

Buchreflexion Cathy O'Neil „Angriff der Algorithmen“

Der englische Titel „Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy“ verrät klarer das durchaus große Anliegen der Autorin.

Näheres zum Buch von Seiten des Verlages erhalten Sie hier: <https://www.hanser-literaturverlage.de/buch/angriff-der-algorithmen/978-3-446-25668-2/>.

Cathy O'Neil kennt die dunkle der Macht, würde man es Star Wars like ausdrücken wollen. Sie war als Mathematikerin bei einem Hedge-Fonds beschäftigt und ist nunmehr Occupy-Aktivistin. Als solche versucht sie der „hellen Seite der Macht“ Vorschub zu leisten, auf der nicht etwa auf Algorithmen verzichtet werden sollte. Vielmehr sollte ein transparenter, klar geregelter, kontrollierter und diskursiver Umgang gepflegt werden. Denn durch Algorithmen werden Menschen digital indiziert und qualifiziert. Dieses ist eine auch ethisch moralische zu begreifende Tatsache, die Cathy O'Neil unter dem Begriff der (Un-) Gerechtigkeit beschreibt und derartige Einordnungen von Personen eingehend in ihrem Buch behandelt.

Häufig hört man in diesem Zusammenhang Argument der Algorithmenanwender, ein mathematischer Algorithmus sei neutraler und damit besser geeignet als Menschen, die durch Fehleinschätzungen und diskriminierenden Vorstellungen „von Natur aus“ schlechter als ein Algorithmus in der Lage seien, Qualifizierungen zu treffen (z.B. bei der Einstellung von Personen). Dieses Argument missachtet, dass auch ein Algorithmus von genauso so fehlerbehafteten Menschen programmiert wird, also deren (fehlerhafte) Annahmen und Vorurteile in die Programmierung des Algorithmus mit einfließen.

Der Begriff im englischen Originaltitel „Weapons of Math Destruction“ (WMD) steht für Algorithmen, die sich rekursiv ausschließlich aus sich selbst speisen und sich damit negativ verstärken. Ihnen fehlt es an aus Sicht des Algorithmus äußeren Lerninformationen, durch die dieser Algorithmus im positiven Sinne lernen könnte. Fehlt dieser Abgleich mit Außeninformationen, werden Selektionen und Entscheidungen, die durch ihn getroffen werden, durch „böartige Feedbackschleifen“ negativ verstärkt und noch unheilvoller: *Anstatt die Wahrheit herauszufinden, wird der Score selbst zur Wahrheit* (Cathy O'Neil „Angriff der Algorithmen“ S. 13, Carl Hanser, München 2017 – ich verwende zukünftig bei Zitaten aus diesem Buch das Sigle AdA).

Beispiele hierfür sind elektronische (e-) Scores (in den USA), die bei einer Internetrecherche für einen Kredit u.a. den Standort des Browsers heranziehen, um Kreditkonditionen entsprechend anzubieten. Aus Sicht von Cathy O'Neil sind die Basis derartige Scores fehlerhafte Daten bzw. Indizien, die nicht hinreichend genau die Faktenlage bestimmen bzw. widerspiegeln. O'Neil schreibt dazu (AdA S. 23):

[...] fehlen den WMD-Entwicklern sehr häufig die Daten für genau die Verhaltensweisen, die sie am meisten interessieren. Also verwenden sie ersatzweise Näherungswerte oder Indikatoren. Sie errechnen statistische Korrelationen zwischen der Postleitzahl oder gewissen Mustern in der Ausdrucksweise einer Person und der Wahrscheinlichkeit, dass sie einen Kredit zurückzahlen oder einen Job gut erledigen wird. Solche Korrelationen sind diskriminierend, und manche von ihnen sind sogar illegal.

Zudem ist die Zielformulierung, die einem Programmierer vorliegt, häufig auf ein unternehmerisches Ziel ausgelegt, wie etwa Risikominimierung, Gewinnmaximierung. Werte sind nicht Element des Briefings und überhaupt auch schwer oder gar nicht? in Algorithmen zu übertragen.

Und Computer, ungeachtet all ihrer Fortschritte bei der Verarbeitung von Sprache und logischen Zusammenhängen, haben nach wie vor große Schwierigkeiten mit Konzepten. Sie »verstehen« Schönheit nur als ein Wort, das verknüpft ist mit dem Grand Canyon, Sonnenuntergängen am Meer und Körperpflegetipps in der Vogue. Sie versuchen vergeblich, »Freundschaft« zu messen, indem sie Dinge wie »Likes« und Verknüpfungen auf Facebook zählen. Und mit dem Konzept »Gerechtigkeit« können sie überhaupt nichts anfangen – die Programmierer wissen nicht, wie sie es in Code umsetzen können, und kaum einer ihrer Chefs verlangt das von ihnen.

Das bedeutet, dass WMDs nicht auf Gerechtigkeit programmiert werden. Und das Ergebnis ist die massenhafte, industrielle Produktion von Ungerechtigkeit. (AdA S.107)

Selbst wenn staatliche Institutionen eine Verbesserung einer gesellschaftlichen Situation erreichen wollen, so wird durch die oben genannten Schwächen das Gegenteil erreicht.

Im Versuch der Qualitätsverbesserung der schulischen Lehre werden in den USA WMDs eingesetzt, die Lehrerscores erzeugen, die ausschließlich über die relativen Notenvergleiche ihrer Schüler im Jahresvergleich konstruiert werden und dabei wichtige Außenkriterien wie z.B. die Änderung der Zusammensetzung, die Anzahl, die Größe der „Stichprobe“ der Schüler, der Vergleich mit Schülern anderer Stadtteile außer Acht lässt.

Allerdings hatte die Zahl der Schüler, die den Test ablegten, im Laufe dieser 17 Jahre stark zugenommen. Die Hochschulen hatten ihre Pforten immer mehr Schülern aus ärmeren Bevölkerungsschichten und ethnischen Minderheiten geöffnet. Immer mehr berufliche Chancen taten sich auf. Das signalisierte gesellschaftlichen Erfolg. Aber natürlich zog dieser Zustrom von Newcomern die Durchschnittsscores nach unten. Als jedoch die Statistiker die untersuchten Personen nach Einkommensgruppen kategorisierten, stiegen die Scores jeder einzelnen Gruppe, bei Armen ebenso wie bei Reichen.

In der Statistik ist dieses Phänomen als »Simpson-Paradoxon« bekannt: wenn ein Datenbestand insgesamt den einen Trend zeigt, sich jedoch, wenn er in Untergruppen niedergebroschen wird, für jede dieser Untergruppen der entgegengesetzte Trend manifestiert. (AdA S.151)

Anstatt also die Leistungen der Lehrer auf einer absoluten Skala zu messen, versuchten sie, in ihrem Modell soziale Ungleichheiten zu kompensieren. Statt Tim Cliffords Schüler mit anderen Schülern in anderen Stadtteilen zu vergleichen, verglichen sie sie mit prognostizierten Modellen von ihnen selbst.

[...]

Wenn 80.000 Achtklässler den Test absolvieren, sollte es doch machbar sein, aus den Ergebnissen zuverlässige Durchschnittswerte für mangelhafte, mittelmäßige und hervorragende Schulen zu ermitteln, oder etwa nicht?

Ja, durchaus. Und wenn Tim Clifford eine große Stichprobe von Schülern unterrichten würde, vielleicht 10.000, könnte es sinnvoll sein, die Ergebnisse dieser Kohorte mit dem Vorjahresdurchschnitt zu vergleichen und daraus Rückschlüsse zu ziehen. Große Stichproben gleichen Ausnahmen und Ausreißer aus. Dann würden sich Trends abzeichnen, zumindest theoretisch. Aber es ist so gut wie unmöglich, dass eine Klasse von 25 oder 30 Schülern der Gesamtheit der Stichprobe entspricht. (AdA S. 152)

Die Stratifizierung, also die Analyse einzelner Sub-Kohorten lag nicht im Blickpunkt der Algorithmusprogrammierer. Diesen Input hätten Statistiker zuliefern müssen. Um der Situation gerecht zu werden, ist es unabdingbar, interdisziplinäres Wissen zu vereinen, um Fehler zu vermeiden, die aus unterstellt gut gemeinten Absichten sogenannte WMD gerieren.

Ich vermute und befürchte, dass vergleichbare Projekte zur Qualitätsverbesserung in Deutschland aufgesetzt werden, bei denen auf Algorithmen vertraut wird, die keine verbessernde Außenquelle nutzen oder Vereinfachungen nutzen, die algorithmisch überhaupt bzw. einfacher abzubilden sind. Das wird den komplexen Situationen nicht gerecht und führt damit nicht zu einer echten Qualitätsverbesserung, sondern im schlimmsten Fall zu einer Umorientierung. Denn diejenigen, die von der algorithmischen „Qualitätsverbesserung“ betroffen sind, werden anstatt sich auf die eigentliche Situation zu konzentrieren, überlegen, die algorithmische „Prüfung“ erfolgreich zu bestehen. Dies wird in den seltensten Fällen zu einer tatsächlichen Qualitätsverbesserung führen. Das Vertrauen in Algorithmenprogrammierung ist ziemlich blind, dazu ist eine interdisziplinäre Herangehensweise teurer, komplexer und zeitaufwändiger.

Eine Unterart – die e-scores – sind ein weiterer Einsatzort von Algorithmen, die zu WMD mutieren, den Cathy O’Neil anspricht. Aufgrund bestehender Datenschutzgesetze sind diese so in Deutschland / in der EU im Vergleich zu Cathy O’Neils Beispielen aus den USA eingeschränkt übertragbar.

Ein in diesem Zusammenhang häufiger Fehler ist die Verwechslung von Kausalität (Ursächlichkeit) und Korrelation (Zusammenhang) (vgl. Ada S.177). Nicht alles was sich ähnlich (bzw. gegensätzlich) entwickelt (korreliert), hängt ursächlich miteinander zusammen. Zudem werden bei e-scores Verallgemeinerungen vorgenommen:

Mit anderen Worten: Die Entwickler von e-Score-Modellen müssen sich damit begnügen, zu versuchen, die folgende Frage zu beantworten: »Wie haben Menschen wie du sich in der Vergangenheit verhalten?« Obwohl sie idealerweise fragen müssten: »Wie hast du dich in der Vergangenheit verhalten?« [...] Ich sollte darauf hinweisen, dass in dem statistischen Universum, in dem solche Indikatoren – oder Näherungswerte – existieren, diese in vielen Fällen funktionieren. [...] Aber kommt es auch vor, dass die Systeme sich bei einer Person irren und sie im falschen »Eimer« landet? Das kann passieren. Und es gibt kein Feedback, welches das System korrigieren könnte. (AdA S. 161)

Genau hier zeigt sich sehr prägnant, was O’Neil als WMD ansieht. Einmal im „falschen“ Eimer, immer im „falschen“ Eimer. E-scores sind darauf ausgerichtet, dass alles so bleibt wie es gewesen ist. Insofern ignorieren e-scores allerdings die Wesensform von Gesellschaft, die sich permanent ändert und in den

USA geradezu zum Gründungsmythos gehört: vom Tellerwäscher¹ zum Millionär, alles ist möglich, wenn Du nur willst. E-scores werden Tellerwäscher immer als Tellerwäscher identifizieren und damit seine Chance verringern, sich positiv verändern zu können. O'Neil sieht darin eine Diskriminierung durch e-scores, die ungerecht ist:

Wenn automatische Systeme unsere Daten durchsieben, um uns für einen e-Score zu taxieren, projizieren sie naturgemäß die Vergangenheit in die Zukunft. Wie wir bei den Rückfallquoten-Strafzumessungs-Modellen und den räuberischen Darlehens-Algorithmen gesehen haben, wird von den Armen erwartet, dass sie für immer arm bleiben, und sie werden entsprechend behandelt – berufliche Chancen werden ihnen verwehrt, sie landen häufiger im Gefängnis und sie werden für diverse Leistungen und Kredite über den Tisch gezogen. Die Diskriminierung ist unerbittlich, in den meisten Fällen verdeckt und unanfechtbar, und sie ist zutiefst ungerecht. (AdA S. 171)

Es gilt dabei zu verhindern, die zu große Bedeutung von Bonitätsscores nach Europa schwappen zu lassen. Auch in Deutschland ist der bekannteste Bonitätsscore ein privatwirtschaftlicher, der hinsichtlich der Erstellung des Scores nur wenig transparent ist: die Schufa ((Schutzgemeinschaft für allgemeine Kreditsicherung, <https://www.schufa.de/>) stellt zwar dar, welche Abfragen in den Score eingehen, allerdings bleiben die Aspekte in der Vollständigkeit und die Bildung des scores selbst eine black box. Gegen öffentliche Transparenz spricht aus Sicht des Unternehmens, dass die Werte mutwillig zum eigenen Vorteil manipuliert werden: <https://www.schufa.de/de/ueber-uns/daten-scoring/scoring/transparente-scoreverfahren/>. Man kann sich fragen, was wäre daran so falsch? So schwächt z.B. aller Wahrscheinlichkeit die Anzahl der Kreditkarten den Bonitätsscore. Was spräche dagegen, nur eine zu nutzen? Die Nichteinhaltung von Ratenzahlungen (z.B. bei der Kreditkarte) schwächt den Bonitätsscore ebenfalls. Was wäre falsch daran, wenn eine Person im Versuch, den Score zu verbessern, sich bemühte, die Ratenzahlung korrekt einzuhalten? Es wird schwierig bleiben, den Faktor Vertrauenswürdigkeit im Score abzubilden. Allerdings wird generell das negative Bedrohungsszenario als allgemeine Grundlage angenommen. Ich frage: taugt die Unterstellung: „ich halte verhalte mich im kleinen einige Zeit korrekt, um im großen Stil einen Kredit zu ergaunern, den ich dann nicht zurückzahle.“ tatsächlich für eine generalistische Annahme? Problematisch wird es, wenn allein die Schufaauskunft dazu führt, ein Ansinnen abzulehnen, denn gerade diese Generalisierung und das Nichtberücksichtigen anderer Aspekte diskriminiert.

Einen weiteren kritischen Aspekt sieht Cathy O'Neil im sich quasi automatisch ergebenden Verlangen nach mehr Daten.

Aber je mehr Informationen die Versicherer sammeln, desto zuverlässigere Vorhersagen werden sie treffen können. Das ist das Wesen der Datenwirtschaft. Die Unternehmen, die die meisten Erkenntnisse aus diesen Daten gewinnen und sie zu Profit machen, werden die Konkurrenz hinter sich lassen. Sie werden Gruppenrisiken genauer prognostizieren können (obwohl sie immer wieder am Individuum scheitern werden). Und je mehr sie von solchen Daten profitieren, desto hartnäckiger werden sie darauf drängen, immer noch mehr davon zu bekommen. (AdA S. 187)

Dieses hartnäckige Drängen kann sich darin zeigen, dass geschäftliche oder vertragliche Beziehungen nur unter der Vorgabe, Daten sammeln zu dürfen, eingegangen werden. Problematisch ist ebenfalls, dass mit „immer mehr Daten“ der Glaube wächst, absolute also korrekt zutreffende Risikoprognosen zu liefern. Jeder der Prognose entgegenstehender individueller Einzelfall wird eher als Fehler des Einzelfalls (also im

Falle der Versicherung: des versicherten Menschen) und nicht des Systems begriffen. Denn die falsche Annahme, dass sich Algorithmen nicht irren (können), bricht sich gerade durch die abstrakte entfernte (von anderen Unbekannten programmiert) dem eigenen Intellekt verschlossene Funktionsweise der Algorithmen Bahn.

Algorithmen wird häufig eine „härtere Wahrheit“ zugesprochen, die faktisch nicht vorhanden ist. O’Neill dazu:

Sehen Sie den Widerspruch? Ein Algorithmus verarbeitet einen Haufen Daten und errechnet daraus eine Wahrscheinlichkeit, dass eine bestimmte Person ein schlechter Mitarbeiter, ein Kreditrisiko, ein Terrorist oder eine miserable Lehrerin sein könnte. Aus dieser Wahrscheinlichkeit wird ein Score destilliert, der das Leben eines Menschen auf den Kopf stellen kann. Aber wenn diese Person sich dagegen wehrt, ist ein »Indiz«, das auf das Gegenteil hindeutet, einfach nicht gut genug. Ihre Beweise müssen eindeutig sein. Die menschlichen Opfer von WMDs, das werden wir immer wieder sehen, müssen wesentlich stichhaltigere Beweise erbringen als die Algorithmen selbst. (AdA S. 17)

Warum werden Algorithmen als „härtere Wahrheit“ angesehen: es fehlt Transparenz und Verständnis, auf welcher Grundlage der Algorithmus von wem programmiert wurde. Da es sich um Mathematik handelt, wird einem Algorithmus wissenschaftliche Stichhaltigkeit, Zuverlässigkeit und Wahrheit unhinterfragt unterstellt. Dieses Missverständnis muss m.E. (wenn nicht anders möglich rechtlich verankert) ausgeräumt werden. Eine neutrale Gegenüberstellung gleichrangiger Argumente muss zwingender Ausgangspunkt von Beschwerden und Auseinandersetzung Algorithmus versus Individuum sein.

O’Neil weist im Zusammenhang der Datensammelwut der Versicherung darauf hin, dass sich dadurch prinzipielle Zielkonflikte ergeben:

Ursprünglich war es das Ziel einer Versicherung, Risiken in der Gesellschaft gleichmäßig zu verteilen. In einer aufs Individuum ausgerichteten Welt zahlen wir jedoch nicht mehr den Durchschnitt, sondern werden mit den voraussichtlichen individuellen Kosten belastet. Anstatt Schicksalsschläge abzufedern, werden die Versicherungsunternehmen fordern, dass wir für solche potenziellen Schicksalsschläge im Voraus zahlen. Das konterkariert den Sinn einer Versicherung. (AdA S. 188)

Es sind – in Deutschland bzw. der EU – gesetzliche Regulierungen notwendig, um z.B. Versicherungen vorzugeben, wie sie persönliche Daten wie Schrittzähler etc. nutzen dürfen. In den USA müssen Angestellte einzelner Unternehmen Zusatzkrankenkassenbeiträge bezahlen, wenn sie ihre Daten wie BMI etc. nicht zur Verfügung stellen. Dieses Umdrehen der Versicherungsbedingungen muss in Europa verhindert werden. O’Neil sieht weitere Bedrohungen, wenn nicht gesetzliche Regelungen getroffen werden:

Wenn die Arbeitgeber erst einmal riesige Bestände an Gesundheitsdaten ihrer Beschäftigten zusammengesammelt haben, wer soll sie denn dann noch daran hindern, Gesundheitsscores zu entwickeln und sie zu nutzen, um Stellenbewerber auszulesen? Da ein großer Teil der Indikator-Daten – seien es nun Schrittzähler oder Schlafmuster – nicht gesetzlich geschützt ist, wäre das theoretisch völlig legal. (AdA S. 193)

Zudem muss man hinterfragen, welche Werte herangezogen werden (dürfen). O'Neil nennt mit dem BMI einen Wert, der gegenüber Ausnahmen eindeutig fehlerhaft ausfällt. Wird hier Wirklichkeit abgebildet, verdreht oder gar zum unternehmerischen Zwecke der Gewinnerwirtschaftung selbst definiert?

Hinzu kommt, dass viele durchtrainierte Sportler einen extrem hohen BMI haben, weil Fett weniger wiegt als Muskelgewebe. In der Alternativwelt des BMI hat LeBron James Übergewicht. S. 194

Ich bin weiter oben im Abschnitt über den Schufa Bonitätsscore bereits darauf eingegangen. Die Grundlagen der Algorithmen werden geheim gehalten. Die Ziele und die Art wie durch die Algorithmen erreicht werden sollen, werden nicht hinterfragt. Warum werden Algorithmen anders behandelt als menschliche Beschwerdeführer? Im Zeitalter der lernenden Maschine bedeutet es die selbstlernende Blackbox Algorithmus noch mehr:

Der springende Punkt ist, dass in den kommenden Jahren wahre Ozeane von Verhaltensdaten direkt in Künstliche-Intelligenz-Systeme einfließen werden. Und diese Systeme werden für den Menschen immer geheimnisvolle Blackboxen bleiben. Während sich diese Entwicklung vollzieht, werden wir kaum jemals erfahren, zu welchen »Stämmen« wir »gehören«, oder warum. Im Zeitalter der intelligenten Maschinen werden uns die meisten Variablen ein Rätsel bleiben. Viele dieser »Stämme« werden von einer Stunde – oder gar Minute – zur nächsten mutieren, wenn die Systeme die Menschen von einer Gruppe in eine andere verschieben. Denn immerhin kann sich ja dieselbe Person um acht Uhr morgens ganz anders verhalten als um acht Uhr abends. Diese automatischen Programme werden immer stärker bestimmen, wie wir von den anderen Maschinen behandelt werden – den Systemen, die entscheiden, welche Anzeigen wir zu sehen bekommen, die Preise für uns festlegen, einen Termin beim Hautarzt machen oder uns zeigen, wie wir am besten von A nach B fahren können. Sie werden hochgradig effizient und allem Anschein nach willkürlich sein und von niemandem zur Rechenschaft gezogen werden können. Niemand wird in der Lage sein, ihre Logik zu verstehen oder zu erklären. (AdA S. 190)

Selbst die Programmierer eines selbstlernenden Algorithmus wird nicht immer und mit der Zeit immer weniger in der Lage sein, Entscheidungen von lernenden Algorithmen nachzuvollziehen. Hier stoßen wir auf ein umfassendes gesellschaftliches und ethisches Problem: lassen wir selbstlernende Maschinen uns Menschen beherrschen?ⁱⁱ Wie gehen wir mit menschlichen „Einsprüchen“ um? Wie sind derartige ethische Regulierungen technisch in einen Algorithmus zu implementieren? Ist das überhaupt möglich? Diese Fragen müssen einem politisch gesellschaftlichen Diskurs zugeführt werden – im klaren Bewusstsein, dass ein klares eindeutiges „richtig“ nur schwerlich und sicher nicht endgültig zu finden ist. Zugleich ist eine also unvollkommene (gesetzliche) Regeln dringend erforderlich, um die nur in Teilen rationale ethisch moralische – menschliche – Werteordnung vor absolut schlussfolgernden aber von Menschen nicht mehr nachvollziehbare Maschinenentscheidungen zu verteidigen und zu schützen. Ich bin nicht alleine mit dieser Forderung, prominente Personen wie zum Beispiel Elon Musk, CEO von Tesla fordern diesen Diskurs ebensoⁱⁱⁱ.

Wenden wir uns mit Cathy O'Neil der Verbreitung von Information (und Meinung) zu.

Stellen wir uns also einmal vor, ich würde mich entschließen, eine Kampagne für strengere Regulierung von WMDs ins Leben zu rufen, und ich poste einen entsprechenden Aufruf auf meiner Facebook-Seite. Welche meiner Freunde würden ihn in ihrem Newsfeed sehen?

Ich weiß es nicht. Sobald ich auf »Senden« geklickt habe, gehört dieser Aufruf nicht mehr mir, sondern Facebook, und dann trifft der Algorithmus des sozialen Netzwerks die Entscheidung, wie er am besten zu verwenden sei. Er berechnet, wie wahrscheinlich der Aufruf jedem meiner Facebook-Freunde gefallen würde. Manche von ihnen, so weiß der Algorithmus, unterschreiben häufig solche Appelle, und leiten sie dann vielleicht sogar weiter an ihren eigenen Freundeskreis. Andere halten sich mit so etwas nicht auf. Aber einige meiner Freunde interessieren sich für das, was ich tue, und klicken häufig auf die Artikel, die ich poste. Der Facebook-Algorithmus berücksichtigt all diese Informationen, wenn er entscheidet, wer meinen Aufruf zu sehen bekommt. Für viele meiner Freunde wird er so weit unten in ihrem Newsfeed vergraben sein, dass sie ihn nie zu sehen bekommen werden.

Das ist es, was passiert, wenn das enorm mächtige Netzwerk, das wir mit 1,8 Milliarden Nutzern teilen, außerdem eine börsengehandelte Aktiengesellschaft ist. Facebook mag sich ja vielleicht wie ein neuzeitlicher Rathausplatz anfühlen, aber dessen ungeachtet bestimmt das Unternehmen nach seinen eigenen Interessen, was wir in seinem sozialen Netzwerk zu sehen bekommen und erfahren. (AdA S. 197f)

O'Neil stellt hier einen wichtigen Fakt fest: der Konzern Facebook, dessen Ziel Gewinne und die Steigerung des Unternehmenswertes ist, muss unterstellt werden, dass die Verteilung meiner Informationen und meiner Meinung diesem unternehmerischen Ziel untergeordnet ist. Die Steuerung wird Algorithmen überlassen, die vorgegebener Maßen darauf programmiert werden, um diese unternehmerischen Ziele zu erreichen. Es gibt zwar die Funktion, den Newsfeed in Facebook (zumindest nicht nach „TOP Meldungen“ sondern „Neueste Meldungen“ anzuzeigen. Diese Einstellung muss allerdings ständig vom Nutzer aktiv wieder und wieder eingestellt werden. D.h. um im Bild des Rathausplatzes zu bleiben, dass dieser Platz durch unsichtbare Schranken nach für uns nicht durchschaubaren Regeln uns von anderen Menschen trennt – oder auch nicht.

Und Facebook kann dieses Wissen nutzen, um das Verhalten von Menschen zu beeinflussen – bei dem es in diesem Fall darum ging, ihre Stimme abzugeben.

Das ist ein beachtliches Machtpotenzial, und Facebook ist nicht das einzige Unternehmen mit einer solchen Machtfülle. Auch andere börsennotierte Konzerne wie Google, Apple, Microsoft und Amazon sowie Mobilfunknetzbetreiber wie Verizon und AT&T besitzen enorme Datenmengen über einen großen Teil der Menschheit – und sie haben dadurch die Möglichkeit, uns nach Belieben in die eine oder andere Richtung zu lenken. (AdA S. 198)

Einige wenige Konzerne haben m.E. zu große Machtfülle, die nicht etwa für die Erhaltung des Staates oder gesellschaftlicher und sozialer Strukturen eingesetzt werden, sondern um privatwirtschaftliche Gewinne zu erwirtschaften. Was unter diesem Primat gesellschaftlich passiert ist Randbedingung (sic!). Eine in dieser Dimension nicht akzeptable Missachtung sozialer und gesellschaftlicher Macht und damit: Verantwortung.

O'Neil spinnt den Faden weiter:

Wie wir gesehen haben, konzentrieren sie [FH: Unternehmen wie die im vorherigen Zitat erwähnten] sich normalerweise darauf, Geld zu verdienen. Aber ihre Gewinne hängen eben auch in hohem Maße von der Politik der Regierung ab. Die Regierung beaufsichtigt sie – oder unterlässt es ganz bewusst; sie genehmigt oder blockiert ihre Fusionen und Firmenkäufe und bestimmt die Steuerpolitik (wobei sie in vielen Fällen die Augen vor den Milliarden verschließt, die in Offshore-Steuerparadiesen geparkt sind). Aus diesem Grund schicken die Technologieunternehmen – wie alle anderen Konzerne auch – Heerschaaren von Lobbyisten nach Washington und schleusen diskret Hunderte von Millionen Dollar an Wahlkampfspenden ins politische System.⁶ Und heute stehen sie außerdem im Begriff, sich die Möglichkeiten zu verschaffen, unser politisches Verhalten – und mit ihm die Zusammensetzung der US-Regierung – gezielt zu beeinflussen, indem sie einfach nur ihre Algorithmen optimieren. S. 198

Die Macht, Information und Meinung zu steuern, kann und wird im Sinne der Unternehmensziele darauf verwendet, möglichst günstige Rahmenbedingungen für das Unternehmen zu erreichen und zu erhalten. Jenseits der „normalen“ Lobbyarbeit und dem Versuch über eigene Informationsgenerierung und -kanäle^{iv} kann z.B. Facebook darüber hinaus nicht selbst erzeugte Information und Meinung gesteuert Milliarden Personen zuteilen. Aus meiner Sicht ist das ein enormes gesellschaftliches Risiko, welches diskutiert werden muss: Unternehmen „missbrauchen“ Nutzerdaten aus eigenem Lobbyinteresse – zum Beispiel der Wahlbeeinflussung (im Interesse des eigenen Unternehmens). Das ist eine Machtausübung „zweiter Ordnung“: Personen gehen etwa wählen, weil Freunde gewählt haben und nicht aus eigener Rason oder Überzeugung - und wählen dasselbe - wie die von Facebook gesteuert dargestellten Freunde, da diese Wahlempfehlung“ den Interessen der Plattform förderlich ist. Für den Nutzer, der diese „Wahlempfehlung“ sieht, erschließt sich der unternehmensinteressensgesteuerte Kontext nicht.

62 Prozent der Nutzer nicht bewusst ist, dass Facebook den Newsfeed manipuliert. (AdA S. 201)

Diese vergleichsweise einfache Manipulation ist vielen Nutzern nicht bewusst. Kann man diesen Umstand grundsätzlich ändern, also: kann man „praktisch alle Nutzer“ diesen Sachverhalt vermitteln und von ihnen erwarten entsprechend verantwortungsbewusst damit umzugehen? Könnte eine Lösung entsprechend der Warnhinweise auf Zigarettenschachteln diese sein: ständig einblenden zu müssen: „diese Informationen sind selektiv von unserem Algorithmus für sie ausgewählt.“ Ohne den derzeit üblichen Hinweis, dass diese Manipulation zum Vorteil des Nutzers selbst geschehe. Genauso gut könnte man dann hinzusetzen „...um so mit Ihrer Nutzung die höchsten Gewinne zu erwirtschaften und wo immer möglich unsere Unternehmensinteressen zu befördern.“

Diese Machtausübung zweiter Ordnung unterscheidet das Vorgehen dieser Unternehmen auch von traditionellen Medien:

Natürlich waren es in diesem Fall nicht die Freunde, welche die Zeitung brachten, sondern Facebook. Man könnte einwenden, dass gedruckte Zeitungen seit Menschengedenken eine ähnliche Macht ausüben: Der Chefredakteur entscheidet, welche Nachrichten auf die erste Seite kommen und wie sie präsentiert werden sollen. Er entscheidet zum Beispiel, ob ausgebombte Palästinenser oder trauernde Israelis in den Mittelpunkt gerückt werden, ein Polizist, der einem Baby das Leben rettet, oder einer, der einen Demonstranten niederknüpelt. Solche Entscheidungen beeinflussen zweifellos sowohl die öffentliche Meinung als auch Wahlen. Das Gleiche gilt auch für Fernsehnachrichten. Aber wenn die New York Times oder CNN eine Geschichte bringen, bekommt sie theoretisch jeder zu sehen. Die journalistische

Entscheidung ist eindeutig, für jeden schwarz auf weiß zu lesen, nicht intransparent. Und später wird öffentlich darüber diskutiert, ob diese Entscheidung richtig war. (AdA S.201)

Wir gestatten heute einigen wenigen Unternehmen unglaublich viel Macht und Einfluss - womöglich durch künstliche Intelligenz ohne menschliche Einflussnahme. O'Neil kleidet diesen Gedanken in folgende Frage:

Was würde also passieren, wenn Facebook am Tag der Wahl mit den Emotionen der Menschen spielen würde? (AdA S.202)

Ich halte es für dringend angeraten, darüber zu debattieren, wie diese Macht eingeschränkt werden kann. Das sie eingeschränkt werden muss, steht für mich außer Frage, da die gesellschaftliche (Aus-)Wirkung für diese Unternehmen wenn überhaupt eine Randbedingung ist.

Jedoch gebrauchen Politiker und Parteien ebensolche WMDs und missachten mit deren Einsatz die für die Erhaltung der (im Wesen nicht perfekten) Demokratie gesellschaftlich-politische Auseinandersetzung. Pointiert gesagt entdemokratisiert sich der der Wahlkampf durch das Verhalten der Politiker und Parteien, vor Wahlen möglichst zielgenau jeden einzelnen (!) Wähler mit auf das Individuum abgestellten Informationspartikeln zu füttern, der die aus richtige Wahl wahrscheinlicher macht. Zudem stellt O'Neill fest:

Diese heranwachsende Wissenschaft des Mikrotargeting [FH: bei politischen Kampagnen] mit ihren Profilen und Vorhersagen passt nur allzu gut in unsere finstere Sammlung von WMDs. Sie ist riesig, intransparent und nicht zur Verantwortung zu ziehen. Sie deckt Politiker und erleichtert es ihnen, viele verschiedene Dinge für viele verschiedene Menschen darzustellen.

...

Die Programme haben unser Wahlverhalten schon prognostiziert, und jeder Versuch, es zu ändern, ist die dafür notwendige Investition nicht wert. (AdA S. 215)

Der demokratische Prozess wird auf dem Altar der erfolgreichen Machtgewinnung geopfert. Gesellschaftlich gesehen ein Pyrrhussieg: denn auf breiten demokratischen Diskurs, Meinungs austausch, Meinungsstreit wird verzichtet, sind diese doch aus Sicht des „unternehmerisch gewinnorientierten“ Wahlkampfes ineffizient, wenig produktiv und nicht Ziel führend, also nicht wahlentscheidend. Der eigentliche demokratische Sinn einer Wahl, eine breite Auseinandersetzung möglichst vieler Personen über die zu verfolgenden gesellschaftlichen Ziele, welcher Politiker, welche Partei welche verfolgt, bleibt aus. Darf Demokratiemüdigkeit verwundern, wenn systematisch grundsätzliche gesellschaftlich demokratische Werte unterminiert werden?

Nicht nur politisch, auch gesellschaftlich haben sich selbst verstärkende Algorithmen negative Auswirkungen. Denn Sie verschärfen die vertikale Undurchlässigkeit der Gesellschaft.

Es wird immer gefährlicher und teurer, in einer von WMDs bestimmten Welt arm zu sein. (AdA S. 217)

[...] in unserer Gesellschaft, in der man mit Geld Einfluss kaufen kann, haben diese WMD-Opfer kaum eine Stimme. [...]

Wir müssen darüber nachdenken, wie wir im modernen Leben Schuld zuweisen und wie mathematische Modelle diesen Zyklus verschärfen. (AdA S. 220)

Mir ist der zuletzt zitierte Satz zu wenig exakt formuliert. Die Schuldfrage wird häufig gar nicht gestellt, daher wird entsprechend häufig auch gar nicht darüber nachgedacht, dass man mit der Verwendung von WMDs Schuld zuweist. Daher müssten die Auswirkungen eines einzusetzenden Algorithmus ganz anders und eben auch ethisch moralisch diskutiert werden. Eine solche Diskussion ist allerdings jenseits aller geschäftlichen Interessen. Faktisch werden die Benachteiligten durch die WMDs zu Benachteiligten abgestempelt (einmal in diesem Eimer, immer in diesem Eimer). Und keine dieser WMDs ist darauf programmiert, Chancen zu geben, sondern Risiken zu minimieren oder Umsätze zu optimieren. Daher wird es eine gesellschaftlich-politische Aufgabe sein, diese ethisch moralische Diskussion zu führen. In dieser wird sicher folgendes Argument eingeführt werden:

Doch menschliche Entscheidungsprozesse lassen zwar häufig zu wünschen übrig, aber sie haben einen wichtigen Vorteil: Sie können sich weiterentwickeln. Automatisierte Systeme bleiben dagegen in der Zeit stecken, bis ein Ingenieur ihnen auf den Grund geht und sie ändert. (AdA S. 221)

Ist das „selbstlernende System“ eine bessere Alternative zur sich selbstverstärkenden WMD? Im Zweifel ist es nicht viel attraktiver, da Entscheidungsprozesse vollends in der mathematischen Blackbox verschwinden, die kein Mensch mehr nachvollziehen kann. Eine menschliche demokratisch freiheitliche Gesellschaft muss aber aus meiner Sicht auf für Menschen exakt nachvollziehbare Entscheidungsprozessen aufbauen, die sich eben auch an ethisch-moralischen Grundwerten (in Deutschland angefangen mit den ersten 20 Paragraphen des deutschen Grundgesetzes) orientieren.

Big Data-Prozesse kodifizieren die Vergangenheit, aber sie können nicht die Zukunft erfinden, denn das erfordert moralisches Vorstellungsvermögen, und das ist etwas, was nur der Mensch einbringen kann. Wir müssen gezielt bessere Wertvorstellungen in unsere Algorithmen einbauen und auf diese Weise Big-Data-Modelle entwickeln, die unseren moralischen Vorgaben folgen. Manchmal wird das bedeuten, Gerechtigkeit über Profit zu stellen. (AdA S. 222)

Es ist eine für mich eine sehr schwierige Vorstellung, moralische Vorgaben in Big Data Modelle zu integrieren. Ebenso überzeugt bin ich, dass wir schnellstens damit anfangen müssen. Ein guter Startpunkt könnte folgender „hypokratischer“ vielmehr sollte er „Zuse-ischer“ Eid heißen, wenn man denn Konrad Zuse als Erfinder des Computers als Urvater der heutigen Programmierer ansieht.

Wie können wir beginnen, die mathematischen Modelle zu regulieren, die in immer höherem Maße unseren Alltag bestimmen? Ich würde vorschlagen, dass dieser Prozess von den Modellentwicklern selbst initiiert werden sollte. Wie jeder Arzt sollte auch jeder Data Scientist einen hippokratischen Eid leisten, der auf die möglichen Missbräuche und Fehlinterpretationen seiner Modelle abzielt. Nach dem Zusammenbruch der Finanzmärkte im Jahr 2008 entwarfen die Finanzingenieure Emanuel Derman und Paul Wilmott einen solchen Eid, der folgendermaßen lautet:

** Ich werde stets daran denken, dass ich die Welt nicht erschaffen habe und dass sie sich nicht an meine Gleichungen hält.*

** Obwohl ich meine Modelle unerschrocken nutzen werde, um Zahlenwerte einzuschätzen, werde ich mich nicht allzu sehr von der Mathematik beeindrucken lassen.*

** Ich werde nie Realität für Eleganz opfern, ohne zu erklären, warum ich das getan habe.*

** Außerdem werde ich die Menschen, die mein Modell nutzen, nicht in falscher Sicherheit wiegen über seine Richtigkeit. Vielmehr werde ich ausdrücklich auf seine Annahmen und Fehler hinweisen.*

** Mir ist klar, dass meine Arbeit enorme Auswirkungen auf Gesellschaft und Wirtschaft haben kann, von denen ich einen Großteil überhaupt nicht absehen kann. (AdA S. 223)*

Demut und Respekt sind zentrale Eigenschaften dieses Eid-Vorschlages. Es kann nur Startpunkt sein, wie im letzten Aspekt dargestellt, reicht er nicht aus. Denn ein Diskurs über die Auswirkungen ist umfassend und überfordert den einzelnen Programmierer.

Die WMD erzeugt zwar Einnahmen für Suchmaschinen, Lead-Sammler und Vermarkter, ist aber aus gesamtwirtschaftlicher Sicht ein Parasit. Ein Regulierungssystem für WMDs müsste solche versteckten Kosten erfassen und darüber hinaus auch zahlreiche nicht-numerische Variablen berücksichtigen. (AdA S. 225)

Einem Programmierer zuzumuten, bei ihrer Programmierung solche ethisch-moralischen Aspekte zu berücksichtigen, finde ich, überfordert eine solche Person (bzw. ein solches Team). Zur Veranschaulichung ein Beispiel jenseits digitaler Algorithmen: wenn die deutsche Bahn Gleise erneuert (wie in den ersten vierzehn Tagen im November 2017 zwischen Köln und Bonn), verursachen diese Bauarbeiten tausende von mehr Reisetunden von den Pendlern. Gesamtwirtschaftlich wäre es sinnvoller (wie nachmittags nach Arbeitsende der Gleisarbeit) auch vormittags den Pendlerstrom durchzuleiten und mit den Arbeiten erst nach 9 Uhr zu beginnen.^v Dieser gesamtwirtschaftliche Nutzen ist dem Unternehmen Deutsche Bahn ein klarer wirtschaftlicher Nachteil.

Eine solche Motivation in Unternehmen zu erwarten, würde bedeuten, dass die Grundfesten des marktwirtschaftlichen Denkens über Bord geworfen werden. Daher erscheint mir der Ansatz nicht gänzlich überzeugend.

M.E. müsste eine Kommission das Ziel eines Algorithmus – mindestens all jene mit großem gesellschaftlichen Einfluss - inhaltlich (WMD ja oder nein?) und ethisch-moralisch prüfen und gegebenenfalls verbieten.

Das impliziert eine Menge offener Fragen: Wann hat ein Algorithmus einen gesellschaftlichen Einfluss, der groß genug ist? Wann prüfe ich einen Algorithmus? Wer prüft diesen? Selbst diejenigen, die über die Zusammensetzung einer solchen Kommission entscheiden, wären mächtige Leute, die zweifelsohne Lobbying ausgesetzt wären. Nichtsdestotrotz muss man (auch im Wissen des möglichen Scheiterns) derartige Institutionen ausprobieren:

Es gibt schon jetzt Initiativen, die fordern, Algorithmen zu prüfen, ähnlich wie eine Betriebsprüfung. So haben zum Beispiel Forscher an der Princeton University das Web Transparency and Accountability Project (»Projekt für Transparenz und Rechenschaftspflicht im Netz«) ins Leben gerufen.¹⁰ Sie programmieren Software-Roboter, die sich online als diverse Arten von Mensch ausgeben – reich, arm, männlich, weiblich oder geistig beeinträchtigt. Indem sie untersuchen, wie diese Roboter behandelt werden, können die Initiatoren des Projekts

Voreingenommenheiten automatisierter Systeme feststellen, von Suchmaschinen bis hin zu Stellenbörsen. Ähnliche Initiativen entwickeln sich auch an anderen Universitäten, etwa an der Carnegie Mellon University und am MIT. (AdA S. 229)

Eine essenzielle Grundlage ist allerdings unabdingbar. Meiner Meinung ist dies auch verhältnismäßig einfach, hier einen politisch-gesellschaftlichen Konsens zu erzeugen:

Der erste Schritt muss also sein, unsere Technologie-Utopie in den Griff zu bekommen, jene grenzenlose und unberechtigte Hoffnung in das, was Algorithmen und Technologie leisten können. Bevor wir Besserung von ihnen erwarten können, müssen wir uns eingestehen, dass sie nicht alles können. (AdA S. 226)

Warum bin ich überzeugt, dass sich ein gesellschaftlicher Konsens erzeugen lässt? Cathy O'Neil stellt wesentliche Argumente am Beispiel des Modells dar, welches in den USA den Unterricht zu verbessern vorgibt:

Das Modell [FH: des schulischen Mehrwerts] ist einfach noch nicht gut genug, um als Grundlage für wichtige Entscheidungen über Menschen dienen zu können, denen wir das Unterrichten unserer Kinder anvertrauen. Solche Entscheidungen erfordern Feingefühl und Kontextwissen, und selbst im Zeitalter von Big Data sollte es Menschen vorbehalten bleiben, sie zu treffen. [...] Erst wenn wir ein Ökosystem mit positiven Feedbackschleifen etabliert haben, können wir erwarten, aufgrund von Daten den Unterricht verbessern zu können. Bis es so weit ist, sind solche Maßnahmen einfach nur willkürliche Bestrafungen. (AdA S. 227)

Essenzielle Entscheidungen, die über Menschen getroffen werden, dürfen nicht ausschließlich maschinell erzeugt werden. Maschinelle Programme können solche Entscheidungen sehr wohl unterstützen, doch im Eingestehen, dass Algorithmen nicht alles können, sollte – so ineffizient dies aus wirtschaftlicher Sicht auch erscheinen mag – eine Entscheidung, die Menschen qualifiziert, ihnen Chancen ermöglicht (oder auch nicht) stets von Menschen letztkontrolliert werden, ganz egal wie schwierig die Berücksichtigung ethisch-moralischer Faktoren sein mag. Denn so schwierig es sein mag, Menschen mit Richtlinien und Gesetzen für die Berücksichtigung ethisch moralischer Faktoren auszustatten, es erscheint mir einfacher als ethisch moralische Faktoren in maschinellen Code zu übersetzen (siehe dazu auch weiter unten das Beispiel Facebooks Trending Topics).

Gesetze müssen erstellt oder angepasst werden, die die Wahrung der Würde des Menschen zum Beispiel in Bezug auf die Verwendung von Gesundheitsmessdaten sichern (vgl. AdA S. 232). Ein weiteres positives Beispiel der maschinellen Unterstützung erwähnt Cathy O'Neil im Bereich der Bekleidungsherstellung.

Ebenso wie viele andere verantwortungsvoll ausgelegte Modelle geht der Sklaverei-Detektor nicht zu weit; er liefert lediglich Hinweise auf verdächtige Fertigungsorte und überlässt den Rest der Ermittlungsarbeit dem Menschen. (AdA S. 235)

Wie weiter oben von mir erwähnt, sollten menschliche Prüfkommisionen eingesetzt werden, um das Ziel eines Algorithmus – mindestens all jene mit großem gesellschaftlichen Einfluss - inhaltlich (WMD ja oder nein?) und ethisch-moralisch prüfen und gegebenenfalls verbieten. Auch diese Idee formuliert O'Neil:

Viele dieser Modelle werden – wie manche der WMDs, über die wir gesprochen haben – mit den besten Absichten entwickelt werden. Um aber diesen guten Absichten gerecht zu werden,

müssen sie transparent sein, es muss stets ersichtlich sein, welche Input-Daten sie nutzen und zu welchen Ergebnissen ihre Zielsetzungen und Algorithmen führen. Und sie müssen unabhängigen Prüfern zugänglich gemacht werden, denn schließlich handelt es sich um sehr mächtige Maschinen, die wir genau im Auge behalten sollten. [...]

Aber wie ich zu zeigen versucht habe, bestehen diese Modelle nicht nur aus Daten, sondern auch aus den Entscheidungen, welche Daten wir berücksichtigen wollen – und welche nicht. Bei solchen Entscheidungen geht es nicht nur um Logistik, Effizienz und Profit, sondern auch um zutiefst moralische Fragen.

Wenn wir vor ihnen zurückweichen und mathematische Modelle als neutrale und unvermeidliche Macht behandeln, gleichsam wie das Wetter oder die Gezeiten, entziehen wir uns unserer Verantwortung. (AdA S. 237)

Wie die Überzeugung, dass algorithmisch getroffene Entscheidungen grundsätzlich besser als menschliche scheitert, zeigt O’Neil am Beispiel Facebook. In 2014 wurden in den USA die sogenannten Trending topics in Facebook eingeführt. Hier ein originaler Teil der Definition Facebooks: “It shows you a list of topics that have recently become popular on Facebook. [...] Our team is responsible for reviewing Trending topics [...]. (<https://www.facebook.com/help/1401671260054622> recherchiert am 21.11.2017, Seite nicht in Deutschland verfügbar).

Ziel des Redaktionsteams war es, die Themen, fake news und tendenziöse Richtungen zu kontrollieren.

Als immer häufiger Berichte auftauchten, dass dieses Redaktionsteam wiederholt politisch rechtsgerichtete Inhalte gelöscht hatte, feuerte Facebook die Journalisten und vertraute Trending Topics einem neuen Algorithmus an – allem Anschein nach, um solche »tendenziösen« Eingriffe zu verhindern.⁸ Sobald er jedoch von der Leine gelassen wurde, benahm sich dieser Algorithmus noch unberechenbarer als die menschlichen Redakteure, und bald zirkulierten mehrere Berichte, dass »fake news« sich in die Trending Topics eingeschlichen hätten. (AdA S. 241)

An diesem Beispiel zeigt sich, dass eine qualitative Auswahl eine Richtung haben muss. Denn ohne diese (wertemäßig nicht neutrale) Richtung kann weder Mensch noch Algorithmus eine über die Zeit nachvollziehbare Auswahl oder Kontrolle vornehmen.

Auch in Algorithmen codierte Prozesse reflektieren Wertvorstellungen und moralische Abwägungen, ebenso wie jeder von Menschen durchgeführte Prozess; sie sehen nur sauberer aus, weil sie diesen Umstand besser verschleiern. (AdA S. 241)

Umso aufmerksamer müssen diese den Algorithmen zu Grunde liegenden ethisch moralischen Vorstellungen expliziert werden, damit zum einen nachvollziehbar wird, auf welcher Grundlage technische Bewertungen und Entscheidungen getroffen werden. Zum anderen wird damit eine Entscheidungsgrundlage geschaffen, an welchen Stellen das menschliche ethisch-moralische Urteilsvermögen der technischen Abwägung vorgezogen werden sollte. Eine allgemeingültige Definition, wie ein Algorithmus gerecht zu programmieren ist oder wie er ethisch-moralisch vertretbare Entscheidungen trifft, gibt es nach meiner Auffassung nicht. Insbesondere da das Ziel „Moral und Ethik“ sich über die Zeit verändert.^{vi}

Wahrscheinlich werden wir nie zu einer einfachen und allgemein akzeptierten Definition dessen kommen, was einen Algorithmus gerecht macht. Aber zum Glück wird dieser Diskurs zumindest inzwischen geführt; jetzt stellt sich uns die Aufgabe, diese Diskussion, die Maßstäbe für die zukünftige Rechenschaftspflicht für Algorithmen umreißt, auf eine zielstrebige, sorgfältige und inkludierende Weise fortzuführen. (AdA S. 243)

Eine derartige Diskussion in wirtschaftliche Produktionsprozesse zu integrieren, ist eine wichtige Aufgabe. Dies kann nur dann gelingen, wenn diejenigen, die Algorithmen verwenden, Restriktionen auferlegt werden.

[...] stellen ein deutliches Plädoyer dafür dar, die Beweislast jenen aufzuerlegen, die solche Algorithmen entwickeln und implementieren. (AdA S. 246)

Denn der Facebook Fall Trending topics zeigt, dass es weder möglich ist „objektive Neutralität“ zu programmieren, noch dass derartige programmierten Systeme nicht durch „schlauere“ Systeme betrogen werden könnten.

Aber wir sollten nicht erwarten, dass der Algorithmus von Facebook in absehbarer Zukunft selbstständig wird erkennen können, ob ein bestimmter Artikel Propaganda enthält oder wahre Fakten. (AdA S. 247)

Hier muss der Mensch eingreifen, nicht nur die Inputquellen sind zu prüfen, sondern auch, ob mit einem Algorithmus überhaupt das definierte Ziel erreicht werden kann oder dadurch unkontrollierbare Gefährdungen entstehen. Wie weiter oben postuliert, darf man nicht der Überzeugung verfallen, dass Algorithmen dem menschlichen Urteil grundsätzlich überlegen wären. Im Falle Trending topics zumindest scheint das bei Facebook der missliche Fall gewesen zu sein.

O'Neil ruft die Politiker zur Ordnung. Sie nimmt den Staat im Zuge seiner ordnungspolitischen Aufgaben in die Pflicht, die Auswirkungen auf die Gesellschaft zu prüfen und diese nachhaltig beeinflussende gesellschaftliche Kraft nicht einigen wenigen Großunternehmen zu überlassen.

Angesichts des in der Politik vorherrschenden postfaktischen Zeitgeists mag sich das wie Wunschdenken anhören. Wenn Daten – und das öffentliche Vertrauen in diese Daten – von Politikern als »alternative facts« verfälscht werden, wie können wir dann noch von der Big-Data-Branche erwarten, mehr Klarheit zu schaffen, statt ebenfalls zur Vernebelung der Fakten beizutragen?

Wir können das erwarten, weil wir es müssen. Zurzeit üben Mammutfirmen wie Google, Amazon und Facebook eine enorme Kontrolle über die Gesellschaft aus, weil sie die Daten kontrollieren. Sie sacken riesige Profite ein, laden dabei aber die Verantwortung, die Fakten zu prüfen, bei anderen ab. (AdA S. 249)

Am letzten Satz des Zitats anschließend: Diese Unternehmen setzen ihre algorithmischen Blackboxes als objektiv und unfehlbar. Algorithmen sind es allerdings - wie O'Neils Buch zeigt – keinesfalls. Weder in wirtschaftlicher Sicht - warum sonst würden sie permanent „verbessert“? Im Besonderen muss die Kraft ihres gesellschaftlichen Einflusses erkannt, diskutiert und kontrolliert werden. Ein weiteres Beispiel, wie tiefgreifend der „algorithmische“ Einfluss, wie gefährlich mangelnde gesellschaftliche Diskussion des Themas sein kann, liefert Christoph Drösser in der ZEIT 40/2017 S.39f (mit Registrierung abrufbar unter

<http://www.zeit.de/2017/40/datenschutz-internet-google-amazon-information-online-haendler>) im Artikel „Ich ist ein anderer“. Der Autor hat die fünf Dimensionen seiner Persönlichkeit nach dem Ocean-Modell (Offenheit, Gewissenhaftigkeit, Extraversion, Verträglichkeit, Neurotizismus) viermal bestimmen lassen: mit zwei ausführlichen Tests (100 Fragen und VisualDNA), auf Basis seiner Facebook-Likes und aufgrund seines Twitter-Profiles (IBM). Die Werte unterscheiden sich teilweise erheblich. Erschreckenderweise wird im Artikel (im letzter Absatz) der (Irr-)Glaube an den „besseren Algorithmus“ auch aus unvermuteter Ecke postuliert: Michal Kosinski ist Psychologe und forscht an der Stanford-Universität in den USA und entwickelte die Persönlichkeits-Testvariante über die Facebook-Likes und würde seinen zukünftigen Job lieber über die Analyse seiner Facebook Likes als eines Personalchefs erhalten. Er unterstellt: „Algorithmen ließen sich verbessern, menschliche Vorurteile dagegen säßen tief“. Diese Aussage missachtet, dass auch Algorithmen nicht vorurteilsfrei programmiert werden können, wie diese Ausführungen dargestellt haben. Dieses Beispiel zeigt einmal mehr die dringende Notwendigkeit, dieses Thema in der vollen Bandbreite aller gesellschaftlichen Ebenen und Gruppen ausführlich zu diskutieren.

Das Trugbild der „objektiven Algorithmen“, die ein gesellschaftlich akzeptiertes wirtschaftliches Interesse der Unternehmen befördern, enttarnt O’Neil mit ihrem Buch. Sie endet mit folgendem Aufruf:

Diese Entwicklungen sind Teil eines größeren Trends, demzufolge Daten privatisiert und privat genutzt werden, um private Profite zu erzeugen und Macht zu gewinnen, während die Öffentlichkeit von diesem Prozess ausgeschlossen und angehalten wird, sich ordentlich zu benehmen und den Algorithmen zu vertrauen. Es ist an der Zeit, sich dagegen zu wehren. [...] Ich rufe dazu auf, ein Regelwerk aufzubauen, mit dem wir jetzt und in Zukunft die Algorithmen zur Rechenschaft ziehen können. (AdA S. 250)

Ich möchte das Zitat mit einem für mich wichtigen Zusatz erweitern: Die gesellschaftliche Verantwortung der Unternehmen, die mit ihrer Macht (der Algorithmen) einhergeht, muss gesellschaftlich – von der Öffentlichkeit, aber insbesondere von den gesellschaftlichen gewählten Vertretern, den Politikern aktiv eingefordert und wenn erforderlich rechtlich erzwungen werden.

Es ist für eine Transparenz der Algorithmen mindestens der sehr einflussreichen Unternehmen wie zum Beispiel Facebook zu sorgen (zumindest für eine gesellschaftlich akzeptierte Prüfkommision), damit die Art der Einflussnahme öffentlich gemacht und diskutiert werden kann.

In der Verwendung und Kreation von Algorithmen muss die verbindende Kommunikation zwischen jenen, die die ethisch-moralischen oder gesellschaftlichen Konsequenzen ausdiskutieren und ermessen und denen die Algorithmen umsetzen, gefördert werden. Dieser Austausch setzt ein hohes Maß an Verständnis beider Seiten voraus, um Algorithmen vielleicht weniger perfekt, aber dafür mit dem geeigneten Maß an Gerechtigkeit auszugestalten. Aus meiner Sicht ist dieses Ziel nur interdisziplinär zu erreichen.

Endnoten:

ⁱ Ich spreche von Diskriminierung, benutze dennoch bewusst sprachlich die nicht genderneutrale Form: der Übersichtlichkeit wegen. Insofern: ausdrücklich schließe ich egal welchen Geschlechts die Sprache ist, sämtliche Geschlechter ein. Es gibt nicht nur zwei Geschlechter, was erfreulicherweise in Deutschland kürzlich höchstrichterlich festgestellt wurde. Das allerdings macht eine genderneutrale Formulierung äußerst kompliziert und schwieriger verständlich. Ich priorisiere den aus meiner Sicht grundsätzlichen Sinn von Sprache: den der (möglichst einfachen) Verständigung. Ich werde dem Diskurs um genderneutrale Formulierung weiter aufmerksam folgen.

ⁱⁱ Die Verselbständigung selbstlernender Maschinen, die ein für Menschen existenzbedrohendes Ausmaß erreicht hat, wird (mal mehr mal weniger) unterhaltsam in der Terminator Filmreihe im Plot umgesetzt.

ⁱⁱⁱ U.a. <http://www.faz.net/aktuell/wirtschaft/kuenstliche-intelligenz/elon-musk-tesla-chef-warnt-vor-kuenstlicher-intelligenz-15182958.html>. Abruf am 12.11.2017.

^{iv} Wie sie etwa ein Unternehmen wie Red Bull exzessiv zur Unterstützung der Unternehmensziele einsetzt.

^v Selbst die dargestellte „Lösung“ ist keinesfalls ethisch-moralisch wasserdicht. Der Zeitverlust einer Person, die nach 9 Uhr z.B. wegen eines Arzttermins dann sehr viel länger unterwegs ist, würde hier dem Pendlerzeitverlust nicht gleichgestellt.

^{vi} Ein illustres Beispiel ist der Umweltschutzgedanke. Bis in die 70er Jahre wurde dieser Gedanke in der Öffentlichkeit praktisch nicht wahrgenommen. In Deutschland postulierten die Grünen Ende der 1970er Jahre die Relevanz des Umweltschutzes. Nach etwa 20 bis 40 Jahren (also etwa innerhalb einer Generation) ist die Diskussion der wirtschaftlichen Entwicklung ohne Berücksichtigung des Umweltaspekts (u.a. bei folgenden Themen: Energiepolitik, Mobilität und Klimaschutz) in der Politik in Deutschland undenkbar.