

Schriften zur Medienpädagogik 55

## Medienbildung für alle

**Digitalisierung. Teilhabe. Vielfalt.**

Marion Brüggemann  
Sabine Eder  
Angela Tillmann(Hrsg.)

## **Schriften zur Medienpädagogik 55**

Dem Bundesministerium für Familie, Senioren, Frauen und Jugend danken wir für die Förderung des vorliegenden Bandes.

### **Herausgeber**

Gesellschaft für Medienpädagogik und Kommunikationskultur in der Bundesrepublik Deutschland (GMK) e.V.

### **Anschrift**

GMK-Geschäftsstelle  
Oberstr. 24a  
D-33602 Bielefeld  
Fon: 0521.67788  
Fax: 0521.67729  
Email: [gmk@medienpaed.de](mailto:gmk@medienpaed.de)  
Website: [www.gmk-net.de](http://www.gmk-net.de)

Für namentlich gekennzeichnete Beiträge sind die Autorinnen und Autoren verantwortlich.

Redaktion: Marion Brüggemann, Sabine Eder, Angela Tillmann, Tanja Kalwar

Lektorat: Tanja Kalwar

Einbandgestaltung und Titelillustration: Katharina Künkel

Druck: Memminger MedienCentrum, Memmingen

© kopaed 2019  
Arnulfstraße 205  
80634 München  
Fon: 089.68890098  
Fax: 089.6891912  
Email: [info@kopaed.de](mailto:info@kopaed.de)  
Website: [www.kopaed.de](http://www.kopaed.de)

ISBN 978-3-86736-555-0  
e-ISBN 978-3-86736-643-4

---

Lorenz Matzat  
**„Algorithmic Accountability“**  
Automatisierte Entscheidungen sichtbar machen

---

*Im Zeitalter der Digitalisierung werden immer mehr Alltagsprozesse durch Algorithmen in Form von Software geregelt. Meist geht um die Entlastung von eintöniger Arbeit, aber auch um die Vorwegnahmen und Vorbereitung von Entscheidungen. Die Folgen dieser Automatisierung bleiben dabei oft im Verborgenen, obwohl sie nicht zuletzt auch die Teilhabe an Gesellschaft betreffen. Die journalistische Methode „Algorithmic Accountability“ tritt deshalb für eine Rechenschaftspflicht ein: So soll nachvollziehbar werden, welche Entscheidungen wie durch Algorithmen getroffen oder geprägt werden.*

Sensoren messen den Verkehr und regeln ihn per Ampeln. Computerprogramme entscheiden in Sekundenbruchteilen, ob sie Aktien verkaufen oder nicht: Der Alltag unserer Gesellschaft ist durchsetzt von automatischen Systemen, die viele kleine Entscheidungen umsetzen und uns dadurch von Arbeit entlasten. Große Aufmerksamkeit erfährt derzeit das Thema „Künstliche Intelligenz“ (KI). Der Begriff wurde vor über 60 Jahren geprägt und hat seitdem schon einige Hypezyklen durchlaufen. In den 80er-Jahren des 20. Jahrhundert wurden damit beispielsweise „Expertensysteme“ beschrieben (vgl. Manhart 2018). Manche meinen, dass der jüngste Zyklus, der maßgeblich auf Fortschritten von „Machine Learning“ und „Neuronalen Netzen“ beruht, bereits sein Potential ausgeschöpft hat und der nächste „KI-Winter“ bevorsteht. Außerhalb der Fachwelt wird KI sehr unscharf verwendet; von mancher Seite werden ihr quasi-magische Fähigkeiten zugesprochen. Bis vor kurzem wurde beispielsweise noch davon ausgegangen, dass selbstfahrende Autos mittels KI kurz vor dem Durchbruch stünden. Mittlerweile ist unklar, wann dies möglich sein wird (vgl. <http://rodneYROOKS.com/predictions-scorecard-2019-january-01/> [Stand: 14.05.2019]). Seit der Ende 2018 veröffentlichten KI-Strategie der Bundesregierung ([www.ki-strategie-deutschland.de](http://www.ki-strategie-deutschland.de)) wird unter dem Begriff KI nun nahezu alles gefasst, was mit der „Digitalisierung“ zu tun hat. Inwiefern Maschinen tatsächlich „Intelligenz“ an den Tag legen können, die menschlicher Autonomie und Internationalität entsprechen, ist nach wie vor umstritten (vgl. Brooks 2019).

Aus gesellschaftlicher Sicht ist hinsichtlich der Übertragung von Aufgaben, die einst Menschen besorgten, wesentlich, dass damit Vorbereitungen und auch Umsetzungen von Entscheidungen an Maschinen delegiert werden. Das können hochgradig komplexe neuronale Netze sein oder auch recht simple Softwareanwendungen, die anhand schlichter Regelwerke Daten miteinander verrechnen, diese gewichten und sortieren. Aus meiner Sicht ist der Begriff „Systeme Automatisierter Entscheidungsfindung“ (automated decision-making, ADM) geeigneter, um die neuen Entwicklungen zu beschreiben (vgl. Spielkamp et al.: 2019).

Es gibt vier wesentliche Verfahren, wie per Software Entscheidungen getroffen werden: Priorisieren, Klassifizieren, Herstellung von Zusammenhängen sowie Filterung. Nahezu alle Computerprogramme, die derzeit unter Schlagworten wie Machine Learning im Bereich der KI vermarktet werden, setzen auf eines oder eine Kombination der genannten mathematischen beziehungsweise statistischen Verfahren. Dabei spielen oft große Datenmengen eine Rolle, die es erlauben, Software anhand bestehender Datensätze zu „trainieren“, um zukünftige Anforderungen erfüllen zu können.

### ***Algorithmic Accountability – eine Methode zur Auseinandersetzung mit Automatisierung***

Es ist an der Zeit für eine Auseinandersetzung mit ADM, da in den letzten zehn Jahren die Automatisierungsprozesse durch Software immens zugenommen haben. Sowohl die Menge der Daten, die automatisierte Entscheidungsfindungen häufig erst ermöglicht haben, als auch die Anzahl der Geräte bzw. Infrastruktur auf denen ADM stattfindet, hat exponentiell zugenommen. Damit steigt zwangsläufig auch der Einfluss dieser Systeme in verschiedensten Bereichen der Gesellschaft. Berührt wird dabei immer wieder auch der Bereich von Teilhabe. Wenn Behörden Systeme einsetzen, die z.B. Kriminalität vorhersagen (predictive policing), Studienplätze vergeben, Sozialhilfebetrug aufdecken oder Kindeswohl monitoren sollen, ist das nicht per se abzulehnen. Die Gefahr ist allerdings, dass diese Systeme schlecht gemacht sind und es zu Diskriminierung kommt, sprich Personen aufgrund einer Zugehörigkeit zu einer Gruppe pauschal von der Teilhabe am gesellschaftlichen Leben ferngehalten werden.

In diesem Zusammenhang kommt Algorithmic Accountability ins Spiel. Maßgeblich geprägt hat die Methode der US-amerikanische Journalismusforscher Nicholas Diakopoulos mit seiner Anfang 2014 erschienenen Veröffentlichung „Algorithmic Accountability Reporting: On the Investigation of Black Boxes“ (vgl. Diakopoulos 2014). In diesem Bericht wird

ein neues Aufgabengebiet für den Journalismus skizziert: Rechenverfahren (Algorithmen) wird Rechenschaft (Accountability) abverlangt. Aufgabe des Journalisten ist es demnach, die Funktionsweise von Softwaresystemen mit gesellschaftlicher Relevanz als Recherchegegenstand zu verstehen, um so Transparenz zu schaffen.

Die Methode ist mit dem Datenjournalismus verwandt, der sich einem oder mehreren Datensätzen widmet und diese ins Zentrum der journalistischen Betrachtung stellt; oft unter Einsatz von (interaktiver) Datenvisualisierung. Entgegen dem Datenjournalismus widmet sich Algorithmic Accountability nicht allein dem Datensatz, sondern den Prozessen der Datenverarbeitung. Diese beginnen schon bei der Beauftragung, Konzeption und Entwicklung von Software für einen bestimmten Zweck. Weiter werden Daten, die zum Testen der Anwendung, aber auch im Normalbetrieb Einsatz finden, berücksichtigt. Schließlich wird auch der Betrieb und Einsatz der datenverarbeitenden Software und deren Auswirkungen untersucht: Es geht um das System der Automatisierten Entscheidung. Die weiter unten beschriebene journalistische Untersuchung des COMPAS-Algorithmus gilt als exemplarisch für diesen Ansatz.

### ***Rechenschaft für Rechenverfahren***

Selbstlernende Algorithmen und große Datensätze bilden eine Black Box, die Algorithmic Accountability notwendig macht. Problematisch ist nicht etwa, dass Aufgaben an Maschinen abgetreten werden; dies kann in vielen Bereichen Erleichterungen für alle Beteiligten bringen: Kosten können gespart, fairere Ergebnisse geliefert und menschliche Stimmungen sowie Vorurteile aus Auswahlprozessen herausgehalten werden. Problematisch sind allerdings die Aufgaben oder Regeln, die durch Algorithmen abgearbeitet werden sollen. So können etwa Wertungen, Vorurteile oder moralische Einstellungen gezielt oder unterschwellig in Software für selbstlernende Algorithmen einfließen. Ein Beispiel wäre der Komplex „Dieselgate“, bei dem Entscheider\*innen und Ingenieur\*innen Testverfahren zur Messung von Motorabgasen täuschten. Gäbe es verpflichtende Verfahren für Algorithmic Accountability, wäre solch ein Betrugsversuch nicht so einfach möglich gewesen. Kritiker\*innen fordern deshalb zum Beispiel ein Audit-System, bei dem Softwarecodes und Trainingsdaten überprüft werden sollten, manche wünschen sich gar einen „Algorithmen-TÜV“ (z.B. Wahlprogramm SPD, Bundestagswahl 2017; vgl. Klingel/Lischka 2017).

Noch komplizierter sind Prozesse, bei denen Algorithmen automatisch an die Stelle von menschlichen Entscheidungen treten oder diese we-

sentlich beeinflussen. Zu diesem Bereich des Automated Decision Making (ADM) gehören auch Predictive-Analytics-Verfahren, die menschliches Verhalten voraussagen sollen. Ein Beispiel dafür ist die Software Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), die in vielen Bundesstaaten der USA eingesetzt wird. Sie liefert unter anderem Richtern Einschätzungen darüber, inwiefern eine zu verurteilende Person mit hoher Wahrscheinlichkeit in Zukunft ein Gewaltverbrechen begehen könnte. Das US-amerikanische Recherchebüro Pro Publica hat die Software im vergangenen Jahr untersucht und gezeigt, wie Algorithmen juristische Entscheidungen ungerechter machen können (vgl. Angwin et al. 2016). So erlangte Pro Publica über eine Informationsfreiheitsanfrage die Risikoeinschätzungen zu mehr als 18.000 Personen, die 2013 und 2014 in einem Landkreis von Florida durch COMPAS bewertet wurden. Aber nur etwa 20 Prozent derjenigen, von denen COMPAS vorhersagte, dass sie in Zukunft ein Gewaltverbrechen ausüben würden, begingen es bislang tatsächlich. Auch wenn alle möglichen Kriminalitätsarten und Verstöße – inklusive Fahren ohne Führerschein – hinzugenommen wurden, lagen nur für etwa 60 Prozent Vergehen vor. Außerdem prognostizierte die Software für US-Amerikaner\*innen schwarzer Hautfarbe fälschlicherweise fast doppelt so oft wie für Weiße, dass sie rückfällig werden würden (vgl. ebd.).

Das COMPAS-Beispiel zeigt die umstrittenen möglichen Konsequenzen von Entscheidungsprozessen, die auf ADM-Verfahren basieren. Die Software-Entwickler von COMPAS behaupteten zwar, die Analyse von Pro Publica sei falsch. Ihre Berechnungsmethoden wollte die Firma Northpointe allerdings nicht offenlegen. Schon diese mangelnde Transparenz nährt Zweifel am Einsatz entsprechender Software bei Justizverfahren, die über menschliche Schicksale entscheiden können. So lässt sich auch nicht klären, wie es zur rassistischen Diskriminierung durch algorithmische Verhaltensvoraussagen für Menschen verschiedener Hautfarbe kam.

Das Machine-Bias-Beispiel macht deutlich, warum Datensätze und Algorithmen, auf denen ADM-Systeme basieren, offengelegt beziehungsweise zugänglich gemacht werden müssen.

### ***Wie werden die Entscheidungen sichtbar?***

In Deutschland ist der Einsatz von „Predictive“ Software bei Strafverfahren noch nicht erlaubt. Die Frage ist allerdings, in welchen anderen staatlichen Verwaltungen und Bereichen Software-Entscheidungen vorbereitet werden, die die persönliche Freiheit Einzelner berühren könnten. Zählt die Software, die Hartz IV-Bescheide ausrechnet, schon dazu? Was ist mit Kreditwürdig-

keitsverfahren (Schufa etc.)? Im Frühjahr 2019 haben wir von AlgorithmWatch versucht, eine Bestandsaufnahme davon in einem „Atlas der Automatisierung“ vorzunehmen. Behandelt werden hier Themengebiete wie Gesundheit, Arbeit, Sicherheit und Überwachung sowie Verkehr (vgl. Matzat et al.: 2019).

Ein sinnvoller Schritt ist sicherlich die bereits geforderte Kennzeichnungspflicht für ADM-Prozesse. Die seit Mai 2018 gültige EU-Datenschutzgrundverordnung enthält dahingehende Regelungen, die allerdings als nicht weitgehend genug kritisiert werden: Etwa seien die Ausnahmen des Verbots kompletter Automatisierung recht weit und die Formulierungen zum Anspruch auf Auskunft über den Einsatz von ADM unzureichend. Denn das Wissen darüber, ob ein\*e Bürger\*in einer automatisierten Entscheidung ausgesetzt war, garantiert nicht, dass die Entscheidung rechtsstaatlichen Verfahren genügt oder diskriminierungsfrei ist. Auch eine Offenlegung von Softwarecodes reicht nicht aus. Damit wäre vielleicht einer Transparenz genüge getan, allerdings können Softwaresysteme so komplex sein, dass ohne beispielsweise die Trainingsdatensätze (siehe COMPAS) nicht nachvollzogen werden kann, wie sie funktionieren. Insofern sollte die Forderung bei ADM nicht Transparenz, sondern Nachvollziehbarkeit lauten. Das bedeutet, die Funktionsweise eines Algorithmus muss dokumentiert sein.

Der Vorschlag, dass den Betroffenen bei automatisierten Entscheidungen deren Gründe erläutert werden, weist in die richtige Richtung. Das ist allerdings nicht immer praktikabel: Wenn beispielsweise ein autonomes Fahrzeug „entscheiden“ würde, jemanden zu überfahren, um jemand anderen zu schützen, nutzt eine Erklärung wenig. So wundert es nicht, dass diese Art von ethischen Fragen zu den Konsequenzen des automatisierten Fahrens in jüngster Zeit eine breitere Öffentlichkeit fand. Sie zeigt eindrücklich, dass es beim Thema ADM nicht nur um abstrakte Vorgänge geht, die sich auf einem Bildschirm in sozialen Online-Netzwerken oder in einem Amtsbrief niederschlagen. Vielmehr haben durch Algorithmen automatisierte Verfahren – auch hinsichtlich der zunehmenden Automatisierung von Arbeit – unmittelbare Konsequenzen in der physischen Welt. So setzt das Bundesamt für Migration und Flüchtlinge (BAMF) beispielsweise eine Software ein, um darüber automatisch in Sprachaufnahmen von einer/einem Asylsuchenden anhand des Dialektes zu erkennen, ob sie oder er aus der von ihr/ihm angegebenen Region stammt. Die Software hat laut BAMF einen Fehlerquotient von 20 Prozent („false positive“); sie wird trotzdem weiterhin eingesetzt. Das heißt in rund 20 von 100 Fällen fließt eine falsche automatisierte Einschätzung in die von einem Menschen getroffene Entscheidung über einen Aufenthaltstitel ein (vgl. Biselli 2018).

Bis die Politik für ADM-Prozesse eine geeignete Regulierung und möglicherweise entsprechende Aufsichtsinstitutionen geschaffen hat, bleiben Gesellschaft und Individuum im digitalen Zeitalter auf Algorithmic Accountability angewiesen. Die Wirkungen und Nebenwirkungen von Algorithmen müssen sichtbar und öffentlich diskutiert werden. Algorithmic Accountability kann dabei für die nötige Transparenz, Nachvollziehbarkeit und Rechenschaftspflicht sorgen. Denn problematisch ist nicht, dass Aufgaben an digitale Maschinen abgetreten werden. Das kann in vielen Bereichen Erleichterungen schaffen, Probleme lassen sich teils besser lösen und Kosten können eingespart werden. Auch sind manchmal, wenn menschliche Stimmungen und Vorurteile aus den Prozessen herausgehalten werden, fairere Ergebnisse möglich, so beispielsweise in Bewerbungsprozessen, wo sich hierüber Effekte von offensichtlicher und unterschwelliger Diskriminierung lindern lassen (vgl. Kramer 2018).

Insofern wäre es an der Medienpädagogik, die journalistische Methode Algorithmic Accountability für die eigene Arbeit zu übersetzen. An erster Stelle stünde sicher die Bewusstseinsmachung: Wo begegnen uns in unserem Alltag Algorithmen und automatisierte Entscheidungen? An welcher Stelle bin ich ihnen bereits unterworfen oder von ihnen beeinflusst? Ein weiterer zu reflektierender Aspekt könnte die Funktionsweise von Algorithmen sein: Wie funktioniert das Filtern und Sortieren konkret? Hier bietet es sich an, das Abstrakte erfahrbar zu machen: Etwa ließen sich diese Prozesse haptisch durch Modellbauten mit Lego nachvollziehen oder über Gruppenspiele bearbeiten, in denen Individuen eine vorher gemeinsam festgelegte Entscheidungskette anhand bestimmter Merkmale durchlaufen müssen. Letztendlich geht es um Ansätze, die den/die Einzel\*en dazu befähigen, in der Gesellschaft und im Digitalen zu erkennen, welchen Kräften sie oder er ausgesetzt ist und welche Rechte und Handlungsoptionen sie oder er hat.

## **Literatur**

Angwin, Julia et al. (2016): Machine Bias. There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. Abrufbar unter: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> [Stand: 18.02.2019].

Biselli, Anna (2018): Eine Software des BAMF bringt Menschen in Gefahr. Motherboard Vide Deutschland. Abrufbar unter: <https://motherboard.vice.com/de/article/a3q8wj/fluechtlinge-bamf-sprachanalyse-software-entscheidet-asyl> [Stand: 05.05.2019].



- Brooks, Rodney (2019): Predictions Scorecard, 2019 January 01. Abrufbar unter: <http://rodneybrooks.com/predictions-scorecard-2019-january-01/> [Stand: 05.05.2019].
- Diakopoulos, Nicholas (2014): Algorithmic Accountability Reporting: On the Investigation of Black Boxes. Tow Center. Abrufbar unter: <https://towcenter.columbia.edu/news/algorithmic-accountability-reporting-investigation-black-boxes> [Stand: 18.02.2019].
- Dreyer, Stephan/Schulz, Wolfgang (2018): Was bringt die Datenschutz-Grundverordnung für automatisierte Entscheidungssysteme? Bertelsmann Stiftung. Abrufbar unter: [https://algorithmenethik.de/wp-content/uploads/sites/10/2018/04/BSt\\_DSGVOundADM\\_dt\\_ohneCover.pdf](https://algorithmenethik.de/wp-content/uploads/sites/10/2018/04/BSt_DSGVOundADM_dt_ohneCover.pdf) [Stand: 05.05.2019].
- Klingel, Anita/Lischka, Konrad (2017): Was die Wahlprogramme über Maschinen sagen, die Menschen bewerten. Ethik der Algorithmen. Abrufbar unter: <https://algorithmenethik.de/2017/09/11/was-die-wahlprogramme-ueber-maschinen-sagen-die-menschen-bewerten/> [Stand: 05.05.2019].
- Kramer, Bernd (2018): Interview mit Tim Kramer: „Der Algorithmus diskriminiert nicht“. Zeit Online. Abrufbar unter: <https://www.zeit.de/arbeit/2018-01/roboter-recruiting-bewerbungsgespraech-computer-tim-weitzel-wirtschaftsinformatiker/komplettansicht> [Stand: 05.05.2019].
- Manhart, Klaus (2018): Eine kleine Geschichte der Künstlichen Intelligenz. Computerwoche. Abrufbar unter: <https://www.computerwoche.de/a/eine-kleine-geschichte-der-kuenstlichen-intelligenz,3330537> [Stand 05.05.2019].
- Matzat, Lorenz et al. (2019): Atlas der Automatisierung. Automatisierte Entscheidungen und Teilhabe in Deutschland. AlgorithmWatch. Abrufbar unter: <https://atlas.algorithmwatch.org/> [Stand 05.05.2019].
- Spielkamp, Matthias et al. (2019): Automating Society – Taking Stock of Automated Decision-Making in the EU. AlgorithmWatch. Abrufbar unter: <https://algorithmwatch.org/en/automating-society/> [Stand: 18.02.2019].

## **Lizenz**

Der Artikel steht unter der Creative Commons Lizenz **CC BY-SA 4.0**. Der Name des Urhebers soll bei einer Weiterverwendung genannt werden. Wird das Material mit anderen Materialien zu etwas Neuem verbunden oder verschmolzen, sodass das ursprüngliche Material nicht mehr als solches erkennbar ist und die unterschiedlichen Materialien nicht mehr voneinander zu trennen sind, muss die bearbeitete Fassung bzw. das neue Werk unter derselben Lizenz wie das Original stehen. Details zur Lizenz: <https://creativecommons.org/licenses/by-sa/4.0/legalcode>