# Staatliche Regulierung durch Big Data und Algorithmen

Sebastian Haunss und Lena Ulbricht

## Inhalt

1 Klärung der Begriffe ..... 2
2 Theoretische Grundlagen ..... 3
3 Praktische Anwendungsfelder ..... 4
4 Perspektiven für Staat und Verwaltung ..... 7
Literatur ..... 9

**Zusammenfassung**

Beim Einsatz von *big data* und Algorithmen für Regulierung in Staat und Verwaltung ist zwischen unterschiedlichen Aspekten der technischen Instrumente und verschiedenen Dimensionen von Regulierung zu unterscheiden. Auf dieser Grundlage lässt sich feststellen, wo bereits Veränderungen und Wirkungen sichtbar werden und welche Reichweite diese Veränderungen haben. Die am weitesten reichenden Veränderungen staatlicher Regulierung durch *big data* und Algorithmen zeigen sich bisher außerhalb Europas, doch auch in deutschsprachigen Ländern gibt es zahlreiche Beispiele.

Anhand der Diskussion verschiedener Anwendungsfälle ist festzustellen, dass *big data* von staatlichen Organisationen bisher weitgehend dafür eingesetzt wird, um konkrete Regeln zu setzen und zu implementieren und im Rahmen algorithmischer Regulierung im staatlichen Bereich Regeln stärker nach Zielgruppen zu

S. Haunss (✉)
SOCIUM Forschungszentrum Ungleichheit und Sozialpolitik, Universität Bremen, Bremen, Deutschland
E-Mail: sebastian.haunss@uni-bremen.de

L. Ulbricht
Weizenbaum-Institut für die vernetzte Gesellschaft, Wissenschaftszentrum Berlin für Sozialforschung, Berlin, Deutschland
E-Mail: lena.ulbricht@wzb.eu

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2020

T. Klenk et al. (Hrsg.), *Handbuch Digitalisierung in Staat und Verwaltung*,
https://doi.org/10.1007/978-3-658-23669-4_2-1

differenzieren. Weitergehende Programme, die eine *big data*-basierte Regulierung vielfältiger Lebensbereiche zum Ziel haben, existieren bisher in demokratischen Staaten noch nicht.

**Schlüsselwörter**
Big data · algorithmic regulation · Regulierung · social credit · Entscheidung

## 1 Klärung der Begriffe

*Big data* meint die massenhafte Sammlung und Auswertung der vielfältigen Daten, deren Umfang, Diversität und Erhebungsfrequenz durch die Digitalisierung aller Lebensbereiche rapide zunimmt (boyd und Crawford 2012). Staatliche Verwaltungen arbeiten schon sehr lange mit Datensätzen, die zumindest zwei der drei gängigen Kriterien für *big data* – *volume*, *variety* und *velocity* – erfüllen: Bei Zensus-Melde-, Steuer- und Sozialversicherungsdaten handelt es sich um Datensätze großen Volumens sowie großer Vielfalt. In der Regel fehlt diesen Daten zwar die dritte Dimension, die hohe Geschwindigkeit der Datengenerierung, diese wird allerdings auch nicht von allen Definitionen vorausgesetzt. Durch die Verfügbarkeit großer und diverser Datensätze im Zuge der Digitalisierung würde sich für staatliche Verwaltungen zunächst einmal wenig ändern. Allerdings geht es in der Debatte um *big data* nicht allein um die Daten selbst, sondern zu einem wesentlichen Teil um die Möglichkeiten der Datenanalyse, die sich in den letzten Jahren dramatisch verändert haben. Fortschritte im Bereich der explorativen Datenanalyse und des maschinellen Lernens erlauben es heute, viel größere Datenbestände in viel kürzerer Zeit zu analysieren. Die sprichwörtliche Nadel im Heuhaufen zu finden, stellt für aktuelle Analysemethoden kein Problem mehr da.

Algorithmen bezeichnen computergestützte Verfahren, die einen Input nach bestimmten Regeln zu einem Output transformieren. Algorithmen können die Bedingungen staatlicher Regulierung verändern. Die sogenannte *algorithmic regulation* beschreibt Entscheidungssysteme, die auf adaptiver computergestützter Informationsverarbeitung beruhen und für Regulierung eingesetzt werden, also für die absichtsvolle Beeinflussung von Verhalten (Yeung 2017). Regulierung umfasst entsprechend ganz verschiedene Praktiken unterschiedlichen Formalitätsgrades: Gesetze und Verordnungen, aber auch informale Instrumente der Verhaltensbeeinflussung, „harte“ sowie „weiche“ Steuerung. Natürlich wird *big data* je nach Instrument in unterschiedlichem Ausmaß eingesetzt. In welcher Hinsicht *big data* und Algorithmen staatliche Regulierung bereits prägen, lässt sich nicht allein mit Blick auf den deutschsprachigen Raum betrachten. Zwar gibt hier erste Erfahrungen, doch ist die Debatte stark von Erfahrungen im Ausland geprägt. Für die Erprobung von *big data* und Algorithmen in neuen Bereichen staatlicher Steuerung spielen Erfahrungen in anderen Ländern meist eine wichtige Rolle.

## 2 Theoretische Grundlagen

*Big data*-Verfahren werden heute schon in ganz verschiedenen Politikfeldern genutzt, um zwischen regelkonformen Handlungen und Regelverstößen zu unterscheiden, um Werturteile zu fällen und um Weltbilder und Sinnzuschreibungen zu fixieren. Man kann also drei Dimensionen der Wirkung von *big data* unterscheiden: eine regulative, eine normative und eine kulturell-kognitive (Ulbricht et al. 2018). Wenn es um ihren Einfluss auf die Regelsetzung geht, ist zu beachten, welche Aspekte von *big data* Regulierung beeinflussen: Geht es eher um die *Art der Daten* oder um die Form der *Analyse der Daten*? Die Größe des Datensatzes ist dabei nicht unbedingt das ausschlaggebende Kriterium. Zu unterscheiden ist vielmehr, ob die Daten domainspezifisch sind, sie also (nur) in dem Bereich und für den Zweck, für den sie erhoben worden sind, genutzt werden, oder ob es sich um unspezifische bzw. domainübergreifende Daten handelt. Private Krankenkassen in den USA ziehen etwa zunehmend unspezifische Daten über ihre Versicherten heran, etwa Konsumdaten, um Tarife und Leistungen zu bestimmen; öffentliche Krankenkassen in den USA sowie alle Krankenkassen in Deutschland beschränken sich weitgehend noch auf gesundheitsspezifische Daten. Hinsichtlich der *Analysemethoden* geht es darum, zu unterscheiden, ob klassische statistische Verfahren und regelgebundene explorative Analysetechniken zur Anwendung kommen oder ob Verfahren künstlicher Intelligenz und maschinellen Lernens (etwa *deep learning* oder hochkomplexe *decision trees*) zum Einsatz kommen, bei denen die Entscheidungsregeln nicht mehr vollständig von außen vorgegeben sind.

Innerhalb der *big data*-basierten Regulierung muss man grundsätzlich unterscheiden, ob es um den Einfluss auf konkrete, einzelne Regeln geht oder um den Einfluss auf den Regelsetzungsprozess selbst. Es stellt sich also die Frage, ob *big data* neue oder veränderte Meta-Regeln hervorbringt. Auf dieser Meta-Ebene könnte die Nutzung solcher Verfahren dazu führen, dass Regeln so konstruiert werden, dass ihre Auslegung und Präzisierung in die Datenanalyse verlagert wird, dass es also zu formal weniger präzisen Regeln kommt, die erst in der Anwendung datengetrieben präzisiert werden. Gleichzeitig könnten Regelsetzungsverfahren auch stärker responsiv gestaltet werden und es könnte damit zu einem schnelleren Regelwandel kommen.

Hinsichtlich des Einflusses auf konkrete, einzelne Regeln ist zu unterscheiden, ob *big data* zu grundsätzlich andersartigen Regeln führt oder ob „nur“ zu einer veränderten Interpretation bzw. Auswirkung je nach gesellschaftlicher Subgruppe von im Grundsatz gleich bleibenden Regeln. Dem Versprechen nach soll *big data* dazu dienen, Regulierung mit feinerer Granularität auf kleinere gesellschaftliche Gruppen anzupassen (Morozov 2014). Regeln müssten dafür beispielsweise nicht die Bevölkerung als Ganzes, sondern nur noch einzelne Bevölkerungsgruppen betreffen. Eine veränderte Regelinterpretation hätte dagegen vor allem eine Verschiebung der Autorität zur Folge. Wenn *big data* im Sinne algorithmischer Regulierung dazu genutzt wird, Entscheidungen über Regelanwendungen zu treffen, verlagert sich die Regulierungskompetenz von den bisherigen Institutionen weg in die Regeln, die die Entscheidung der Maschine bestimmen. Das wäre eine Autoritätsverlagerung weg

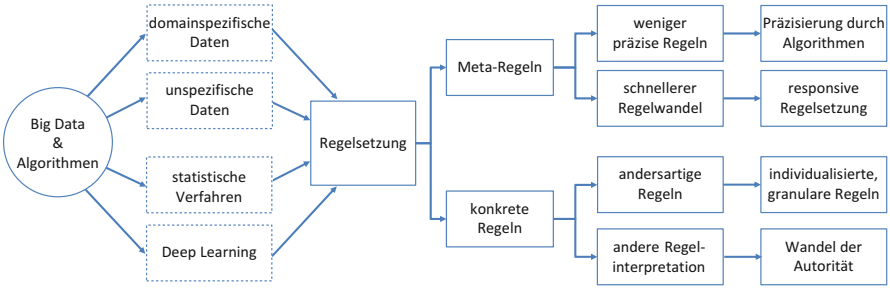


**Abb. 1** Der Einfluss von Big Data auf Regulierung. Quelle: Eigene Darstellung

von Verwaltungsmitarbeiter*innen hin zu Entwickler*innen der *big data*-Verfahren oder – im Falle von *deep learning* – in die Maschine selbst. Abb. 1 fasst diese unterschiedlichen Aspekte schematisch zusammen.

Ein Wandel der Regelsetzung oder -interpretation kann unterschiedlich weit reichen: von der vollständigen Übertragung von Verantwortung auf die Maschine bis hin zu hybriden Formen geteilter Verantwortung zwischen Mensch und Maschine. Um zu ermessen, wie weitreichend der Einfluss von *big data* und Algorithmen auf Regulierung ist, kommt es daher darauf an, Automatisierung und Eingriffstiefe zu beurteilen: Am weitreichendsten sind Systeme, in denen Entscheidungen vollautomatisch, ohne menschliche Intervention getroffen werden. Weniger weitreichend ist der Einfluss bei Systemen, die zwar Vorschläge machen, die Entscheidung aber letztendlich bei der menschlichen Anwenderin des Systems belassen. Darüber hinaus geht es darum, ob *big data*-basierte Systeme nur einzelne, spezifische und abgegrenzte Lebensbereiche betreffen oder ob sich ihr Einfluss gleich auf mehrere Lebensbereiche erstreckt.

## 3 Praktische Anwendungsfelder

Am weitesten ist algorithmische Regulierung in Deutschland bislang in der Sicherheitspolitik verbreitet; andere Politikbereiche erwägen und erproben entsprechende Verfahren erst. Dabei spielen Erfahrungen aus anderen Ländern eine wichtige Rolle. Eine wichtige Funktion von *big data* ist, Bürger*innen zu Regulierungszwecken zu klassifizieren. Bislang beruhten entsprechende staatliche Klassifizierungen meist auf relativ großen und fixierten Kategorien, wie etwa Geburtskohorten, Einkommensgruppen und Staatsangehörigkeit. Nunmehr erproben staatliche Akteure die Klassifizierung zunehmend feiner granulierter Gruppen, zum Teil in dynamischer Weise. In den USA, Großbritannien und Neuseeland werden *big data* und Algorithmen beispielsweise eingesetzt, um besser einschätzen zu können, wie groß das Risiko ist, dass ein Kind oder Jugendlicher in nächster Zeit Gewalt erfährt oder vernachlässigt wird. Ausgewertet werden Daten aus der Sozialverwaltung, gepaart mit Informationen über die Strafgeschichte von Familienangehörigen und zuweilen angereichert durch Befragungen von betroffenen Jugendlichen. Daten über Online-Verhalten und

aus sozialen Netzwerken werden bislang nicht verwendet. Je nach Anwendungskontext werden die Informationen mit einfachen statistischen Verfahren, aber auch unter Verwendung maschinellen Lernens ausgewertet. Am Ende steht meist ein Risikowert, der allerdings keine automatischen Folgen hat. Der Score ist vielmehr eine Information unter vielen, die Sozialarbeiter heranziehen, um zu entscheiden, welche der folgenden Maßnahmen sie treffen, wenn sie eine Eingabe bezüglich eines Kindes erhalten: keine Handlung, eine weitere Prüfung der Unterlagen, Entsendung eines Sozialarbeiters vor Ort, erste Maßnahmen zur Entfernung eines Kindes aus seiner Familie etc. Es handelt sich also nicht um eine vollautomatisierte Umsetzung, sondern um ein sogenanntes *human in the loop*-Modell, bei dem ein Mensch die letzte Entscheidung trifft. Allerdings gibt es Hinweise darauf, dass Menschen nur selten von der maschinellen Empfehlungen abweichen, die sogenannte *over-reliance*. Das liegt zum Teil an dem Versprechen von Objektivität, das derartige Verfahren mit sich bringen, aber auch an der Verantwortungsstruktur: Bei Fehlern, wenn etwa Kinder trotz Meldungen ans Sozialamt zu Opfern werden, ist die Sozialverwaltung oder das zuständige Personal weniger unter Legitimierungsdruck, wenn es der maschinellen Empfehlung gefolgt ist, als wenn es von ihr abgewichen ist. Entsprechend monieren Kritiker des Einsatzes von *big data* und Algorithmen in der Kinder- und Jugendhilfe, dass die maschinelle Empfehlung eine zu große Bedeutung erhalte, die aufgrund der Intransparenz, ungewissen Wirksamkeit und möglichen unerwünschten Effekten problematisch sei (Mckinlay 2017). Tatsächlich gibt es in den USA Staaten, die beschlossen haben, eine entsprechende Software wieder auszusetzen. Ähnliche algorithmische Verfahren werden in den USA auch zur Vorhersage der Rückfallwahrscheinlichkeit von Straftätern eingesetzt. Dazu zählt etwa der *COMPAS Score*, der zahlreiche spezifische sowie unspezifische Daten und ein Verfahren des maschinellen Lernens verwendet. Auch hier treffen Staatsanwält*innen und Richter*innen die letzte Entscheidung und der Risikoscore ist *de jure* nur eine Information von vielen, die die Entscheidung begründen sollen. Doch es gibt auch Kritik, dass der maschinell erzeugte Score im Entscheidungsprozess letztlich den Ausschlag geben könnte (Angwin et al. 2016).

Das bisher wohl weitreichendste Programm zum Einsatz von *big data* in staatlichen Verwaltungen ist der vom Staatsrat der Volksrepublik China 2014 veröffentlichte Plan für ein *social credit system* (State Council 2014). Darin werden Richtlinien für die Schaffung eines umfassenden *scoring*-Systems formuliert, mit dessen Hilfe bis zum Jahr 2020 die Kreditwürdigkeit von Organisationen und Personen in insgesamt 20 Bereichen, von Industrie und Handel, Steuern, Fertigungssicherheit über Umweltschutz, Lebensmittel, Medizin bis hin zu Sozialversicherung, Bildung und Forschung beurteilt werden soll. Kern des geplanten Systems ist die Erhebung von Daten in verschiedenen gesellschaftlichen Bereichen und Kontexten und die systematische Verknüpfung dieser Daten über Bereichsgrenzen hinweg. Ziel des Projekts ist die Schaffung von Vertrauen und einer „Aufrichtigkeitskultur“, es geht also über rein ökonomische Erwägungen hinaus. Bisher wird das System in China noch nicht flächendeckend umgesetzt, aber es existieren in einzelnen Großstädten und Regionen Pilotprojekte, von denen manche schon zwischen 2007 und 2011 begonnen haben (Creemers 2018). Diese Pilotprojekte gehen teilweise weit über den

im staatlichen Planungsdokument skizzierten Rahmen hinaus und erfassen auch alltägliche Verhaltensdaten der Bürger*innen. Im Rahmen des „Massenkredits“ in Suining (Creemers 2018) sowie im *social credit system* Rongchengs (Mistreanu 2018) wurde jeweils allen Bürger*innen im Einzugsbereich ein Startwert von 1000 zugewiesen, der dann durch regelkonformes und vorbildliches Verhalten erhöht werden konnte. Verspätete Kreditrückzahlungen, aber auch eher alltäglichen Regelverstöße wie Schwarzfahren oder bei Rot über die Ampel gehen führen dagegen zu Punktabzügen. Unterschreitet man einen bestimmten Wert, kommt man auf eine „schwarze Liste“ und wird damit von bestimmten Leistungen ausgeschlossen. Voraussetzung für diese weitreichenden Systeme sind ein eindeutiger Identifizierungscode für jede*n Bürger*in sowie der Ausbau biometrischer Datenbanken, mit deren Hilfe alltägliche Regelverstöße, die von Überwachungskameras und Sensoren aufgezeichnet werden, automatisch Individuen zugeordnet werden können. Allerdings stammen die Daten in den bisherigen Pilotprojekten zum allergrößten Teil nicht von Sensoren und Kameras, sondern werden ganz traditionell von Verwaltungsmitarbeiter*innen eingegeben. In solchen Pilotprojekten entstehen Systeme, die sehr weitreichend sind und die individuelle Verhaltensdaten aus einer großen Zahl von Lebensbereichen miteinander verknüpfen. Mit diesen *social credit system* entwickelt sich in gewisser Weise ein „transparenter“ Überwachungsstaat, der das Verhalten seiner Bürger*innen in (fast) allen Lebensbereichen zentralisiert und standardisiert erfasst und zugleich mit dem dadurch entstehenden Score auch umfassend Auskunft über das (aggregierte) Ergebnis der Überwachung erteilt. In Kombination mit *data mining*-Techniken und maschinellem Lernen, die in den bisher existierenden Systemen erst sehr begrenzt zum Einsatz kommen (Creemers 2018), könnte hier ein wahrhaft dystopisches *big data*-basiertes Kontrollinstrument entstehen, mit dem sich Muster normabweichenden Verhaltens zuverlässig identifizieren lassen – auch wenn noch offen ist, welche der in den Pilotprojekten eingesetzten Techniken und Datenquellen tatsächlich landesweit zum Einsatz kommen werden. In Deutschland und anderen Ländern Europas oder Nordamerikas gibt es keine vergleichbaren Projekte zu Erhebung und Speicherung von individuellen Daten. Allerdings haben die Enthüllungen von Geheimdienstpraktiken der US-amerikanischen NSA und des Britischen GCHQ durch Edward Snowden gezeigt, dass die staatlichen Fähigkeiten zur Sammlung und Verknüpfung persönlicher Daten aus den verschiedensten Lebensbereichen schon weit fortgeschritten sind, wenngleich auch bisher keine Programme existieren, diese Daten in ähnlicher Weise zur Kontrolle der eigenen Bevölkerung zu nutzen.

Am weitesten fortgeschritten ist der Einsatz von *big data* in Deutschland in der Sicherheitspolitik. Ein Großteil der Bevölkerung – alle, die mit dem Flugzeug reisen – wird mit Verfahren der Mustererkennung darauf untersucht, ob sie möglicherweise ein Risiko darstellen, an Terrorismus oder schwerer Kriminalität beteiligt zu sein. Zu diesem Zweck werten Sicherheitsbehörden die sogenannten „Fluggastdaten“ aus, die von Fluggesellschaften und Reiseveranstaltern erhoben werden müssen. Aktuell werden diese noch mit Datenbanken über verdächtige Personen abgeglichen; perspektivisch sollen aber Verfahren eingesetzt werden, die durch maschinelles Lernen Muster erkennen. Anders als sonst in der Ermittlungsarbeit üblich, werden bei

Fluggästen keine Daten zur strafrechtlichen Vorgeschichte erhoben, sondern eine lange Liste an Konsum-, Verhaltens- und Netzwerkdaten. Analog zur entsprechenden Praxis in den USA und anderen angelsächsischen Ländern sollen die durch *big data* ermittelten Muster dazu dienen, allen Fluggästen einen Risikowert zuzuschreiben. Auf dessen Grundlage werden sodann unterschiedliche Regeln angewendet: Manche Personen werden eingehender überwacht, ohne etwas davon zu bemerken, andere werden an Flughäfen peniblen Kontrollen ausgesetzt; wieder andere werden am Besteigen des Flugzeugs gehindert. Ein Teil der Flugreisenden, jene mit geringem Risikowert, können wiederum mit einer schnellen Abfertigung rechnen. Hier zeigt sich, wie anhand von *big data* eine Granularisierung von Regeln entsteht. Auch früher wurden Fluggäste an Flughäfen unterschiedlich behandelt. Neu ist, dass große Bevölkerungsgruppen geprüft werden und dass der Risikowert einer Person aus Datenspuren entsteht, die diese in ganz anderen Kontexten hinterlassen hat.

Mit Blick auf Meta-Regeln, also parlamentarische Gesetzgebungsverfahren, Entscheidungsverfahren innerhalb von Verwaltungseinheiten, politischen Parteien oder Interessengruppen, spielt *big data* bislang noch eine untergeordnete Rolle. Unter dem Titel *legal technology* kommt zwar zunehmend Software zum Einsatz, die Juristen beim Verfassen und Prüfen von rechtlichen Normen helfen soll. Dabei geht es aber (noch) nicht um die algorithmisch gesteuerte Definition der Inhalte als vielmehr um die interne und externe Kohärenz rechtlicher Normen, etwa um Verweise auf andere Regeltexte. Allerdings zeigt sich der Markt dynamisch und es werden Ideen und Technologien entwickelt, in denen Mustererkennung und maschinelles Lernen Normentwicklung und -interpretation unterstützen oder gar übernehmen sollen, die *legal technologies 3.0* (Goodenough 2015).

Während die formalen Verfahren der Regelsetzung durch *big data* noch nicht beeinflusst werden, gibt es Hinweise, dass dies bei informalen Verfahren bereits anders aussieht. Bei der inhaltlichen Vorbereitung von politischen Entscheidungen bedienen sich staatliche Behörden, Parteien sowie Interessengruppen zunehmend Analyseverfahren, die auf *big data* beruhen (Kreiss 2016; Karpf 2016). Ziel der Analysen ist es, Probleme zu identifizieren, politische Präferenzen zu erfassen, die Auswirkungen politischer Interventionen zu messen etc. Diese Verfahren der *people analytics*, des *micro-targeting* und der *sentiment analysis* ersetzen andere Verfahren der Bevölkerungs- und Problemanalyse nicht, treten jedoch hinzu und nehmen so Einfluss auf Regelsetzungsverfahren (siehe auch den Artikel über „Data Mining“ in diesem Handbuch). Auch in Deutschland gehören sie ins Repertoire politischer Akteure (Kruschinski und Haller 2017).

## 4 Perspektiven für Staat und Verwaltung

Bisher sind es vor allem privatwirtschaftliche Akteure, die versuchen mit Hilfe von *big data* das Verhalten ihrer Kund*innen zu verstehen und zu beeinflussen. Politik und Verwaltungen demokratischer Staaten greifen bisher noch recht zurückhaltend auf große und diverse Datensätze und die mit *big data* assoziierten Analysemethoden (Verfahren des maschinellen Lernens) zurück. Parteien nutzen insbesondere in

den USA schon extensiv *big data* zur Identifizierung und Beeinflussung wechselbereiter Wähler*innen. Einzelne autoritäre Regime bauen aktuell weitreichende Programme zur Nutzung großer und vielfältiger Datenmengen zur umfassenden Kontrolle der eigenen Bevölkerung auf. Grundsätzlich ist der staatliche Einsatz von *big data* im Bereich der inneren und äußeren Sicherheitspolitik am weitesten fortgeschritten. Der Grund dafür ist zum einen die lange Tradition des Einsatzes mustererkennender Verfahren wie beispielsweise der Rasterfahndung und zum anderen die in diesen Bereichen leichter aufrechtzuerhaltende Fiktion, die umfassende Auswertung personenbezogener Daten würde nur einen kleinen, devianten Teil der Bevölkerung betreffen. In Deutschland ist die Sicherheitspolitik das einzige Politikfeld, in dem *big data-basierte* Analysen bisher regelmäßig angewandt werden. Der gemeinsame Nenner algorithmischer Regulierung im staatlichen Bereich ist es, Regeln stärker zu differenzieren: nach geografischen Gebieten oder nach Subpopulationen. Die entsprechende Praxis von Überwachung, Disziplinierung und *social sorting* wird bereits kritisch durch sozialwissenschaftliche Forschung begleitet.

Die Erwartung, dass *big data* im staatlichen Bereich zu einer Renaissance kybernetischer Steuerung führen würde (Mayer-Schönberger und Cukier 2013), hat sich indes als überzogen erwiesen. Beispiele für vollautomatisierte algorithmische Regulierung gibt es bislang nicht, stets gibt es einen „menschlichen Filter" (*human in the loop*). Es gibt freilich Kontexte, in denen die Anreize so gesetzt sind, dass die menschlichen Entscheider*innen sich *de facto* kaum über die maschinelle Empfehlung hinwegsetzen können.

Mit Blick auf die Rolle von *big data* für Meta-Regeln zeigt sich deutlich, dass *big data*-basierte Verfahren andere Verfahren der Evidenzproduktion bislang nicht verdrängen, sondern nur hinzutreten. Somit ist ihr Einfluss begrenzt und es entsteht Raum dafür, ihre Vor- und Nachteile zu erproben. Unterschiedliche Strukturen der (Sozial-)Staatlichkeit fördern oder hemmen den Einsatz von *big data*. In Deutschland wirkt beispielsweise der umfassend ausgebaute Wohlfahrtsstaat einer Ausbreitung privatwirtschaftlicher *scoring*-Verfahren entgegen (Fourcade und Healy 2017). Zudem setzen strenge Datenschutzregeln der Verknüpfung und Zweckentfremdung von Daten in Deutschland enge Grenzen für die Verwendung von *big data* durch staatliche Akteure. Jenseits der Sicherheitspolitik erproben deutschsprachige Länder algorithmische Regulierung in Form prädiktiver Scores im öffentlichen Sektor, etwa zur Prävention von Langzeitarbeitslosigkeit (Österreich), häuslicher Gewalt (Schweiz) und Einbruchdiebstahl (Deutschland, Österreich, Schweiz). Unspezifische Daten kommen jedoch bisher nicht zum Einsatz. Da mit der weiteren Ausbreitung vernetzter Sensoren (*Internet of Things*) die Verfügbarkeit personenbezogener, nutzergenerierter und domainunspezifischer Daten noch deutlich zunehmen wird, hängt die weitere Entwicklung des Einsatzes von *big data* in Staat und Verwaltung wesentlich davon ab, ob gleichzeitig auch Schranken der Verwendung dieser Daten etabliert und implementiert werden, wie sie beispielsweise in der europäischen Datenschutzgrundverordnung enthalten sind.

## Literatur

Angwin, Julia, Jeff Larson, Surya Mattu, und Lauren Kirchner. 2016. Machine Bias. *ProPublica*. www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. Zugegriffen am 9.01.2019.

boyd, danah, und Kate Crawford. 2012. Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society* 15(5): 662–679.

Creemers, Rogier. 2018. *China's social credit system. An evolving practice of control*. https://ssrn.com/abstract=3175792. Zugegriffen am 17.05.2018.

Fourcade, Marion, und Kieran Healy. 2017. Categories all the way down. *Historical Social Research* 42(1): 286–296.

Goodenough, Oliver R. 2015. Legal technology 3.0. *Huffington Post, 04.02.2015,* www.huffingtonpost.com/oliver-r-goodenough/legal-technology-30_b_6603658.html. Zugegriffen am 17.01.2019.

Karpf, David. 2016. *Analytic activism. Digital listening and the new political strategy*. New York: Oxford University Press.

Kreiss, Daniel. 2016. *Prototype politics. Technology-intensive campaigning and the data of democracy*. New York: Oxford University Press.

Kruschinski, Simon, und André Haller. 2017. Restrictions on data-driven political micro-targeting in Germany. *Internet Policy Review* 6(4). https://doi.org/10.14763/2017.4.780.

Mayer-Schönberger, Viktor, und Kenneth Cukier. 2013. *Big data. A revolution that will transform how we live, work and think*. London: John Murray.

Mckinlay, Steve T. 2017. Evidence, explanation and predictive data modelling. *Philosophy & Technology* 30(4): 461–473.

Mistreanu, Simina. 2018. Life inside china's social credit laboratory. *Foreign Policy*. https://foreignpolicy.com/2018/04/03/life-inside-chinas-social-credit-laboratory/. Zugegriffen am 15.05.2018.

Morozov, Evgeny. 2014. The rise of data and the death of politics. *The Observer*, 19. Juli 2014. https://www.theguardian.com/technology/2014/jul/20/rise-of-data-death-of-politics-evgeny-morozov-algorithmic-regulation. Zugegriffen am 09.01.2019.

State Council. 2014. *Planning outline for the construction of a social credit system (2014–2020)*, GF No. (2014)21, Beijing: State Council of the People's Republic of China. China Copyright and Media. https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/. Zugegriffen am 15.05.2018.

Ulbricht, Lena, Sebastian Haunss, Jeanette Hofmann, Ulrike Klinger, Jan-Hendrik Passoth, Christian Pentzold, Ingrid Schneider, Holger Straßheim, und Jan-Peter Voß. 2018. Dimensionen von Big Data. Eine politikwissenschaftliche Systematisierung. In *Big Data und Gesellschaft. Eine multidisziplinäre Annäherung*, Hrsg. Barbara Kolany-Raiser, Reinhard Heil, Carsten Orwat und Thomas Hoeren, 151–231. Wiesbaden: Springer VS.

Yeung, Karen. 2017. ‚Hypernudge'. Big data as a mode of regulation by design. *Information, Communication & Society* 20(1): 118–136.