Körber Stiftung

Deutscher Studienpreis | 1. Preis Geistes- und Kulturwissenschaften

Selbstauferlegte Gedankenlosigkeit der algorithmischen Wende? Kriminalwissenschaftliche Untersuchung über die Automatisierung der Kriminalprognose

Verbrechen verhindern, bevor sie geschehen? Dieses verlockende Narrativ ist nicht länger nur Stoff des Science-Fiction-Genres. Computergestützte Systeme mit dem Ziel, kriminelles Verhalten vorauszusagen, werden heute auch in Deutschland eingesetzt. Die Arbeit geht den gesamt-gesellschaftlichen Auswirkungen dieser Technologie mit interdisziplinärem Zugang (rechts-wissenschaftlich, kriminologisch, soziologisch) auf den Grund. Sie kommt zu dem Ergebnis, dass die Technologie eine Aushöhlung einer Vielzahl verfassungsrechtlicher Garantien mit sich bringt und selbst bei Einhaltung eines engen rechtlichen Rahmens Risiken für Rechtsstaat und Gesellschaft bestehen. Insbesondere droht sich die Kriminalitäts-kontrolle bei unreflektiertem Einsatz der Technologie in eine – in An-lehnung an Hannah Arendt – selbstauferlegte algorithmische Gedanken-losigkeit zu begeben. Die Arbeit schließt mit einem Vorschlag für Mindestanforderungen, an denen sich computergestützte Kriminal-prognosen in Zukunft orientieren sollten.

Lucia Sommerer promovierte an der Georg-August-Universität Göttingen im Fachgebiet Rechtswissenschaften.

Körber Stiftung

Der vorliegende Beitrag wurde beim Deutschen Studienpreis 2020 mit dem 1. Preis in der Sektion Geistes- und Kulturwissenschaften ausgezeichnet. Er beruht auf der 2019 an der Georg-August-Universität Göttingen eingereichten Dissertation »Personenbezogenes Predictive Policing: Kriminalwissenschaftliche Untersuchung über die Automatisierung der Kriminalprognose« von Dr. Lucia Sommerer.

Selbstauferlegte Gedankenlosigkeit der algorithmischen Wende? Kriminalwissenschaftliche Untersuchung über die Automatisierung der Kriminalprognose

I. Hintergrund

»Scientia potestas est« –
»Wissen ist Macht«
Francis Bacon

Inwieweit ist unsere Zukunft vorbestimmt? Kann Vorbestimmtes auch vorausgesagt werden? Dies sind Fragen, die die Menschheit seit jeher umtreiben. Von der Konsultation des Delphischen Orakels der Antike bis hin zum modernen datengestützten Staat – stets wollen wir wissen, was die Zukunft für uns bereithält. Zur Beantwortung dieser Frage in ihrer jüngsten Ausformung wird nun seit einiger Zeit auf prädiktive Algorithmen zurückgegriffen. Prädiktive Algorithmen, d.h. computergestützte Systeme mit dem Ziel, menschliches Verhalten vorauszusagen, sind in unserem Alltag allgegenwärtig. Jeder, der eine Suchmaschine verwendet, sich um einen Kredit bewirbt oder online ein Produkt kauft, setzt sich prädiktiven Algorithmen aus. Nun werden prädiktive Algorithmen von der Polizei auch zur Vorhersage individuellen kriminellen Verhaltens herangezogen (sog. personenbezogenes Predictive Policing), eine Entwicklung, die angesichts ihres transformativen Potentials als »algorithmische Wende« der Kriminalitätskontrolle bezeichnet werden kann. Der narrative Rahmen, Verbrechen zu unterbinden, bevor sie geschehen, ist verlockend. Doch darf er Wissenschaft und Praxis nicht blind machen für die tatsächlichen Beschränkungen prädiktiver Algorithmen und die Risiken dieser Form der »Unsicherheitsabsorption« für Gesellschaft und Rechtsstaat.

Die rechtliche Zulässigkeit des personenbezogenen Predictive Policing in Deutschland ist äußerst kritisch zu bewerten (II.). Selbst bei Einhaltung enger rechtlicher Grenzen wohnen dem Einsatz der neuen Technologien im Interesse von Effizienz und Neutralität bisher wenig beachtete gesamtgesellschaftliche Risiken inne, die Kriminalitätskontrolle und Gesellschaft vor grundlegend neue Herausforderungen stellen (III.). Insbesondere droht die Gesamtheit der Herausforderungen bei unreflektiertem Einsatz und Ausdehnung der Technologie die Kriminalitätskontrolle in eine – in Anlehnung an Hannah Arendt – »selbstauferlegte (algorithmische) Gedankenlosigkeit« zu führen. Um diesen Risiken zu begegnen, hat die vorliegende Arbeit einen Katalog an Mindestanforderungen für die Praxis entwickelt, an denen sich

personenbezogenes Predictive Policing in Zukunft orientieren sollte (IV.).

Die Aktualität der Thematik kann dabei nicht überbewertet werden. Seit 2017 erlaubt das Fluggastdatengesetz mit dem sog. Fluggastdatenmusterabgleich das personenbezogene Predictive Policing in Deutschland. Ebenfalls seit 2017 setzt das Bundeskriminalamt mit dem System RADAR-iTEeine Vorstufe des personenbezogenen Predictive Policing im Bereich des islamistischen Terrorismus ein, und 2018 änderte Hessen sogar sein Polizeigesetz, um den Einsatz der Software Palantir Gotham zu ermöglichen, die in den USA bereits zum personenbezogenen Predictive Policing verwendet wurde. Keines der Projekte konnte bei seiner Einführung auf eine umfängliche juristisch-kriminologische Aufbereitung in der deutschen Wissenschaftslandschaft zurückblicken. Durch die Einführung einer neuen Technologie ohne vorangegangenen und begleitenden wissenschaftlichen Diskurs können sich in die Technologie und ihre Verwendung in der Praxis jedoch problematische Modalitäten und Grundannahmen einschleifen, denen im Nachhinein nur noch schwer beizukommen ist. Es geht dabei nicht nur darum, Rechtskonformität der Technologie sicherzustellen, sondern zur demokratischen Diskussion zu stellen, welche Wertentscheidungen überhaupt in einer Technologie verkörpert werden sollen. Ist eine Technologie einmal eingeführt, kann eine dann erst entstehende Diskussion nicht mehr in gleichem Maße auf die Ausgestaltung des Programms Einfluss nehmen. Entscheidende Weichenstellungen sind dann bereits ohne Berücksichtigung rechtlich-kriminologischer Erwägungen getroffen. Einer Überrumpelung der Gesellschaft, eben ohne wissenschaftliche Bearbeitung und diskursiven Prozess im Vorfeld, auf dem Gebiet des personenbezogenen Predictive Policing möchte diese Arbeit vorbeugen.

II. Rechtliche Grenzen des personenbezogenen Predictive Policing

Wie so viele erfolgreiche Aphorismen ist auch Bacons Eingangszitat über Wissen und Macht v.a. Abbreviatur. Wissensbestände als solche mögen latente Machtchancen darstellen, doch die Verbindung von »Imperium und Empirie« ist keine selbstverständliche, sondern beruht auf gezielten Aneignungs-, Verarbeitungs- und Anwendungsprozessen. Im Sinne einer demokratisch-distributiven Machtorganisation bedürfen gerade diese konkreten Prozesse rechtlicher Begleitung und Einhegung: Aus dem durch die Verfassung garantierten Recht auf informationelle Selbstbestimmung folgt dabei, dass personenbezogenes Predictive Policing, das eine Vielzahl sensibler Daten über Bürger verarbeitet, nur bei Vorliegen einer konkreten Gefahr und nur zum Schutz hochrangiger Rechtsgüter eingesetzt werden darf, sowie, dass ausreichend Kontrollen die Rechts-konformität des Verfahrens sicherzustellen haben.

Aus dem Rechtsstaatsprinzip und dem Menschenwürdekern des Rechts auf informationelle Selbstbestimmung, der es untersagt, Menschen zum bloßen Objekt einer Maschine zu machen, fließt zudem ein Verbot der Verwendung systemimmanent intrans-

parenter Algorithmen zur Kriminalitätskontrolle, d.h. der Verwendung von Algorithmen, deren innere Zusammenhänge selbst von Experten im Nachhinein nicht mehr im Detail nachvollzogen werden können. Dies schließt somit z.B. den zukünftigen Einsatz neuronaler Netze, einer besonders intransparenten Form künstlicher Intelligenz, für personenbezogenes Predictive Policing aus.

Personenbezogenes Predictive Policing ist schließlich entgegen weitverbreiteter Annahmen nicht *von Natur aus* neutraler und objektiver als Entscheidungen menschlicher Polizeibeamter, sondern reproduziert bestehende gesellschaftliche Vorurteile, z.B. gegenüber Minderheiten, wenn diese Vorurteile in den Daten, die dem jeweiligen Algorithmus zugrunde liegen, nicht erkannt und beseitigt werden. Mit Blick auf eine Verletzung des verfassungsrechtlichen Diskriminierungsverbots besteht angesichts der Komplexität und der Intransparenz der verwendeten algorithmischen Systeme die Gefahr, dass der Überprüfungsmaßstab der Gerichte zu einer bloßen Willkürkontrolle verkümmert, ein Überprüfungsmaßstab also, der hinter der gegenwärtigen Intensität gerichtlicher Kontrollen zurückbleibt.

Insgesamt stellt personenbezogenes Predictive Policing damit ein Risiko im Sinne einer Aushöhlung einer Vielzahl verfassungsrechtlicher Garantien dar. Es kann deshalb in Deutschland nur in engem rechtlichem Rahmen und unter strenger Einhaltung der in IV. zu skizzierenden Transparenz- und Kontrollinfrastrukturen gestattet werden.

III. Gesamtgesellschaftliche Risiken der »Algorithmischen Wende«

Überblick

Selbst bei Einhaltung des rechtlichen Rahmens wohnen personenbezogenem Predictive Policing und der dadurch angestoßenen »algorithmischen Wende« der Kriminalitätskontrolle Risiken für Gesellschaft und Rechtsstaat inne. Die neuen Herausforderungen durchziehen die Kriminalitätskontrolle dabei auf verschiedenen Ebenen. Sie zeigen sich zum einen auf der Ebene des Strafrechts. Dort kann es durch die neuen datensammelnden und -auswertenden Technologien zu einer Fokusverschiebung von der konkreten Tat einer Person hin auf eine allgemein verwerfliche »Lebensführung« einer Person kommen und zum Wiederaufleben der im Dritten Reich beliebten Figur der »Lebensführungsschuld«, d.h. die ausschlaggebende Berücksichtigung des Lebenswandels einer Person bei der Bestimmung ihrer »Strafwürdigkeit«. Dies droht in eine unzulässige Allgemeinabrechnung mit der gesamten Lebensführung und einer »Durchleuchtung« der Bürger zu münden. Zum anderen manifestiert sich die »algorithmische Wende« auch auf der Ebene der gesellschaftlichen Wahrnehmung von Polizeiarbeit. Dort kann es – wie erste Studien andeuten – durch einen von Bürgern empfundenen Mangel an Verfahrensgerechtigkeit intransparenter prädiktiver Algorithmen zu einem Legitimitätsverlust der Polizei kommen. Auf Ebene der Institutionen droht zudem ein Relevanzverlust der Kriminologie gegenüber

den Computerwissenschaften. Damit verbunden treten Fragen nach gesellschaftlichen Ursachen kriminellen Verhaltens hinter bloßen statistischen Korrelationen zurück. Um die technische Entwicklung computerwissenschaftlich fundiert, kritisch und aus unabhängiger Perspektive begleiten zu können, muss die Kriminologie sich detailliert mit Prozessen maschinellen Lernens auseinandersetzen und eine neue Subdisziplin der »Digitalen Kriminologie« ausbilden, wozu diese Arbeit einen ersten Anstoß leisten möchte.
Die Verwendung von prädiktiven Algorithmen befördert schließlich und am zentralsten – in Anlehnung an Hannah Arendt – eine »selbstauferlegte Gedankenlosigkeit« der Kriminalitätskontrolle, d.h. eine Situation, in der der Mensch das eigene Denken ausblendet und sich, eingebunden in ein hierarchisches System, völlig auf die Entscheidungen anderer verlässt, ohne diese selbst zu hinterfragen oder an eigenen Wertmaßstäben zu messen.

Insbesondere »selbstauferlegte Gedankenlosigkeit«

Mit dem Begriff der »Gedankenlosigkeit« beschrieb Arendt, wie gewöhnliche Menschen im Dritten Reich durch das Ausschalten ihres selbständigen Denkens Kriegsverbrechen begehen konnten, ohne dezidiert »böse« Absichten zu haben. Wesentlicher Faktor für das Zustandekommen einer solchen Haltung war die Einbindung in einen bürokratischen Apparat. Der NS-Kriegsverbrecher Eichmann etwa berief sich wiederholt darauf, lediglich Anweisungen befolgt zu haben. Die große Gefahr sah Arendt dabei in der Unfähigkeit von Menschen, über die Tragweite ihres eigenen Tuns nachzudenken. Diese Unfähigkeit könne unter gewissen Umständen nahezu jeden durchschnittlichen Menschen befallen, worin Arendt gerade die »Banalität« des Bösen sieht.
Es soll hier prädiktiven Algorithmen selbstverständlich nicht unterstellt werden, dass sie linear zu Verbrechen gegen die Menschlichkeit führen. Und dennoch ist das Konzept der »Gedankenlosigkeit« auch im algorithmischen Kontext nützlich, da es zum Ausdruck bringt, wie sich Menschen in einem System auf die Entscheidungen anderer verlassen, diese nicht hinterfragen, sie schlicht befolgen. Als Rechtfertigung wird auf die höhere Autorität und das Bedürfnis der Regelbefolgung im Interesse des Funktionierens des Systems verwiesen. Diese Situation ist mit dem Umgang von Menschen mit dem Ergebnis algorithmischer Berechnungen durchaus vergleichbar.

Obwohl algorithmengestützte Systeme zunächst nur als ein dem Nutzer untergeordnetes Werkzeug konzipiert wurden, kann ihnen in der Praxis eine umfassendere Bedeutung zugemessen werden. Sie können eine Rolle ähnlich einer dem Nutzer übergeordneten Autoritätsfigur annehmen, wie die eines Vorgesetzten, dessen »Anordnungen« ohne Hinterfragen ausgeführt werden. Studien zu Entscheidungsunterstützungssystemen im medizinischen Bereich sowie bei Flugzeugpiloten haben gezeigt, dass es Menschen sehr schwerfällt, sich gegen das Ergebnis algorithmischer Berechnungen (ähnlich derer

des personenbezogenen Predictive Policing) zu entscheiden, und zwar selbst dann, wenn das Ergebnis des Algorithmus eigentlich nur einen von mehreren Entscheidungsfaktoren darstellen sollte. Dieses als »Automation Bias« bezeichnete und empirisch erforschte Phänomen führt dazu, dass Menschen es unterlassen, zusätzlich zu einem algorithmischen Ergebnis selbständig Informationen einzuholen und zu bewerten, ja sogar deutlich gegen das Ergebnis des Algorithmus sprechende Anhaltspunkte bewusst ignorieren. Der eigenen Expertise wird dabei weniger vertraut als dem Ergebnis des komplexen, undurchsichtigen algorithmischen Prozesses, das dem Menschen mit der gesetzten Selbstverständlichkeit des Faktischen entgegentritt. Dies gilt umso mehr, wenn mit Blick auf Zeitdruck und Rationalisierung die Entscheidung gegen »die Maschine« einen erhöhten Zeit- und Rechtfertigungsaufwand mit sich bringt als die Entscheidung mit »der Maschine«. Im Ergebnis führt dies dazu, dass algorithmische Berechnungen, die lediglich als Entscheidungsunterstützung (bloßes Werkzeug) gedacht waren, faktisch die Entscheidung des Menschen völlig determinieren und der Mensch die eigene Verantwortung an die algorithmische Vorgabe auslagert. In dieser Konstellation wird letztendlich der Algorithmus mit seiner vermeintlich sicheren Wahrscheinlichkeit zur entscheidenden Autoritätsfigur. Der menschliche Handelnde gibt sich angesichts der Zumutung einer Entscheidung einer »Gedankenlosigkeit« anheim. Unsicherheit fordert vom Menschen normative Entscheidungen, die er sich durch die Gewissheit, die der Algorithmus zu bieten scheint, abnehmen lässt. Je stärker ein Prozess automatisiert wird, desto leichter kann es bei Menschen zu Gedankenlosigkeit und Gleichgültigkeit gegenüber seinen Ergebnissen kommen. Je stärker die Kriminalitätskontrolle automatisiert wird, desto leichter ist es für die das System nutzenden Beamten, sich für auf Grundlage des Systems vollzogene Handlungen nicht mehr verantwortlich zu fühlen.

Es ist somit festzuhalten, dass algorithmengestützte Entscheidungssysteme der Logik ihrer Struktur nach dazu neigen, die Gedankenlosigkeit der Nutzer zu fördern, d.h. dazu führen, dass Nutzer keine umfassende Abwägung aller Faktoren einer Situation vornehmen und die Folgen ihrer Handlungen nicht vollumfänglich bedenken. Selbstauferlegt ist diese Gedankenlosigkeit insofern zu nennen, als wir uns als Gesellschaft in der »algorithmischen Wende« durch die Einführung von personenbezogenem Predictive Policing sehenden Auges selbst in ein System begeben, das Gedankenlosigkeit fördert.

IV. Mindestanforderungen: Transparenz-Trias

Den mit personenbezogenem Predictive Policing einhergehenden Risiken für das Recht auf informationelle Selbstbestimmung, das Diskriminierungsverbot und das Rechtsstaatsprinzip sowie für die Gesellschaft als Ganzes muss durch neue Transparenzinfrastrukturen entgegengetreten werden. Diese neuen Strukturen müssen dabei

Entwicklungs- wie auch Einsatzmodalitäten prädiktiver Algorithmen umfassen. Es ist von zentraler Bedeutung, eine bewusste Ausgestaltung algorithmischer Systeme nicht erst im Stadium des Einsatzes in Angriff zu nehmen, sondern bereits das Entwicklungsstadium zu lenken. Denn die wichtigsten Weichenstellungen eines algorithmischen Systems finden im Entwicklungsstadium statt.

Die neuen Transparenzinfrastrukturen haben dabei eine Transparenz-Trias von öffentlichen Registrierungspflichten, subjektiven Betroffenenrechten und staatlichen Kontrollstellen zu umfassen. Erstens sind auf dem Wege der Protokollierung festzuhaltende Entscheidungen der Designphase eines Algorithmus sowie seine abstrakten Wirkprinzipien noch vor Einsatz des Algorithmus in öffentlich einsehbarer Weise bei einer staatlichen Registrierungsstelle zu hinterlegen. Zweitens ist dem einzelnen Betroffenen einer algorithmischen Analyse Zugang zu einer Begründung seines spezifischen Ergebnisses zu geben. Und Drittens hat, zusätzlich zu diesen Offenlegungspflichten, als Kernstück der neuen Transparenzinfrastrukturen eine systematische staatliche Kontrolle von personenbezogenem Predictive Policing stattzufinden, sowohl vor seinem ersten Einsatz als auch in regelmäßigen Abständen während des Einsatzes. Eine neu zu schaffende staatliche Kontrollstelle hat dabei den Algorithmus systematisch u.a. auf Verletzungen des Rechts auf informationelle Selbstbestimmung und des Diskriminierungsverbotes zu untersuchen. Darüber hinaus hat die Kontrollstelle eigene Standards, u.a. für Prognosegenauigkeit und Fehlerrate (wie viele fälschlich als »hochgefährlich« eingestufte Menschen sind akzeptabel, um eine tatsächlich »hochgefährliche« Person durch den Algorithmus aufzuspüren?), festzulegen und deren Einhaltung zu überprüfen.

Die weitere Entwicklung von personenbezogenen Predictive Policing-Systemen in Deutschland sollten erst nach Schaffung solch einheitlicher Standards vorgenommen werden, denn erst die Ausarbeitung von präzisen Qualitätsstandards macht eine klare Sanktionierung mangelhafter Systeme möglich. Hersteller von personenbezogenen Predictive Policing-Systemen können ihr Modelltraining nur unter Einhaltung entsprechender Kriterien optimieren, wenn diese zuvor einheitlich festgelegt wurden.

Diese zentrale Mindestanforderung einer Transparenz-Trias sowie zahlreiche weitere Mindestanforderungen – wie etwa zu technischen Maßnahmen, um die Unvoreingenommenheit des Algorithmus sicherzustellen, oder zur Auswahl der Datengrundlage eines Algorithmus – hat meine Arbeit in einer kompakten, rechtlich fundierten »Checkliste« zusammengefasst, die sich an computerwissenschaftliche Entwickler, Entscheidungsträger in Polizei und Politik sowie Gerichte richtet. Die »Checkliste« kann bei der Entscheidung über Entwicklung, Einsatz sowie Überprüfung der Rechtskonformität von Predictive Policing-Projekten in der Praxis unmittelbar herangezogen werden.

Mit den Mindestanforderungen möchte meine Arbeit – im Anschluss an die verfassungsrechtlichen und grund-

lagenorientierten theoretischen Analysen unter (II.) und (III.) – nicht nur Ausgangspunkt einer Neuorientierung der Debatte um Polizeitechnologien sein und erstmals umfassend Handlungsoptionen in der komplexen Balance zwischen Sicherheit und Freiheit in der »algorithmischen Wende« aufzeigen, sondern sie möchte der Praxis durch die Checkliste unmittelbar ein Hilfsmittel an die Hand geben, um so den Wissenstransfer in die anwendende Kriminalpolitik zu sichern.

V. Ausblick

Wie dargelegt, wohnen personenbezogenem Predictive Policing grundlegend transformative Auswirkungen auf die Kriminalitätskontrolle und den Umgang der Gesellschaft mit Kriminalität als Ganzes inne. Eine einmal eingeführte und sich bewährende personenbezogene Predictive Policing-Technologie hat das Potential, auch in andere Bereiche der Kriminalitätskontrolle übertragen zu werden. Ein einmal zu präventiven Zwecken von der Polizei implementiertes System könnte etwa ohne große technische Änderungen zur repressiven Polizeiarbeit zur Aufklärung von Straftaten verwendet oder für Risikoprognosen vor Gericht herangezogen werden. Die detaillierte Auseinandersetzung mit den rechtlichen und tatsächlichen Grenzen prädiktiver Algorithmen bei der Polizei, z.B. mit der Frage nach Kontrollinfrastrukturen, verspricht somit Erkenntnisse, die auch jenseits des Polizeirechts nutzbar sein werden.

Die eben skizzierten Herausforderungen müssen dabei bereits heute adressiert werden. Es mag zwar sein, dass zukünftige Generationen oder zumindest deren wohlsituierte Eliten amüsiert schmunzeln werden, wenn sie auf meine kritischen Analysen einer algorithmengesteuerten Kriminalitätskontrolle im frühen 21. Jahrhundert zurückblicken. Möglicherweise empfinden sie dabei das gleiche Vergnügen, mit dem wir heute Warnungen Platos vor der Technologie des Schreibens oder Trithemius' vor den Auswirkungen der Druckerpresse lesen. Es plagten die Warnenden Sorgen, dass sich der Kontrollverlust, der mit der systematischen und massenweisen Verlagerung von Wissensproduktion und -speicherung von Individuen auf externe Medien verbunden ist, negativ auf Mensch und Gesellschaft auswirken würde.

Ein weniger begünstigtes Segment zukünftiger Gesellschaften, das aus Bürgern bestehen wird, die nicht in der Lage waren, von digitalen Fortschritten zu profitieren, und unter negativen algorithmischen Annahmen über sich zu leiden haben, werden auf die Weichenstellungen dieses Jahrhunderts jedoch möglicherweise vorwurfsvoller zurückblicken und die Frage stellen, warum keine Sicherungen für Rechtsstaatlichkeit und Grundrechtsschutz in die neuen Technologien mit aufgenommen wurden. Die vorgestellte Arbeit über algorithmengesteuerte Systeme in der Kriminalitätskontrolle soll einen ersten Schritt zur Schaffung solcher (rechtlicher und technischer) Sicherungen bereits im Stadium des Algorithmendesigns darstellen.