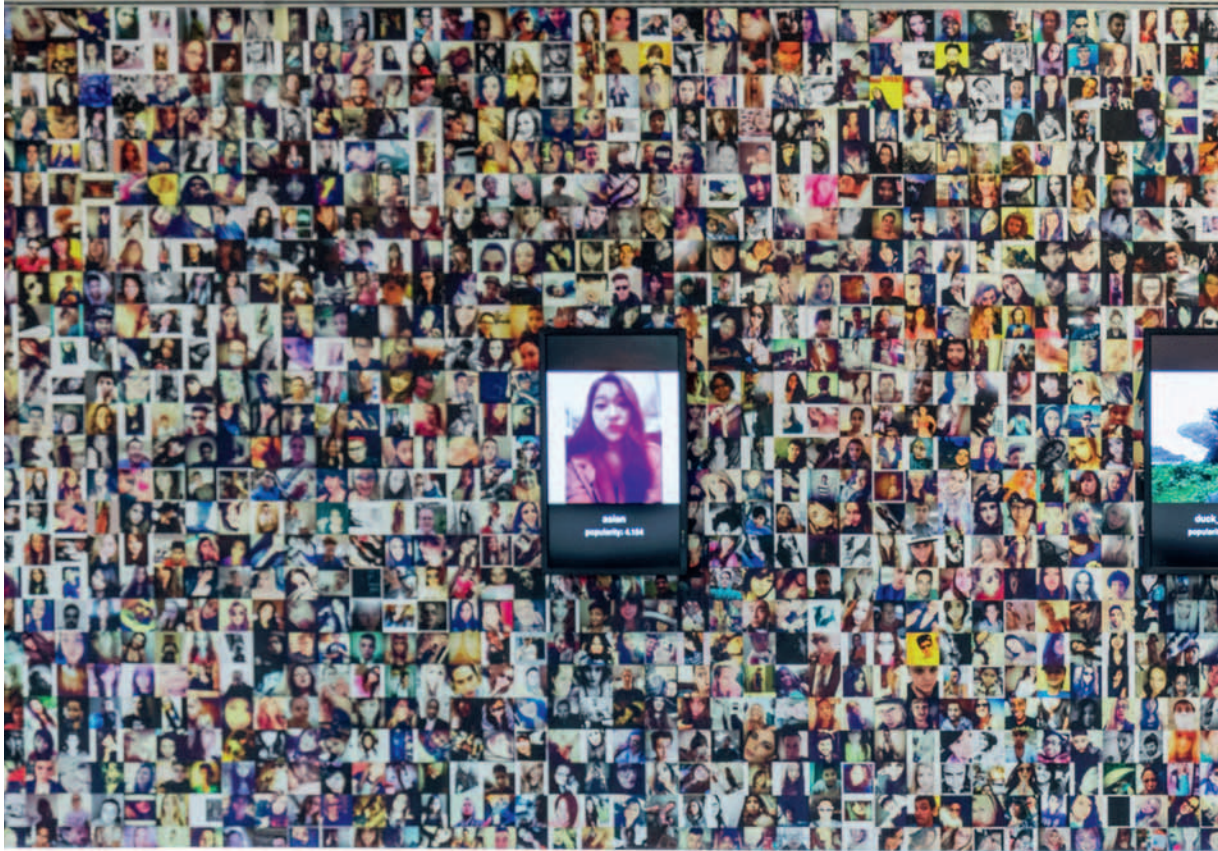
Künstliche Intelligenz hat also mit folgenden Problemen zu kämpfen: a) ist die Auswahl der Trainingsdatensätze oft lückenhaft bzw. zeichnet sich durch mangelnde Diversität aus (nur Gesichter weißer Männer, nur Daten aus dem Globalen Norden, etc.), und b) sind die Annotationen (z.B. im Fall von Bildern von menschlichen Gesichtern oder Körpern) zum Teil rassistisch und mit Vorurteilen beladen. Es gibt nicht so etwas wie einen objektiven oder ‚neutralen Algorithmus‘: Künstliche Intelligenz wird immer die Wertvorstellungen ihrer Schöpfer\*innen reflektieren.

## KÜNSTLERISCHE TAKTIKEN

Bewusstmachung, Öffnen und Unterlaufen der Black Box KI

Und es geht noch weiter: Wir können uns fast gar nicht dagegen wehren, von dieser entfesselten Megamaschine – die eine Black Box ist, wie wir gesehen haben – erfasst zu werden. Hier kommt die Problematik der automatisierten Gesichtserkennung ins Spiel: „Wer in London Apple Pay auf dem iPhone X aktiviert, konnte schon 2018 fast alles mit dem Gesicht bezahlen: U-Bahn-Fahrten, die Pizza-Lieferung nach Hause oder den Wochenendeinkauf im Supermarkt.“<sup>19</sup> An Flughäfen kann man die Grenzkontrolle an Schaltern passieren, die mit automatisierter Gesichtserkennung ausgestattet sind. Doch gleichzeitig stößt diese ‚Facial Recognition‘ auf immer größeren Widerstand. Die britische Menschenrechtsorganisation Liberty zum Beispiel klagte gegen die polizeiliche Nutzung von automatisierter Gesichtserkennung.<sup>20</sup> In Amerika formulierten Mitarbeiter\*innen von Amazon und Microsoft im Sommer 2018 Protestbriefe, um den Verkauf von Gesichtserkennungssoftware an US-Behörden zu stoppen.<sup>21</sup> Und auch in Deutschland wendet sich seit 2020 ein Bündnis aus zivilgesellschaftlichen Organisationen gegen den Plan des Innenministeriums, an 135 Bahnhöfen und 14 Flughäfen automatisierte Gesichtserkennung einzusetzen.<sup>22</sup>

Künstler\*innen haben sich bereits sehr früh mit Technologien der automatisierten Gesichtserkennung auseinandergesetzt. So entwickelte der Künstler und Ingenieur Adam Harvey ab 2011 mit dem recht bekannt gewordenen Projekt *CV Dazzle*<sup>23</sup> eine Make-up-Technik – bestehend aus kubistischen Formen, die sich über charakteristische Gesichtsmarkierungen legen –, die verhindert, dass Gesichtserkennungsalgorithmen auf biometrische Profile



07 Kate Crawford und Trevor Paglen, *Training Humans*,  
Ausstellungsansicht Osservatorio Fondazione Prada,  
Mailand, 2019/2020, Foto: Marco Cappeletti, Courtesy:  
Fondazione Prada — siehe auch S. 128







08

09



08–09 Kate Crawford und Trevor Paglen, *Training Humans*, Ausstellungsansicht Osservatorio Fondazione Prada, Mailand, 2019/2020, Foto: Marco Cappeletti, Courtesy: Fondazione Prada — siehe auch S. 128



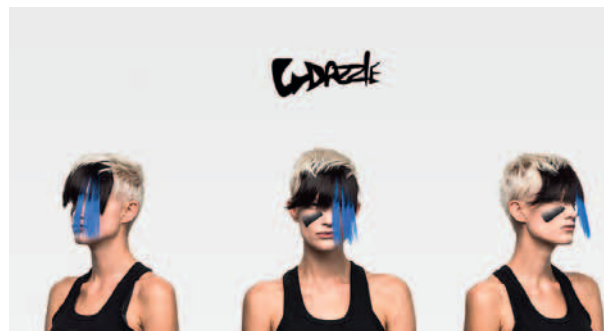
10 Adam Harvey und Jules LaPlace, *MegaPixels: Face Recognition Training Datasets* (heute: *exposing.ai*), 2019, <https://ahprojects.com/megapixels>



zugreifen können. Sein aktuelles Projekt *MegaPixels: Face Recognition Training Datasets*<sup>24</sup> (2017–2020, zusammen mit Jules LaPlace), heute *Exposing.ai* (seit 2021, ebenfalls mit LaPlace), basiert auf jahrelanger Forschung zu Bild-Trainingsdatensätzen, die für Gesichtserkennung und verwandte biometrische Analysetechnologien verwendet werden. Nach dem Aufspüren und Analysieren Hunderter dieser Datensätze kristallisierte sich ein Muster heraus: Millionen von Bildern wurden von Flickr.com heruntergeladen, wo biometrische Daten im Überfluss vorhanden sind. *Exposing.ai* ist eine Suchmaschine, mit der Nutzer\*innen überprüfen können, ob ihre Flickr-Fotos in Dutzenden der am häufigsten verwendeten öffentlichen Gesichts- und biometrischen Bilddatensätze verwendet wurden, die für das Training dieser Systeme eingesetzt werden: „Wenn Sie Flickr.com-Nutzer sind und zwischen 2004 und 2020 Fotos hochgeladen haben, die Gesichter oder andere biometrische Informationen enthalten, wurden Ihre Fotos möglicherweise zum Trainieren, Testen oder Verbessern von Überwachungstechnologien mit künstlicher Intelligenz für den Einsatz in akademischen, kommerziellen oder verteidigungsbezogenen Anwendungen verwendet.“<sup>25</sup> – so wird eine der FAQs auf der Projektwebsite beantwortet. [10, 11]

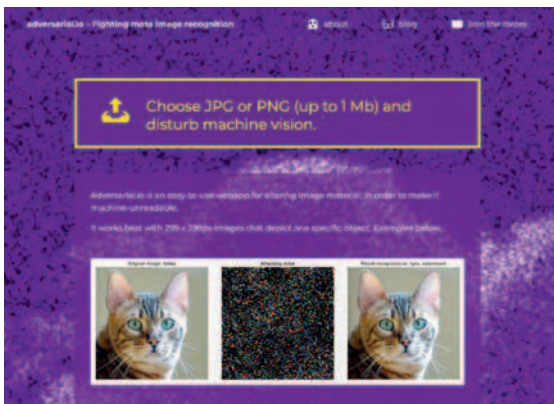
Und, so könnte man anfügen, Ihr Flickr-Foto wurde wahrscheinlich von einem oder einer prekär beschäftigten Mechanical Turk Mitarbeiter\*in annotiert, also für die Maschine lesbar gemacht. Elisa Giardina Papa konzentriert sich in ihrer dreikanaligen Videoinstallation *Cleaning Emotional Data* (2020)<sup>26</sup> auf solche neuen Formen unsichtbarer, prekärer, entfremdeter, schlecht bezahlter und ausgelagerter

Arbeit, die in den Ökonomien der KI entstehen, und in denen die Künstlerin im Winter 2019 selbst arbeitete. Diese Mechanical Turk-Jobs, die Amazon selbst ganz unironisch als „artificial artificial intelligence“ bezeichnet, sind mit der Kategorisierung und ‚Säuberung‘ riesiger Mengen visueller Daten beschäftigt, welche für das Training von Algorithmen zur Emotionserkennung verwendet werden. Zu Elisa Giardina Papas Aufgaben gehörte die Taxonomisierung von Emotionen, die Annotation von Gesichtsausdrücken und die Aufnahme ihres eigenen Gesichtes, um dreidimensionale Charaktere zu animieren. *Cleaning Emotional Data* dokumentiert diese Mikroaufgaben und zeichnet gleichzeitig eine Geschichte der Emotionen nach, die die Methoden und psychologischen Theorien hinterfragt, welche der Erfassung von Gesichtsausdrücken zugrunde



11 Adam Harvey, *CV Dazzle*, *CV Dazzle Look 5*, im Auftrag für *New York Times Op-Art*, 2014, <https://ahprojects.com/cvdazzle>





12

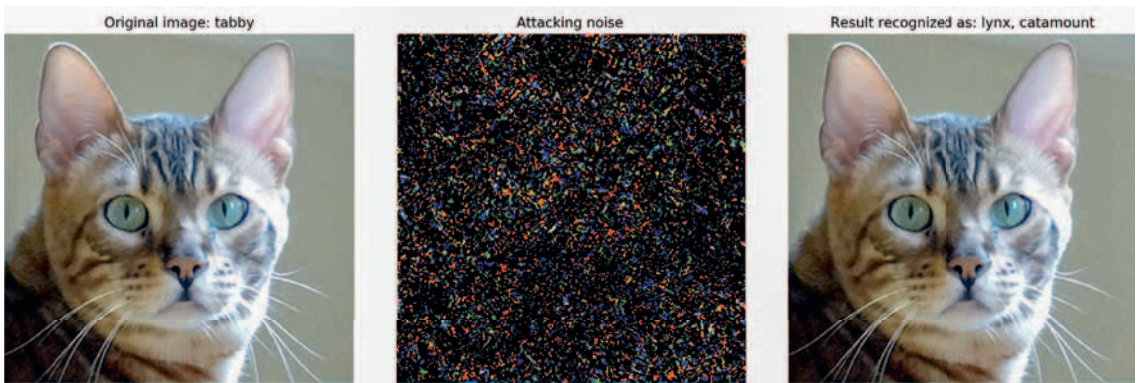
liegen. ‚Emotionale Lesbarkeit‘ wird zunehmend eingesetzt, um entweder die Stimmungen von Verbraucher\*innen zu identifizieren oder potenziell gefährliche Bürger\*innen aufzuspüren, die eine Bedrohung für den Staat darstellen könnten. [14–15]

Ganz anders der Fokus und die Taktik, quasi fast eine Handlungsanweisung, die Francis Hunger und Flupke entwickelt haben: In dem Projekt *adversarial.io* (2020)<sup>27</sup> geht es nicht um das Auffinden von (eigenen) Gesichtern in Trainingsätzen, sondern um das Einfügen von Rauschen in Bilddateien, bevor diese ins Netz gestellt werden: *adversarial.io* ist eine Web-App, mit deren Hilfe man Bilder dergestalt verändern kann, dass sie unlesbar für Maschinen werden, die das Convolutional Neural Network (das künstlich neuronale Netzwerk) Google Inception V.3 zur Bildanalyse und Objekterkennung verwenden. Lädt man z.B. das Bild einer Katze hoch, kalkuliert der Algorithmus von *adversarial.io* ein Noise-Muster, das die Beschreibungsklasse des Bildes („cat“) in die nächste Klasse verschiebt („lynx“). Dieses ‚gegnerische‘ Rauschen ist eine ganz leichte Veränderung des Bildes, die aber das, was die Maschine wahrnimmt, über einen gewissen Schwellenwert schiebt und so quasi

in einer anderen Beschreibung resultiert. Während das Maschinelle Sehen überlistet wird, bleibt die Veränderung des Bildes für das menschliche Auge unsichtbar. Wenn Nutzer\*innen alle ihre Bilder, bevor sie diese im Netz und/oder in den ‚Sozialen Medien‘ verwenden wollen, durch *adversarial.io* schicken, und erst dann, also in veränderter Form, hochladen würden, wären die Bilder für jegliche automatisierte Bilderkennung mittels Google Inception V.3 wertlos. [12–13]

Die Künstlerin Erica Scourti wiederum hat bereits recht früh in einem Selbstversuch nachgewiesen, wie wertend und sexistisch die Ergebnisse von Suchmaschinen sind. Für das Video *Body Scan* (2014) nahm Scourti mit ihrem iPhone Bilder ihres Körpers auf und ließ diese durch verschiedene Suchmaschinen und Apps laufen, die diese Bilder mit Informationen im Netz zu korrelieren versuchten. Scourti interessierte sich dabei für die in der Bildsuche – wie in vielen algorithmischen Prozessen – codierten normativen Aspekte. Trocken kommentiert die Künstlerin im Video die teils witzigen bis sexistischen Suchergebnisse: Natürlich macht der Algorithmus bei Bildern weiblicher Körperteile (v.a. der Brüste) immer Vorschläge, wie diese zu verbessern (v.a. zu vergrößern) seien. *Body Scan* macht die Objektifizierung und die Standardisierung des weiblichen Körpers sichtbar. Die Arbeit ist intim und autobiografisch, verweist aber gleichzeitig auch auf größere gesellschaftliche Kräfte und technologische Entwicklungen.<sup>28</sup> [14]

Zwei Künstlerinnen, die auf weiße Flecken oder fehlende Daten im Bereich der Künstlichen Intelligenz hinweisen, sind Mimi Onuoha und Nora Al-Badri. Mimi Onuoha sammelt in *The Library of Missing Datasets* (2016) und *The Library of Missing Datasets v2.0* (2018) „fehlende Datensätze“.<sup>29</sup> Das Projekt ist ein physisches Archiv für die Leerstellen, die in ansonsten datengesättigten Räumen um uns herum existieren. Daten werden, so Onuoha, aus verschiedenen Gründen nicht gesammelt, z.B.) wenn es auf der Seite derjenigen, die für die Sammlung zuständig



12–13 Francis Hunger und Flupke, *adversarial.io*, 2020



14

sind, kein explizites Interesse gibt (z.B. Daten über Polizeigewalt bei den Behörden), b) wenn die Daten, die gesammelt werden sollen, sich der einfachen Quantifizierung widersetzen (z.B. existieren keine Daten darüber, wieviel Bargeld (US\$) außerhalb der Vereinigten Staaten zirkuliert; ebenso schwierig ist es, Emotionen oder institutionellen Rassismus zu quantifizieren), c) wenn der Akt des Sammelns mit mehr Arbeit verbunden ist als der Nutzen, den das Vorhandensein der Daten vermeintlich bringt (z.B. Daten zu sexuellen Übergriffen und Belästigungen) und d) wenn das Nichtvorhandensein von Daten von Vorteil ist: „Jeder fehlende Datensatz ist ein Beweis für diese Tatsache. So wie das Vorhandensein von Daten jemandem nützt, so tut es auch das Fehlen. Es ist wichtig, dies im Hinterkopf zu behalten.“ In manchen Fällen kann jedoch, so Onuoha, das Nichtvorhandensein von Daten auch ein Schutz sein.

Nora Al-Badri und Jan Nikolai Nelles haben auf der Grundlage ihres (weithin bekannt gewordenen) ‚illegalen‘ 3D-Scans der berühmten Nofretete<sup>30</sup> im Neuen Museum in Berlin den *NefertitiBot*<sup>31</sup> (2018) programmiert. Dieser auf dem\*der persönlichen Assistent\*in Susi AI<sup>32</sup> basierende Chatbot wurde mit von den Künstler\*innen geskripteten Dialogen ausgestattet, die im Chat mit den Nutzer\*innen dekoloniale Antworten liefern. Er kann als utopische Stimme der Subalternen beschrieben werden, die sich dem kolonialen (nach wie vor dominanten) Museums-Narrativ widersetzt. *NefertitiBot* ist ein experimenteller Bot, durch den, so die Künstler\*innen,

„materielle Objekte anderer Kulturen in Museen des Globalen Nordens für sich selbst zu sprechen beginnen, indem sie die gewalttätige koloniale Patina abschütteln, die den institutionellen Narrativen innewohnende Fiktion dekonstruieren und die Politik der Repräsentation herausfordern.“ Und weiter heißt es in der Projektbeschreibung von Nora Al-Badri und Jan Nikolai Nelles: „Sobald Objekte – von kolonialen und umstrittenen Sammlungen – für sich selbst zu sprechen beginnen und Maschinen beginnen, Vorurteile zu transzendieren, könnte uns das bis ins Mark treffen ...“<sup>33</sup> [15, s.S. XXX]



14–15 [oben u. unten](#): Elisa Giardina Papa, *Cleaning Emotional Data*, 2020, Ausstellungsansicht *Algotaylorism*, Kunsthalle Mulhouse, Frankreich, 2020, Courtesy: Elisa Giardina Papa, Foto: Sébastien Bozon, Elisa Giardina Papa, *Cleaning Emotional Data*, 2020, Videostill, Courtesy: Elisa Giardina Papa

Bis dahin ist es jedoch noch ein langer Weg. Viele Künstler\*innen arbeiten daran, die Black Box der Künstlichen Intelligenz sichtbar zu machen und diese im besten Fall zu öffnen und unter die Motorhaube zu schauen. Sie verweisen auf die mangelnde Vielfalt in den Trainingsdaten, die zu verzerrten Ergebnissen führen, welche aber oft – da man von KI als einer ‚objektiven‘ Instanz ausgeht – nicht als solche wahrgenommen werden. Künstler\*innen machen diese mangelnde Diversität sichtbar. Sie verweisen auch auf gelernte Verzerrungen und Vorurteile in der Gesichts- und Mustererkennung, indem sie auf rassistische und mit Vorurteilen beladene, von Menschen gemachte Annotationen hinweisen. Solange es keinen objektiven, neutralen Datenpool gibt, mit dem unsere KIs trainiert werden, wird KI durch automatisierte Diskriminierung und einprogrammierte Vorurteile immer das partielle Weltbild ihrer Schöpfer\*innen reflektieren.

Die Geschichte von Tay sollte uns allen eine Warnung sein: Man muss den Input für Künstliche Intelligenz sehr genau kontrollieren, sonst kommen unten dumme kleine Nazis raus. Oder die KI verweigert Ihnen eine lebenswichtige Nierentransplantation.<sup>34</sup> Warum? Einfach weil Sie die falsche Hautfarbe haben. Denn KI verstärkt existierende Ungleichheiten. In diesem Fall erkennt das System in den US-amerikanischen Gesundheitsdaten das Muster einer kürzeren Lebenserwartung schwarzer Patienten (das auf einer schlechteren Gesundheitsversorgung dieses Bevölkerungsteils der USA beruht) – und investiert die Spenderniere lieber in den Patienten mit einer höheren Lebenserwartung.

INKE ARNS



Direktorin des HMKV in Dortmund. Seit 1993 freie Kuratorin und Autorin mit den Schwerpunkten Medienkunst und -theorie, Netzkulturen, Osteuropa. Nach einem Aufenthalt in Paris (1982–1986) Studium der Slawistik, Osteuropastudien, Politikwissenschaften und Kunstgeschichte in Berlin und Amsterdam (1988–1996), 2004 Promotion an der HU Berlin. Sie kuratiert(e) viele Ausstellungen im In- und Ausland – u. a. am bauhaus Dessau, MG+MSUM Ljubljana, KW Berlin, Ujazdowski Castle Warschau, HMKV Dortmund, HKW Berlin, Muzeum Sztuki Lodz, BOZAR Brüssel – und ist Autorin und Herausgeberin zahlreicher Beiträge zur Medienkunst und Netzkultur. 2021 Gastprofessur für kuratorische Praxis an der Kunstakademie Münster.

Inke Arns, Kulturkonferenz Ruhr, 2019, © Rubert Oberhäuser/Regionalverband Ruhr (RVR).

## ANMERKUNGEN

1 <https://web.archive.org/web/20160414074049/https://www.tay.ai/>, abgerufen am 11. April 2021.

2 Es wird nichts wirklich ‚gelernt‘, sondern es werden vor allem statistische Korrelationen erzeugt.

3 <https://de.wikipedia.org/wiki/DeepDream>, abgerufen am 11. April 2021.

4 N. Katherine Hayles formuliert es so: „(...) the system can know the world only through the modalities dictated by its designer. Although it might work on these data to create new results, the scope of novelty is limited by having its theater of operations – the data that create and circumscribe its world – determined in advance without the possibility of free innovation“ (N. Katherine Hayles, „Computing the Human“, in: *Theory, Culture & Society* 22, 2005, Nr. 1, S. 131–151, hier: S. 137), <https://doi.org/10.1177/0263276405048438>, abgerufen am 11. April 2021.

5 Rosa Menkman, „Hinter weißen Schatten“, in: *Computer Grrrrs*, hrsg. v. Inke Arns, Marie Lechner, Dortmund: Kettler, 2021, S. 32–35.

6 Menkman 2021, S. 32.

7 Lorna Roth, „Looking at Shirley, the Ultimate Norm: Colour Balance, Image Technologies, and Cognitive Equity“, in: *Canadian Journal of Communication*, Jg. 34, Bd. 1, März 2009, S. 111–136. Siehe hierzu auch das Video *Color film was built for white people. Here's what it did to dark skin* (2015), abgerufen am 11. März 2021 <https://www.vox.com/2015/9/18/9348821/photography-race-bias>.

8 Vgl. dazu auch Frederike Kaltheuner, Nele Obermüller, *DatenGerechtigkeit*, Berlin: Nicolai Publishing & Intelligence GmbH, 2018.

9 Vgl. Frederike Kaltheuner, Nele Obermüller, „Diskriminierende Gesichtserkennung: Ich sehe was, was du nicht bist“, in: *Netzpolitik*, 10. November 2018, abgerufen am 28. März 2021, <https://netzpolitik.org/2018/diskriminierende-gesichtserkennung-ich-sehe-was-was-du-nicht-bist/>.

10 Menkman 2021, S. 34.

11 Hito Steyerl, in: „Die Autonomie der Bilder oder Dass Bilder töten können, wussten wir schon immer, aber jetzt sind sie selbst am Abzug“, Hito Steyerl und Trevor Paglen, Vortrag im Centre Pompidou, 10. November 2018, in: Hito Steyerl: *I Will Survive*, hrsg. v. Florian Ebner, Susanne Gaensheimer, Doris Krystof, Marcella Lista, Leipzig: Spector Books, 2020, S. 229–241, hier S. 232.

12 Der deutsche Filmemacher Harun Farocki (1944–2014) prägte den Begriff der „operativen Bilder“ 2003. Vgl. Harun Farocki, „Der Krieg findet immer einen Ausweg“, in: *Cinema 50. Essay*, Marburg: Schüren Verlag, 2005, S. 21–33.

13 Francis Hunger, „Working Paper 2: Computer Vision und die Bilddatensammlung ImageNet in Anwendung auf operative, historische Bilder“, im Rahmen des Forschungsprojekts *Training the Archive*, Ludwig Forum Aachen und HMKV Hartware MedienKunstVerein, Dortmund, 2021. Hunger verweist auf Andreas Broeckmann, *Machine Art in the Twentieth Century*, Cambridge, Mass.: The MIT Press, 2016, hier insbesondere das Kapitel „Operational Images“, S. 128–134.

14 <http://www.fondazioneprada.org/project/training-humans/?lang=en>, abgerufen am 11. April 2021.

15 ImageNet ist einer der in den letzten zehn Jahren am häufigsten verwendeten Trainingssätze für maschinelles Lernen, vgl. <http://www.image-net.org/>, abgerufen am 11. April 2021.

16 Kate Crawford und Trevor Paglen, „Excavating AI: The Politics of Training Sets for Machine Learning“ (September 19, 2019), <https://excavating.ai>, abgerufen am 11. April 2021.

17 Crawford und Paglen, 2019.

18 Crawford und Paglen, 2019. Aufgrund der massiven Kritik von unterschiedlichen Seiten wurde das Trainingsset ImageNet inzwischen zurückgezogen und überarbeitet, und diese Kategorien wurden entfernt. Dies zeigt, dass Kritik also durchaus zu Veränderungen führen kann. Vgl. „An Update to the ImageNet



14 Erica Scourti, *Body Scan*, 2014, im Rahmen der Ausstellung *Computer Grrrls*, kuratiert von Inke Arns und Marie Lechner, HMKV Hartware MedienKunstVerein, Dortmund, 2018/2019, Foto: Hannes Woidich



15 Nora Al-Badri und Jan Nikolai Nelles, *NefertitiBot*, 2018, Foto: Jonas Blume

Website and Dataset", 11. März 2021, <http://www.image-net.org/update-mar-11-2021.php>, abgerufen am 11. April 2021.

Außerdem wurde eine neue Version publiziert, in der die Gesichter abgebildeter Personen mit einem Unschärfe-Filter unkenntlich gemacht wurden. Vgl. Will Knight, „Researchers Blur Faces That Launched a Thousand Algorithms“, *Wired*, 15. März 2021, <https://www.wired.com/story/researchers-blur-faces-launched-thousand-algorithms/>, abgerufen am 11. April 2021.

19 Kalthener und Obermüller, 2018.

20 „Resist Facial Recognition“, <https://www.libertyhuman-rights.org.uk/campaign/resist-facial-recognition/>, abgerufen am 1. April 2021.

21 Vgl. Kalthener und Obermüller, 2018.

22 „Gesichtserkennung stoppen“ fordert ein Verbot dieser hochproblematischen Technologie in Deutschland, <https://digitalcourage.de/blog/2020/buendnis-fordert-verbot-von-gesichtserkennung>, abgerufen am 11. April 2021.

23 <https://cvdazzle.com/>, abgerufen am 11. April 2021.

24 <https://megapixels.cc/>, abgerufen am 11. April 2021.

25 <https://exposing.ai/about/faq/>, abgerufen am 11. April 2021.

26 *Cleaning Emotional Data (2020)* ist zusammen mit *Technologies of Care (2016)* und *Labor of Sleep (2017)* Teil einer

Trilogie, in der sich Elisa Giardina Papa sich mit der Neudefinition von Arbeit und Care-Work durch digitale Ökonomien und Automatisierung beschäftigt. Vgl. <http://www.elisagiardinapapa.org/>, abgerufen am 11. April 2021.

27 <https://adversarial.io/>, abgerufen am 11. April 2021.

28 Vgl. *Computer Grrrls*, hrsg. v. Inke Arns und Marie Lechner, Dortmund: Kettler, 2021.

29 <https://github.com/MimiOnuoha/missing-datasets>, abgerufen am 11. April 2021.

30 *The Other Nefertiti*, 2015, Intervention, 3D-Druck, Video, 17 Min., <https://www.nora-al-badri.de/works-index>, abgerufen am 11. April 2021.

31 <http://nefertiti-chat.surge.sh/>, abgerufen am 11. April 2021, zum Starten „dream nefertiti“ eingeben.

32 Susi AI ist ein „intelligent Open Source personal assistant“, vgl. <https://dev.susi.ai/>, abgerufen am 11. April 2021.

33 Vgl. dazu auch das Projekt *Babylonian Vision (2020)* von Nora Al-Badri, <https://www.nora-al-badri.de/works-index>, abgerufen am 11. April 2021.

34 „How an Algorithm Blocked Kidney Transplants to Black Patients“, in: *Wired*, 26. Oktober 2020, abgerufen am 1. April 2021, <https://www.wired.com/story/how-algorithm-blocked-kidney-transplants-black-patients/>.