[7]

Big Data, Hype und Kritik: Über argumentative Strategien und Strohmänner

Bettina Berendt

„Big Data" sind überall – und mit ihnen Heilsversprechen und Katastrophenszenarien zu den Auswirkungen von immer mehr Datensammlung, immer mehr analytischen Auswertungen und immer mehr Anwendungen. Neben kommerziellen Zwecken wird zunehmend auch die Verhinderung von Straftaten zum Ziel von Anwendungen, die auf Vorhersagen über menschliches Verhalten basieren. Aber was genau können und dürfen solche Vorhersagen eigentlich? Der vorliegende Beitrag argumentiert, dass naheliegende Argumente oft zu kurz greifen und imaginäre Gefahren von Big Data beschwören, dabei aber wirkliche Risiken außer Acht lassen. Am Beispiel einer detaillierten Kritik einer populärwissenschaftlichen Textpassage zum *Predictive Policing* und *Profiling* wird versucht, zu einem besseren Verständnis von Stärken und Schwächen datenbasierter Entscheidungsfindung beizutragen.

> *Big Data verspricht also ein besseres, weniger diskrimierendes und stärker individualisiertes Profiling. Das klingt akzeptabel, wenn es nur darum geht, unerwünschtes Verhalten zu verhindern, wird aber sehr gefährlich, sobald wir Big-Data-Vorhersagen dazu verwenden, um über Schuld und Strafe zu entscheiden für eine Tat, die noch gar nicht begangen wurde.*

Mit diesen Worten beschreiben Mayer-Schönberger und Cukier (2013b, 200) in ihrem populärwissenschaftlichen Bestseller über Big Data ein zentrales „Risiko" dieser ihrer Darstellung nach ansonsten so vielversprechenden neuen Technologien der Datenauswertung. Die Ausführungen schließen an den Film *Minority Report* an, in dem die Hauptfigur in letzter Sekunde zur Verhinderung eines Mordes festgenommen wird, von dem sie nichts weiß, der aber von den Orakeln als sicher von ihr zu begehend vorhergesagt wurde (Spielberg 2002). Die Autoren betrachten diese Dystopie als möglichen Endpunkt einer derzeit vieldiskutierten und zunehmend eingesetzten Praxis in der Polizeiarbeit, dem *Predictive Policing*, in dem computerisierte Auswertungen von Daten zur Vorhersage und Verhinderung möglicher Straftaten eingesetzt werden.

Auf den ersten Blick kann man dieser Textpassage zustimmen, drückt sie doch eine Abwägung von Vor- und Nachteilen von Big Data aus und bezieht Position gegen eine Dystopie. Auf den zweiten Blick jedoch ist die Passage problematisch, weil sie zentrale Elemente des Big-Data-Hypes eben nicht hinterfragt und durch die Beschränkung der Kritik auf rechtsstaatliche Selbstverständlichkeiten konsensfähig und damit zahnlos bleibt. Darüber hinaus drohen auf dem Weg von den „Versprechen" hin zu den „Risiken" wichtige rechtsstaatliche Prinzipien, insbesondere

Gesetzesvorbehalt und Verhältnismäßigkeit, auf der Strecke zu bleiben.

Diese Vermengung von Ebenen schadet dem allgemeinen Verständnis von „Big Data" und seiner Problematik, und insofern erscheint die Passage als repräsentativ auch für andere aktuelle Anwendungsbereiche von Big Data[1]. Der vorliegende Artikel versteht sich daher als Einladung zu einem kritischen und interdisziplinären *close reading* von Passagen wie der oben zitierten – denn nur auf Grundlage eines fundierten Verständnisses von Big Data kann eine produktive öffentliche Debatte über seine Einsatzgebiete entstehen. Zu diesem Zweck werden eine nichttechnische Einführung in das maschinelle Lernen von Vorhersagen gegeben sowie einige relevante rechtliche Grundpfeiler zur Verwendung von Daten beschrieben.[2] Im Fließtext werden die Bestandteile der eingangs zitierten Passage zum Zwecke des besseren Verständnisses wiederholt und analysiert.

Begriffe: Predictive Policing und Profiling

> Das Versprechen von Big Data ist, dass wir das tun, was wir immer schon getan haben – Profiling[3]

Predictive Policing beinhaltet die Verwendung von Data Mining durch die Polizei zur Vorhersage höherer Wahrscheinlichkeiten von Verbrechen sowie der Verwendung dieser Vorhersagen zur Entscheidungsunterstützung[4] – z.B. zur Einsatzplanung von

1 In Berendt 2015 problematisiere ich dies im Kontext einer Buchrezension.

2 Aus Platzgründen kann über keines der hier angesprochenen Themen ein tiefergehender Überblick gegeben werden. Zum Predictive Policing empfehle ich den Überblick von Ferguson (2016) und den Blog von Pilpul (2016), zum maschinellen Lernen / Data Mining z.B. Witten, Frank und Hall (2011) und zu Big Data die umfassende und kritische Darstellung von Kitchin (2014).

3 So beginnt die eingangs zitierte Passage in der englischen Originalversion: „The promise of big data is that we do what we've been doing all along – profiling" (Mayer-Schönberger und Cukier 2013a, 160).

4 „any policing strategy or tactic that develops and uses information and advanced analysis to inform forward-thinking crime prevention" (Uchida 2009, 1).

Streifen, in zunehmendem Maße aber auch für Zwecke wie Entscheidungen über Freigang.

Korrelatives/prädiktives Denken war schon immer Teil der polizeilichen Arbeit. Von statistischen Verfahren wird dabei erwartet, dass sie ‚objektiver' seien als die Erfahrung und Intuition von Polizisten (als ein Beispiel von vielen siehe die Wortwahl in *The Economist* [2014]).

Ein besonders kontroverser Anwendungsbereich ist das *Profiling*. Das Profiling (von potenziellen Straftätern) im Allgemeinen bezeichnet die polizeiliche Praxis, Individuen als einer Straftat verdächtig auszuwählen auf Basis einer Gruppe von Eigenschaften, von denen angenommen wird, dass sie mit Straftaten assoziiert sind (ACLU 2005). Bei Einsatz von Big-Data-Methoden bedeutet dieses, für Individuen die Vorhersage zu machen, dass sie einer Straftat verdächtig seien, auf Basis eines zuvor aus einer Datenmenge von Eigenschaften gelernten Klassifikators/Prädiktators, und aufgrund dieser Vorhersage eine Aktivität (z.B. Aufnahme der Personalien, Durchsuchung, Festnahme) durchzuführen. Das Profiling ist schon lange im Kreuzfeuer der Kritik, insbesondere wenn es sich um *Racial Profiling* zu handeln scheint, bei dem die Gruppe der Eigenschaften mit Ethnizität, Religion oder Nationalität verknüpft ist.

Anwendungsbereiche des Profiling reichen von der Vorhersage von Eigentumsdelikten über Gewaltdelikte bis hin zum Terrorismus. Diese Systeme machen Vorhersagen für Situationen, aber auch für identifizierte Individuen. Wie weit kann man mit diesen Vorhersagen gehen?

Von argumentativen Strohmännern und Panikmache

[Es] wird aber sehr gefährlich, sobald wir Big-Data-Vorhersagen dazu verwenden, um über Schuld und Strafe

> zu entscheiden für eine Tat, die noch gar nicht begangen wurde.[5]

In der Tat würde es gefährlich werden, wenn wir dieses täten. Aber ist das durch den zitierten Satz benannte Risiko überhaupt realistisch? Die Autoren führen in der im Buch nachfolgenden Diskussion aus, dass dieses gegen eine Reihe von Grundüberzeugungen verstoßen würde und beziehen sich dabei auf Konzepte wie den freien Willen und den „Kern von Gerechtigkeit"; sie deklarieren das Ansinnen als „abstoßende Idee". Aber worauf beziehen sie sich damit? Auf einen gesunden Menschenverstand, auf implizite kulturelle Übereinkünfte, die dann doch veränderlich sind, was die Dystopie möglich erscheinen lässt? Es erscheint mir sinnvoller, zu fragen, ob wir diesbezüglich auf stärkere Garantien zählen können. Spezifisch möchte ich mich auf die in der Bundesrepublik Deutschland durch die „Ewigkeitsklausel" des Art. 79 Abs. 3 Grundgesetz als nicht veränderbar deklarierte Rechtsstaatlichkeit beziehen und hierbei auf den Teilgrundsatz der materiellen Gerechtigkeit, der damit auch einen gerechten Umgang mit Schuld und Strafe fordert. Diese Grundpfeiler finden sich auch in anderen Rechtsstaaten wieder. Die Frage wird damit: Steht die im Zitat postulierte Verwendung von Big Data für die Organe eines Rechtsstaats überhaupt zur Debatte?

Die Antwort ist „nein", aus zwei Gründen.

(1) Schuld? Erstens kann die Auswertung von Daten nicht zur Begründung des Schuldvorwurfs bei einer Straftat dienen. Schuld ist eine subjektive Komponente der Straftat, die Vorwerfbarkeit eines gesteuerten Handelns. Bewertet werden die fehlerhafte Willensbildung des Täters und ihre Ursachen. Es ist nicht möglich, die Willensbildung eines Menschen als Ergebnis einer

5 Wie am Ende der Einleitung erläutert, sind dieses und die weiteren Zitate am Beginn von Abschnitten Teile der eingangs zitierten Passage aus (Mayer-Schönberger und Cukier 2013b, 200). Um den Lesefluss nicht zu stören, wurde auf Wiederholungen der bibliographischen Angabe verzichtet.

Datenauswertung darzustellen. Daten können nur die Grundlage einer menschlichen Bewertung der Willensbildung sein.

(2) Strafbar? Zweitens kann Verhalten, das noch nicht geschehen ist, nicht strafbar sein. Eine Straftat liegt vor, wenn eine strafbare Handlung begangen wurde und ggf. ein strafbarer Erfolg eingetreten ist. Die Straftat muss zwar nicht vollendet worden sein; auch der Versuch ist strafbar. Strafbar kann ein (Versuchs-) Verhalten aber erst sein, wenn der Täter die Schwelle zum „Jetzt geht es los" überschritten hat.[6] Die Auswertung von Daten zur Erkennung von Gefährdungslagen, in denen Personen statistisch eher dazu neigen, Straftaten zu begehen, kann daher per definitionem nicht dazu verwendet werden, Verhalten zu bestrafen, das noch nicht geschehen ist. Hingegen kann Software dazu verwendet werden, Gefährdungslagen zu analysieren, Vorhersagen über mögliche Straftaten abzuleiten, und durch präventive Maßnahmen die Begehung von Straftaten zu verhindern suchen. Solche Verwendungen von Big-Data-Analysen unterliegen Beschränkungen, auf die im Abschnitt „Daten, Grundrechtseingriffe und Rechtsstaatlichkeit" näher eingegangen wird.

Das *Minority-Report*-Szenario ist damit, in einem Rechtsstaat, ein Strohmann-Argument. Dies ist darum bedauerlich – und es wird gefährlich – als dass seine rhetorische Verwendung dazu geeignet ist, Angst zu schüren, aber gleichzeitig die Aufmerksamkeit von den wirklichen Risiken ablenkt. Ein erstes wirkliches Risiko ist ein Fehlverständnis dessen, was Algorithmen und Daten leisten können und aussagen; ein zweites eine Dekontextualisierung dessen, was die Handlungen des Sammelns von Daten und des

6 Die Frage, wann es ‚losgeht', muss natürlich in Bezug auf die Straftat beantwortet werden. Hierbei ist zu beachten, dass auch die *Vorbereitung* einer Straftat strafbar sein kann – wenn die Vorbereitungshandlung selbst eine Straftat i.S.d. StGB darstellt. Eine der Vorbereitung vorgelagerte *Planung* kann nur in sehr seltenen Fällen strafbar sein, und nur, wenn sie sich als selbst strafbare Handlung äußert (z.B.§129a StGB, Bildung terroristischer Vereinigungen). Und der Nachweis, dass eine solche strafbare (z.B. Vorbereitungs-)Handlung vorliegt, ist mit oder ohne Software schwierig.

Nutzens von Vorhersagen bedeuten und inwieweit sie daher in einem Rechtsstaat zulässig sind. Ersteres wird nach einer kurzen Darstellung der informatischen Grundlagen von Klassifikatoren/Prädiktoren und ihrer Evaluation (folgender Abschnitt) im Abschnitt „‚Versprechen' ist gut – Nachfragen ist besser" problematisiert. Letzteres ist Thema von „Daten, Grundrechtseingriffe und Rechtsstaatlichkeit".

Intermezzo: Maschinelles Lernen von Prädiktoren aus Daten

An dieser Stelle müssen einige Grundlagen von Klassifikatoren und ihrer Evaluation eingeführt werden. Ein Klassifikator kann als Entscheidungsregel betrachtet werden: Wenn *dies* (beobachtet wird), dann (denke oder tue) *das*. Eine solche Entscheidungsregel kann auf beliebigen Gründen basieren: Intuition, Vorurteil, frühere Erfahrungen eines Menschen oder eben auch Statistik. Im letzteren Fall nennen wir diesen Klassifikator „maschinell gelernt" oder „data-mined" und begründen die Verwendung als Entscheidungsregel damit, dass auf Basis der Klassifikation von Daten aus der Vergangenheit Vorhersagen für die Zukunft getroffen werden können – der Klassifikator agiert dann auch als Prädiktor. Es wird also angenommen, dass die Daten (bzw. die sie verursachenden Phänomene oder Menschen) der Vergangenheit und der Zukunft sich gleich verhalten. Unabhängig davon, wie ein Klassifikator konstruiert wurde, kann er über seine Vorhersagequalität auf neuen Daten evaluiert werden.

Wie wird ein Klassifikator/Prädiktor maschinell gelernt? Dies soll an einem einfachen (und doch viel benutzten) Beispiel gezeigt werden: Es gibt zwei Klassen von Entitäten, und jede Entität gehört zu genau einer dieser beiden Klassen. Betrachten wir zunächst eine fiktive Menge an Personen, deren Klasse sowie verschiedene beschreibende Eigenschaften bekannt sind. Diese Menge wird in Tabelle 1 gezeigt.

ID	Hautfarbe	Pullover-farbe	Schuhe	Hände	Urlaub in ...	Krimineller?
1	Grün	Rot	Stiefel	Schwitzig	Frankreich	Ja
2	Grün	Rot	Flip-flops	Schwitzig	Italien	Ja
3	Grün	Weiß	Sandalen	Trocken	Spanien	Ja
4	Grün	Gelb	Sandalen	Normal	Spanien	Ja
5	Grün	Weiß	High heels	Trocken	Frankreich	Nein
6	Grün	Weiß	Flip-flops	Trocken	Frankreich	Nein
7	Blau	Weiß	Stiefel	Trocken	Frankreich	Nein

Tabelle 1: Eine fiktive Datenmenge zum Klassifikatorlernen

Aus diesen Daten könnte ein einfaches Klassifikationsmodell gelernt werden, nämlich dass jeder, der einen roten Pullover trägt und Schweißhände hat, ein Krimineller ist, ebenso wie jeder, der Sandalen trägt und in Spanien Urlaub macht. Dieses Modell wäre auf den Trainingsdaten (die Daten in Tabelle 1) 100 % genau, denn immer, wenn eine der beiden Prämissen wahr ist, dann ist auch die Schlussfolgerung korrekt. Es könnte auch ein vereinfachtes Modell gelernt werden, etwa, dass alle grünhäutigen Menschen Kriminelle seien und alle anderen nicht – auch dieses, dem klassischen Vorurteil ähnlichere Modell, wäre auf den Trainingsdaten noch in 5 von 7 Fällen, also zu 71 % genau.

Aber auch ein Lernalgorithmus mit einem anderen induktiven Bias als dem Vorzug für ein möglichst einfaches Modell würde, wenn er nur aus positiven Instanzen (also aus den Kriminellen 1–4) lernen würde, schließen, dass „Grün" allein zur Vorhersage ausreicht. Es ist daher unabdingbar, dass die Trainingsdaten positive und negative Beispiele enthalten. Im vorliegenden Fall zeigen 5 und 6, dass „Grün" allein nicht trennscharf genug ist. Die Verteilung positiver und negativer Instanzen sollte möglichst balanciert sein, denn wenn man auf einem Datenset mit 1 % einer Klasse und 99 % der anderen lernt, so wird das Resultat zu spezifisch auf diese kleine Gruppe zugeschnitten sein. Auch müssen die positiven und negativen Instanzen in anderen Hinsichten vergleichbar sein, denn sonst könnte man etwa aus einer

Zusammenstellung männlicher Krimineller und weiblicher unbescholtener Bürger lernen, dass alle Männer Kriminelle seien.

Diese Zahlen sagen uns aber noch nichts darüber, wie gut das Modell wäre, wenn es zur Vorhersage auf neuen Daten genutzt würde. Letzteres ist der Maßstab der Evaluation, und die Genauigkeit (sowie andere Maße) auf Testdaten, die nicht mit den Trainingsdaten überlappen, muss angegeben werden, um einen verlässlichen Eindruck von der Qualität eines Klassifikators zu erhalten.

Betrachten wir nun ein mögliches Beispiel der Performanz auf Testdaten. Tabelle 2 zeigt die Grundstruktur der Evaluierung durch die Anzahlen der Entitäten aus der Testmenge.

Individuen sind …	… klassifiziert als Kriminelle	… klassifiziert als unschuldig	Zeilensumme
… in der Tat Kriminelle	Echte Positive: 4	Falsche Negative: 6 (irrtümlich für unschuldig gehalten)	Wahre Gesamtzahl von Kriminellen: 10
… in Wirklichkeit unschuldig	Falsche Positive: 100 (irrtümlich für kriminell gehalten)	Echte Negative: 900	Wahre Gesamtzahl von Unschuldigen: 1000
Spaltensumme	Positive: 104	Negative: 906	Gesamt: 1010

Tabelle 2: Eine fiktive Vertauschungsmatrix zur Evaluierung eines Klassifikators

In diesem fiktiven Beispiel wird ein Klassifikator, der aus Tabelle 1 gelernt wurde (oder anders erstellt wurde), auf eine neue Datenmenge von 1010 Menschen angewandt. Er klassifiziert 104 von ihnen als Kriminelle und 906 als unschuldig. Diese Vorhersage ist in 904 (4 + 900) Fällen korrekt, daher ist die *Genauigkeit* des Modells 904/1010 = 89,5 %. Allerdings ist die *Präzision* des Modells für die Klasse „Kriminelle“ nur 3,8 %: von den 104 als Kriminelle klassifizierten Individuen sind nur 4 in der Tat kriminell. Der *Recall* für diese Klasse ist 40 % (4 von 10 der tatsächlich Kriminellen werden gefunden). Im Vergleich hierzu hat der Basisprädiktor „immer nein“ eine Genauigkeit von 99 % (1000/1010) und für die

Klasse der Kriminellen einen Recall von 0. Es muss daher stets überprüft werden, ob „Genauigkeit" alltagssprachlich gemeint ist (und sich dann auf jedes dieser Maße beziehen könnte) oder spezifisch in der Fachterminologie (und dann Fragen von Präzision und Recall ausblendet).

Aufgrund all dieser Faktoren erfordert das maschinelle Lernen, um gute Klassifikatoren lernen zu können, eine Sammlung *vieler* Daten inklusive negativer Beispiele und normalisiert dadurch eine *umfassende*, *verdachtsunabhängige* Überwachung (siehe auch Coudert 2015). *Viel* aufgrund der Basisannahmen der Statistik, dass Fehler sich nur in großen Stichproben ausgleichen, *umfassend*, um durch möglichst viele Attribute möglichst viel Information abbilden zu können, und *verdachtsunabhängig / mit negativen Beispielen*, weil ein Klassifikator nicht nur aus positiven Beispielen gelernt werden kann.

Darüber hinaus darf nicht vergessen werden, dass jedwede Daten mitnichten ‚gegeben' sind, sondern immer aufgrund bestimmter Entscheidungen gesammelt und definiert werden. Es muss stets gefragt werden, wie und durch wen die ‚Grundwahrheit' (etwa, dass jemand kriminell ist) definiert und gemessen wird. Besonders problematisch (genauer gesagt, zirkulär) wird es, wenn die Grundwahrheit ex post definiert wird, mit Hilfe der Vorhersage des Klassifikators, denn dadurch werden die berichteten Erfolge (wie z.B. eine hohe Präzision) überschätzt und letztlich bedeutungslos. Dies scheint in der Berichterstattung über Drohnenangriffe zu geschehen. Drohnenangriffe sind als Big-Data-Anwendung beschrieben worden („We kill based on metadata"[7]) und können als eine Extremform von Predictive Policing im Sinne von Mayer-Schönberger und Cukier betrachtet werden: Das Ziel ist die Verhinderung von Verbrechen, Individuen stehen im Fokus, und es gibt scheinbar hohe Erfolgsraten. Letzteres ist zumindest der von Presseberichten über (hohe)

7 Michael Hayden, ehemaliger Direktor der NSA, in einer Podiumsdiskussion (Matthew Keys Live 2014).

Zahlen getöteter „Militanter“ erweckte Eindruck. Wenn aber „Militanter“ definiert wird als ‚jedweder Mann im wehrfähigen Alter, den wir töten, egal, ob und was wir sonst über ihn wissen‘ (Becker und Shanemay 2012; Greenwald 2014), dann werden die Zahlen in einer Vertauschungsmatrix uninterpretierbar.

„Versprechen“ ist gut – Nachfragen ist besser

> Big Data verspricht […] ein besseres, weniger diskrimierendes und stärker individualisiertes Profiling.

Dieses wird mit Techno-Optimismus ausgeführt:

> [M]it Big Data versucht [man], auf Einzelne und nicht auf ganze Gruppen abzuzielen. Damit soll der große Nachteil des Profilings überwunden werden, nämlich die Pauschalverdächtigkeit durch Gruppenzugehörigkeit. […] Mit Big Data können wir der Zwangsjacke der Gruppenidentitäten entkommen und sie durch zutreffendere Einzelvorhersagen ersetzen.“ (Mayer-Schönberger und Cukier 2013b, 200)

Diese Sichtweise suggeriert jedoch, dass (a) das traditionelle Profiling auf Gruppenidentitäten beruhe, die durch einfache Eigenschaften definiert sind (wie z.B. beim Racial Profiling), Eigenschaften, über die darüber hinaus persistente Stereotypen in der Gesellschaft bestehen. Im Gegensatz dazu (b) wähle das Mining von Big Data Individuen aus, und das darüber hinaus korrekt. Diese beiden Annahmen über Data Mining sind jedoch nicht richtig.

(3) Individualisierter? Wie oben dargestellt, identifiziert ein maschinell gelernter Klassifikator eine Konstellation von Eigenschaften, aufgrund derer er eine Zielgröße vorhersagt. Diese mag ‚granularer‘ sein als die Konstellationen des traditionellen Profiling[8] und die resultierenden Gruppen mögen

8 So die Formulierung in der englischen Originalfassung: „With big data we can escape the straitjacket of group identities, and replace them with much

weniger Menschen beinhalten – dennoch konstituieren die (im Beispiel) „schweißhändigen Rote-Pullover-Träger" genauso eine Gruppenidentität wie die „Menschen arabischen Aussehens" oder ähnliche Ziele traditionellen Profilings. Selbst wenn diese neue Gruppe nur eine Person enthalten sollte, so erscheint der Begriff „Individuum" eine irreführende Wortwahl für ‚das Zusammentreffen einer großen Zahl von Indikatoren in einer Person'.[9]

(4) Weniger diskriminierend? Auch bestehen Zweifel an der Behauptung, dass diese Form des Profilings weniger diskriminierend sei. Angwin, Larson, Mattu und Kirchner (2016) berichten über eine von ihnen durchgeführte Untersuchung eines Predictive-Policing-Systems, in dem bei sonst vergleichbaren Eigenschaften Schwarze für unterschiedlichste Delikte eine substantiell höhere Chance hatten, verdächtigt zu werden. Dass Big Data diskriminierende Effekte haben oder gar erst erzeugen können, ist bzgl. so unterschiedlicher Domänen wie Websuche oder mobilen Stadt-Apps gezeigt bzw. argumentiert worden (Sweeney 2013; Crawford 2013). In der informatischen Forschung sind Verfahren des *discrimination-aware* (*fairness-aware*, …) *data mining* entwickelt worden, die dazu beitragen könnten, solche Effekte zu verringern. Sie basieren jedoch darauf, dass die zu schützenden Gruppenattribute bekannt sind (etwa Ethnizität), und beruhen auf einem gegenüber juristischen Konzepten stark vereinfachten (und in manchen Kontexten falschen) Begriff von Diskriminierung bzw. ihrer Vereinfachung (siehe Berendt und Preibusch 2014). Auch können veränderte Lernalgorithmen wenig ausrichten, wenn die Daten, aus denen sie lernen, Ergebnisse früherer diskriminierender Entscheidungen sind. Schließlich kann

more granular predictions for each individual" (Mayer-Schönberger und Cukier 2013a, 161).

9 In der Medizin hat eine analoge Diskussion schon zu einer Verfeinerung der Terminologie geführt. So vermeidet man das Schlagwort „individualisierte/personalisierte Medizin" und spricht stattdessen etwa von „stratifizierter Pharmakotherapie", wenn über bestimmte Merkmale (Biomarker, Enzyme) Untergruppen von Patienten definiert werden, für die dann an diese Merkmale angepasste Therapieschemata erarbeitet werden (Ditzel 2013).

das datenbasierte Profiling zu der kafkaesken Situation führen, dass neue diskriminierte Gruppen entstehen (etwa die schweißhändigen Rote-Pullover-Träger), die ihrer Diskriminierung kaum gewahr sind, darüber aufgrund proprietärer Algorithmen und (noch?) fehlender rechtsstaatlicher Kontrollinstrumente auch nichts erfahren können und keinen rechtlichen Schutz genießen (da es kein Gesetz speziell für Hände und Pullover gibt). Generell ist es schwierig nachzuprüfen, ob Algorithmen diskriminieren, denn i.d.R. sind sie proprietär und/oder werden aus Gründen der Sicherheit nicht zu unabhängiger Prüfung freigegeben (z.B. Stroud 2014).

(5) Besser? Ob das resultierende Vorhersagemodell ‚besser' (genauer/korrekter) ist als ein traditionelles und seine Vorhersagen ‚zutreffender' sind, ist eine empirische Frage. Grundsätzlich können alle Vorhersagemodelle falsch positive und falsch negative Ergebnisse haben. Auf keinen Fall kann eine allgemeine Aussage über relative Korrektheit aus der epistemologischen Genese (menschliche Expertise, maschinelles Lernen) eines solchen Modells abgeleitet werden. Speziell im Bereich der ‚Terroristenidentifikation' ist argumentiert worden, dass die Überlegenheit maschinell gelernter Modelle noch nicht demonstriert worden sei (z.B. Solove, 2011; Cahall, Bergen, Steman und Schneider 2014).

Unabhängig von der Qualität der Vorhersagen finden allerdings weder die Sammlung von Daten noch die Nutzung von Vorhersagen in einem abstrakten Raum statt. Dieses führt zur Frage, unter welchen Bedingungen diese Handlungen legitim sind.

Daten, Grundrechtseingriffe und Rechtsstaatlichkeit

> Das [Profiling] klingt akzeptabel, wenn es nur darum geht, unerwünschtes Verhalten zu verhindern[.]

Das mag für manche Ohren akzeptabel *klingen*, aber was bedeutet eine solche Aussage? Interessanter ist die Frage, ob es akzeptabel *ist* bzw. ob staatliche Organe unter den genannten Umständen unbegrenzt profilen dürfen. Das ist nicht der Fall.

(6) Unerwünscht? Die nonchalante Verwendung des Wortes „unerwünscht" überrascht. Unerwünscht von wem? Wäre es zulässig, jemanden zu profilen, ggf. anzuhalten und/oder festzunehmen, um etwa zu verhindern, dass er Kaugummi kaut (weil das jemand anders nicht wünscht)? Wie oben dargelegt, ist es zulässig, durch präventive Maßnahmen die Begehung von *Straftaten* zu verhindern.

(7) Heiligt der Zweck die Mittel? Aber auch zur Verhinderung von Straftaten ist mitnichten jede Maßnahme „akzeptabel". Die Sammlung von Daten wie auch Maßnahmen zum Profiling und zur Verhinderung von Straftaten greifen i.d.R. in Grundrechte ein (z.B. Privacy, Datenschutz und Bewegungs-/Handlungsfreiheit). Die präventiven Maßnahmen müssen eine Reihe von Bedingungen erfüllen: Sie müssen gesetzlich geregelt sein und verhältnismäßig sein, d.h, ein „legitimes Ziel" haben sowie geeignet, erforderlich und angemessen sein (siehe z.B. Van Alsenoy, Kuczerawy und Ausloos 2013, 70). Die Konsequenzen für die Grundrechte müssen also gegen die verfolgten Ziele abgewogen werden. Auch dürfen sie allein der Verhinderung von Straftaten dienen (also den Aufgabenbereich der Polizei nicht überschreiten).

Anders ausgedrückt: Gefährlich wird es, wie Solove (2011) darlegt, wenn im Interesse der „Sicherheit" rechtsstaatliche Verfahren missachtet werden – und das sind altbekannte Strategien der Macht, die zunächst einmal nichts mit Big Data oder philosophischen Betrachtungen über den freien Willen (à la *Minority Report*) zu tun haben. Gefährlich wird es auch, wenn wir rhetorisch vereinfachen und suggerieren, dass dies „akzeptabel" sein könnte. Darüber hinaus sollten angesichts des großen Einflusses der Technologie die Verfahren weiterentwickelt werden, siehe z.B. den Vorschlag zu *technological due process* von Citron

(2007) oder die Analyse von Ferguson (2015), wie Big Data die Einschätzung einer *reasonable suspicion* verändert.

Fazit

Wenn wir Big Data und deren Folgen bewerten wollen, dann müssen wir achtgeben. Nicht nur hinsichtlich einer Abwägung von ‚Wert' und ‚Risiko', sondern auch hinsichtlich dessen, was genau ein Risiko darstellt. Das Heraufbeschwören von Horrorszenarien bringt uns nicht weiter, wenn wir gleichzeitig simplizistische ‚Lösungen' komplexer gesellschaftlicher Probleme bejahen und Heilserwartungen an die Objektivität und Korrektheit von Statistik und Computern richten. Kritisieren und verbessern müssen wir vielmehr auch diese, und hierzu sind die Detailkenntnisse und der Dialog vieler Disziplinen, insbesondere, aber nicht ausschließlich, der Informatik, Rechtswissenschaft, Psychologie, Soziologie und Politik vonnöten. Nur so haben wir die Chance, Profiling, Predictive Policing und andere Big-Data-Anwendungen im Sinne einer freiheitlich-demokratischen Rechtsordnung zu entwickeln und einzusetzen.

Danksagung. Ich danke Geoffrey Rockwell und Rob Kitchin für Kommentare zu einer früheren Version dieses Textes, Fanny Coudert für Diskussionen und Literaturempfehlungen sowie Ariane Loof und Patrick Berendt für argumentativen und Text-Input insbesondere zum Abschnitt „Von argumentativen Strohmännern und Panikmache".

Literatur

ACLU. 2005. „Racial Profiling: Definition". https://www.aclu.org/racial-justice/racial-profiling-definition. Letzter Zugriff am 05. Juli 2016.

Angwin, Julia, Jeff Larson, Surya Mattu und Lauren Kirchner. 2016. „Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks." *ProPublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Letzter Zugriff am 05. Juli 2016.

Becker, Jo und Scott Shanemay. 2012. „Secret ‚Kill List' Proves a Test of Obama's Principles and Will." *The New York Times*, 29. Mai. http://www.nytimes.com/2012/05/29/world/obamas-leadership-in-war-on-al-qaeda.html?pagewanted=1&_r=2&. Letzter Zugriff am 05. Juli 2016.

Berendt, Bettina. 2015. „Big Capta, Bad Science? On Two Recent Books on ‚Big Data' and Its Revolutionary Potential". 5. März. https://people.cs.kuleuven.be/~bettina.berendt/Reviews/BigData.pdf. Letzter Zugriff am 05. Juli 2016.

Berendt, Bettina und Sören Preibusch. 2014. „Better Decision Support Through Exploratory Discrimination-aware Data Mining: Foundations and Empirical Evidence". *Artificial Intelligence and Law* 22 (2): 175–209.

Cahall, Bailey, Peter Bergen, David Sterman und Emily Schneider. 2014. *Do NSA's Bulk Surveillance Programs Stop Terrorists?* New America Foundation. http://www.newamerica.net/sites/newamerica.net/files/policydocs/Bergen_NAF_NSA%20Surveillance_1_0_0.pdf. Letzter Zugriff am 05. Juli 2016.

Citron, Danielle Keats. 2007. „Technological Due Process". *Washington University Law Review*, 85: 1249–1313. Verfügbar auf SSRN: http://ssrn.com/abstract=1012360. Letzter Zugriff am 12. Januar 2017.

Coudert, Fanny. 2015. „‚Precrime Police' Is Not for 2054, It's for Now: How to Regulate ‚Data Intensive Policing'?". *Amsterdam Privacy Conference*, Amsterdam, 23–26 October 2015.

Crawford, Kate. 2013. *Strata 2013: Kate Crawford, Algorithmic Illusions: Hidden Biases of Big Data*. https://www.youtube.com/watch?v=irP5RCdpilc. Letzter Zugriff am 05. Juli 2016.

Ditzel, Peter. 2013. „Stratifizierte Pharmakotherapie – was heute schon möglich ist". *Deutsche Apotheker-Zeitung* 21. https://www.deutsche-apotheker-zeitung.de/daz-az/2013/daz-21-2013/stratifizierte-pharmakotherapie-was-heute-schon-moeglich-ist. Letzter Zugriff am 05. Juli 2016.

The Economist. 2014. „Parole and Technology: Prison Breakthrough". *The Economist*, April 19. http://www.economist.com/news/united-states/21601009-big-data-can-help-states-decide-whom-release-prison-prison-breakthrough. Letzter Zugriff am 05. Juli 2016.

Ferguson, Andrew Guthrie. 2015. „Big Data and Predictive Reasonable Suspicion". *University of Pennsylvania Law Review* 1632: 327–410.

Ferguson, Andrew Guthrie. 2016. „Policing Predictive Policing". *Washington University Law Review* 94, Forthcoming. http://ssrn.com/abstract=2765525. Letzter Zugriff am 05. Juli 2016.

Greenwald, Glen. 2014. „On Media Outlets That Continue to Describe Unknown Drone Victims as ‚Militants'". *The Intercept*, November 18, 2014. https://firstlook.org/theintercept/2014/11/18/media-outlets-continue-describe-unknown-drone-victims-militants. Letzter Zugriff am 05. Juli 2016.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences.* London: Sage.

Matthew Keys Live. 2014. *Former NSA boss: „We Kill People Based on Metadata"*. https://www.youtube.com/watch?v=UdQiz0Vavmc. Letzter Zugriff am 05. Juli 2016.

Mayer-Schönberger, Viktor und Kenneth Cukier. 2013a. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray (Publishers).

Mayer-Schönberger, Viktor und Kenneth Cukier. 2013b. *Big Data: Die Revolution, die unser Leben verändern wird*. München: Redline Verlag.

Pilpul, Martin. 2016. „Anekdoten aus der berechneten Zukunft". http://blog.pilpul.me/tag/predictive-policing/. Letzter Zugriff am 05. Juli 2016.

Solove, Daniel. 2011. *Nothing to Hide: The False Trade-off between Privacy and Security*. Yale, CT: Yale University Press.

Spielberg, Steven. Regie. 2002. *Minority Report*. USA.

Stroud, Matt. 2014. „The Minority Report: Chicago's New Police Computer Predicts Crimes, But Is It Racist?" *The Verge*, February 19, 2014. http://www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist. Letzter Zugriff am 05. Juli 2016.

Sweeney, Latanya. 2013. „Discrimination in Online Ad Delivery". *ACM Queue* 11 (3). http://queue.acm.org/detail.cfm?id=2460278. Letzter Zugriff am 05. Juli 2016.

Uchida, Craig. D. 2009. *A National Discussion on Predictive Policing: Defining Our Terms and Mapping Successful Implementation Strategies*. Rockville, MD: National Institute of Justice. NCJ 230404. https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=252437. Letzter Zugriff am 05. Juli 2016.

Van Alsenoy, Brendan, Aleksandra Kuczerawy und Jef Ausloos. 2013. „Search Engines after ‚Google Spain': Internet@Liberty or Privacy@Peril?" *TPRC 41: The 41st Research Conference on Communication, Information and Internet Policy*. Verfügbar auf SSRN: http://ssrn.com/abstract=2321494. Letzter Zugriff am 05. Juli 2016.

Witten, Ian. H., Eibe Frank und Mark A. Hall. 2011. *Data Mining: Practical Machine Learning Tools and Techniques*. 3. Auflage. Burlington, MA: Morgan Kaufmann.