Künstliche Intelligenz und Strafrechtspflege – eine Orientierung

Von Dr. **Lukas Staffler**, Rechtsanwalt **Oliver Jany**, Zürich*

Die fortlaufende Digitalisierung und Technisierung ermöglicht es der Strafrechtspflege, neue Wege in ihrem Tätigkeitsfeld zu beschreiten. Die Entwicklung schreitet rasant voran. Die Strafrechtswissenschaften sind deshalb aufgerufen, sich mit dem Phänomen der sog. Künstlichen Intelligenz näher zu befassen. Mit diesem Beitrag wird eine Orientierung über diese Entwicklung gegeben und das Anwendungspotential kritisch beleuchtet.

The advancing digitalisation and technologicalisation enables the criminal justice system to break new ground in its field of activity. The development is progressing rapidly. The criminal law sciences are therefore called upon to take a closer look at the phenomenon of so-called artificial intelligence. This article provides an orientation on this development and critically examines its potential for application.

I. Einleitung

Von der Kinoleinwand[1] in die Realität – so könnte man den Siegeszug der Künstlichen Intelligenz betiteln.[2] Der rasante technologische Fortschritt lässt „intelligente" Computersysteme in unseren Alltag Einzug halten. Bereits heute ist es möglich, autonom im Auto gefahren zu werden, während Algorithmen lenken, Straßenschilder erkennen, die Geschwindigkeit und Spur halten, und man selbst mithilfe eines Transkriptionsassistenten eine E-Mail diktiert, die als digitaler Text aufbereitet und übersetzt wird. Dieser neue Alltag wirft neue Fragen auch für die Strafrechtswissenschaften auf, die seit Beginn der Entwicklung als Forschungsschwerpunkt die strafrechtliche Haftung autonomer Fahrassistenten[3] oder robotischer Systeme[4] in den Blick genommen haben.

Dabei verändert sich indes auch der juristische Alltag an sich. Großkanzleien, die nicht nur in Wirtschaftsstrafverfahren oft mit enormen Datenmengen konfrontiert sind, verwenden komplexe Datenanalyse-Tools („e-discovery") wie Relativity, Reveal, nuix, Clearwell sowie Brainspace.[5] Nur so lassen sich die Daten überhaupt noch innerhalb einer wirtschaftlich rentablen Zeit analysieren – wenngleich strafrechtlich geschulte Datenspezialisten den Review-Prozess kompetent begleiten müssen. Doch auch auf Seiten der Strafverfolgung erhält die neue Technologie Einzug in den Arbeitsalltag. Davon zeugt etwa das OSZE-Polizei-Expertentreffen (APEM) von 2019, das sich dem Generalthema der Künstlichen Intelligenz verschrieben hat.[6] Bei der Gesamtschau der Beiträge[7] wird klar, dass die staatlichen Sicherheitsstrukturen des Staates die künstliche Intelligenz in vielfältiger Weise für sich entdecken.

Letztlich scheint es unvermeidbar, dass sich die Strafrechtswissenschaften mit dem Thema der Künstlichen Intelligenz näher befassen. Der vorliegende Beitrag bietet dazu eine Orientierung, indem er drei Themenkomplexe fokussiert: Eine inhaltliche Durchdringung des Begriffs „Künstliche Intelligenz" (unter II.), einen systematischen Überblick zu den Anwendungsfeldern in der Strafrechtspflege (unter III.) und einen Ausblick über Entwicklungen und Herausforderungen im Umgang mit der neuen Technologie (unter IV.).

II. Begriffserklärung

Eine Annäherung an den Begriff „Künstliche Intelligenz" ist deshalb schwierig, weil der Begriff nicht hinreichend determiniert und daher – letztlich aus Mode- oder Marketingzwecken – inflationär gebraucht wird.[8] Viele Computerprogramme, die zu den künstlich intelligenten gezählt werden, sind beides nicht. Das soll anhand der Bedeutung des Wortpaares deutlich werden.

Zwar gibt es keine allgemeingültige Definition von Intelligenz (aus dem lateinischen „intellegere", also „erken-nen", „einsehen" bzw. „wählen zwischen"), doch ist damit umgangssprachlich die Fähigkeit gemeint, aus einer „kogniti-ven Anstrengung" Erkenntnisse abzuleiten und bestenfalls das Handeln danach auszurichten. Da Algorithmen unbestrit-ten nicht zu eigenen tatsächlichen „Erkenntnissen" gelangen, muss es um einen künstlichen, also aus der Natur nachgebil-

* *Dr. Lukas Staffler* ist Habilitand und Rechtsanwalt/Avvocato (RAK Bozen/Bolzano, Italien). Rechtsanwalt *Oliver Jany* ist Doktorand. Die *Verf.* sind Wissenschaftliche Mitarbeiter am Lehrstuhl für Strafrecht und Strafprozessrecht unter Einschluss des internationalen Strafrechts von Prof. *Dr. Frank Meyer*, LL.M. (Yale), an der Universität Zürich.

[1] Vgl. *Irsigler/Orth*, APuZ 2018, 39.

[2] Zur Geschichte der Künstlichen Intelligenz siehe *Burgstaller/Hermann/Lampesberger*, Künstliche Intelligenz, Rechtliches und technisches Grundwissen, 2019, S. 1 ff.

[3] Grundlegend etwa Gless/Seelmann (Hrsg.), Intelligente Agenten und das Recht, 2016; *Lohmann*, Automatisierte Fahrzeuge im Lichte des Schweizer Zulassungs- und Haftungsrechts, 2016; *Zurkinden*, recht 2016, 144; *Schorro*, SchwZStR 2017, 81; *Thommen*, Zeitschrift für Strassenverkehr 2018, 22; *Thommen/Matjaz*, in: Jositsch/Schwarzenegger/Wohlers (Hrsg.), Festschrift für Andreas Donatsch, 2017, S. 189; zur Unterscheidung von autonomen und automatisierten Fahrsystemen siehe *Armbrüster*, ZRP 2017, 83 (83 f., 85 f.).

[4] Grundlegend etwa *Gaede*, Künstliche Intelligenz – Rechte und Strafen für Roboter?, 2019; *Simmler/Markwalder*, Criminal Law Forum 2018, 1; *dies.*, AJP 2017, 171; *dies.*, ZStW 129 (2017), 20.

[5] https://www.brainspace.com/ (4.4.2020) nutzt unter anderem eine spezielle „unsupervised machine learning" – und damit eine Technologie der Künstlichen Intelligenz.

[6] 2019 OSCE Annual Police Experts Meeting, Artificial Intelligence and Law Enforcement: An Ally or an Adversary?

[7] Die Präsentationen sind abrufbar unter https://polis.osce.org/2019APEM (4.4.2020).

[8] Um ein Beispiel zu nennen, ist auf eine „Zahnbürste mit künstlicher Intelligenz" eines bekannten Markenherstellers hinzuweisen.

deten, den natürlichen Denkprozess imitierenden Begriff gehen.[9] Künstliche Intelligenz lässt sich daher nicht als tatsächlich selbstständig denkende Maschine begreifen. Denn Computerprogramme geben nach herkömmlichem Stand der Technik jene Antworten wieder, die als reactio auf eine actio vom Menschen selbst vorgegeben wurde. Die Prüfung der Parameter („wenn"), die zu einer automatisierten Entscheidung („dann") führen, macht die Entscheidung offenkundig nicht intelligent.

Ein Computerprogramm wird auch nicht dadurch „intelligent", dass es durch Sensoren den Input selbstständig wahrnehmen kann.[10] Denn die Datenerfassung ist kein einer „kognitiven Anstrengung" vergleichbarer Prozess, sondern imitiert unsere Sinne. Selbst wenn die Abläufe komplexer und klüger erscheinen, sind die Algorithmen nicht zur Imitation des menschlichen Denkprozesses, zur Wandlung und Adaption imstande.

Von echter „Künstlichen Intelligenz" ist daher nur dann zu sprechen, wenn ein Algorithmus eine Reaktion aus unbekannten (d.h. vom Menschen nicht ex ante eingespeisten) Parametern generiert. In diesen Fällen zeichnet ein Programm nicht nur das nach, was ihm von Menschenhand vorgegeben wurde,[11] sondern kreiert selbstständig Antworten auf bis dato unbekannte Fragen. Diese, das menschlichen Denken imitierende Tätigkeit verbleibt allein als Anwendungsbereich der KI. Viele mit dem Siegel „KI" etikettierte Programme gehören vor diesem Hintergrund eher in den Bereich der Digitalisierung. Der Anwendungsbereich von KI nach dem vorher skizzierten technischen Verständnis ist daher derzeit erheblich geringer, als es den Anschein macht.

Im Folgenden soll noch vertiefter auf die technischen Grundlagen eingegangen werden. Dazu bedarf es Begriffserklärungen zu „Algorithmus" (1.) und „Künstlicher Intelligenz" (2.) sowie zu den beiden, mit KI verbundenen Begriffen „Maschinelles Lernen" (a) und „Deep Learning" (b). Auf diese Weise sollen der KI-Begriff und seine Implikationen entmystifiziert und die Basis für eine sachliche Befassung mit der Thematik geschaffen werden.

1. Algorithmus

Algorithmus beschreibt eine endliche Abfolge von formalen Arbeitsschritten in Form von Regeln (d.h. logischen Operationen und Anweisungen), die es ermöglichen, ein Ergebnis aus der anfänglichen Eingabe von Informationen zu erhalten. Einfach ausgedrückt sind Algorithmen automatisierte Verfahren, die nach vorab präzise definierten Schritten ablaufen, um ein konkretes Ergebnis zu erreichen. Algorithmen folgen dabei deterministischen Mustern, d.h. bei gleicher Eingabe erzielen sie dasselbe Resultat.[12] Legt man Algorithmen Datenbestände zugrunde, so legen sie mittels statistischer Analyse Zusammenhänge zwischen Eingangs- und Ausgangswert offen, die dann als Grundlage für weitergehende Modelle (z.B. Prognosemodelle) genutzt werden können.

Gleichwohl liegt die praktische Schwierigkeit zentral darin, Aspekte der analogen Welt in Form eines Algorithmus darzustellen. Bereits vermeintlich einfache Aufgabenstellungen (z.B. Bilderkennung von Straßenschildern in Abgrenzung zur Umgebung sowie Klassifizierung dieser Schilder) erfordern die Programmierung komplexer Algorithmen. Bei Zugrundelegung eines hinreichend großen und gleichzeitig gekennzeichneten Datensatzes können Computer allerdings eine statistische Funktion erzeugen, um neue Eingangsdateien nach der trainierten Klassifikation entsprechend zuzuordnen.

Bildlich gesprochen stellen Algorithmen das Werkzeug dar, mit dem die Künstliche Intelligenz arbeitet.

2. Künstliche Intelligenz

Künstliche Intelligenz (KI) ist ein Teilgebiet der Informatik. Der Sammelbegriff, der auf die entsprechende Wortschöpfung aus einem Workshop am Darthmouth College in Hanover (New Hampshire) aus dem Jahr 1956 zurückgeht,[13] bezeichnet das automatisierte Verhalten bzw. die Fähigkeit von Maschinen, mittels komplexer Algorithmen menschliches Entscheidungsverhalten „nachzuahmen". Die Maschine trifft ihre Entscheidung nicht mehr aufgrund eines vorgegebenen Protokolls („wenn-dann"-Relation), sondern reaktiv aus einer nahezu endlos großen Menge an Möglichkeiten. Dadurch haben die Maschinen die Fähigkeit, kleinere Abweichungen von der Norm im Entscheidungsfindungsprozess mit einzubeziehen und können auf diese Weise ein größeres Repertoire an Problemlösungen generieren. Die KI-Leistung, Daten und Informationen mittels Algorithmen zu analysieren, ist durch Autonomie charakterisiert. Das System hat also Fähigkeit, in unbekannten (a priori nicht festgelegten) Rahmen im Sinne der festgelegten Zielsetzung zu agieren, indem es die neue Umgebung erfasst und das bisherige Know-How auf der Grundlage des neu generierten Umgebungswissen anpasst.[14] Wesentliche Charakteristik von KI ist, dass diese Systeme durch die autonome Leistung zugrundeliegende Algorithmen eigenständig (d.h. ohne menschliches Zutun) anpassen können. KI-Systeme sind also in der Lage, selbstständig Algorithmen zu entwerfen oder bestehende Algorithmen zu verändern.

Innerhalb des Sammelbegriffs KI unterscheidet man zwischen schwacher und starker KI.[15] Schwache KI bezeichnet

[9] Vgl. https://www.duden.de/rechtschreibung/kuenstlich (4.4.2020).

[10] Vgl. dazu *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 4 ff.

[11] Nach diesem Maßstab ist freilich auch jede Sprechpuppe intelligent, weil sie vorab aufgenommene Sätze akustisch wiedergibt.

[12] *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 14.

[13] *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 3.

[14] *Yuan*, RW 2018, 477 (481); *Fateh-Moghadam*, ZStW 131 (2019), 863 (875 f., 878); a.A. hingegen noch *Hilgendorf*, in: Beck (Hrsg.), Jenseits von Mensch und Maschine, 2012, S. 119 (120 Fn. 2), wonach Autonomie im technischen Sinne die Unabhängigkeit von menschlicher Eingabe im Einzelfall bedeuten soll.

[15] Zur Unterscheidung zwischen schwacher und starker Künstlicher Intelligenz siehe *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 12 f.; *Gaede* (Fn. 4), S. 24; *Fateh-Moghadam*, ZStW 131 (2019), 863 (879); jeweils m.w.N.

die Fähigkeit des KI-Systems, sich fortzuentwickeln und zu „lernen", allerdings nur bereichsspezifisch für das ihr zugewiesenen Aufgabenprofil. Die meisten KI-Systeme entsprechen nach dem heutigen Stand der Technik der schwachen KI, sie treten daher auch nicht grundsätzlich in Konkurrenz zum Menschen, sondern unterstützen ihn in seiner Effizienz und Leistungsfähigkeit.[16] Eine starke KI hingegen ist in der Lage, außerhalb ihres zugewiesenen Aufgabenprofils zu operieren, was ihre Einsatz- und Entwicklungsmöglichkeiten nahezu kaum begrenzen lässt. Da der Zenit einer solchen starken KI die Entwicklung eines eigenen Bewusstseins darstellt, würden solche KI-Systeme tatsächlich in Konkurrenz mit dem Menschen treten.[17]

a) Machine Learning

KI basieren auf autonomen Systemen, deren Leistungsfähigkeit über die linearen „Wenn-dann-Verfahren" hinausgehen. Maschinelles Lernen beschreibt hierbei[18] die Fähigkeit von KI-Systemen, auf der Grundlage großer Datenmengen die zugrundeliegenden Algorithmen einerseits zu entwickeln und diese andererseits zu trainieren, um im Rahmen der bereits vorhandenen Datenauswertung Assoziationen für die Zukunft herzustellen.[19] Unterschieden werden dabei bisweilen fünf verschiedene Analyseformen von Datensätzen, nämlich Klassifizierung (nach Attributen), Regression (statistische Untersuchung der Beziehung zwischen Variablen), Segmentierung (Einteilung der Daten in Gruppen oder Clustern nach ähnlichen Eigenschaften), Assoziation (Offenlegung von Korrelationen zwischen verschiedenen Attributionen) und Sequenzen (Offenlegung von Sequenzen oder Episoden im Datensatz, um Assoziationen oder Muster im Laufe der Zeit zu erkennen).[20] Zentrale Bedeutung kommt dabei der „Übersetzung" von Informationen in eine für die Maschine verstehbare „Sprache" zu. Hierfür gibt es verschiedene Ansätze, wobei – soweit ersichtlich – der sog. Argument-Based-Machine-Learning-Zugang für den Rechtspflegebereich überaus vielversprechend erscheint.[21]

Einsatzgebiete maschinellen Lernens anhand großer Datenmengen sind – vom Datenkontext abhängig – etwa Mustererkennung, Bildauswertung, Sprach-Text-Übertragung oder Prognosemodellerstellung. Der menschliche Einfluss auf die Programmierung bzw. das Training des Algorithmus variiert. Je nachdem, ob der Mensch in das maschinelle Training mittels Justierungen oder Korrekturen eingreift, wird zwischen überwachtem und unüberwachtem maschinellen Lernen unterschieden.[22]

b) Deep Learning

Deep Learning beschreibt eine Methode des maschinellen Lernens, die derzeit überaus beeindruckende Ergebnisse gerade im Umgang mit unstrukturierten Daten bereithält, weshalb diese Technik insbesondere bei modernen Fahrzeugen oder Mobilgeräten eingesetzt wird.[23] Vorbild dieser Technik ist das menschliche Gehirn, das aus Milliarden von Nervenzellen besteht und durch Lernen Verbindungen (Synapsen) zwischen diesen Zellen aufbaut oder abschwächt. Deep Learning versucht, diesen biologischen Vorgang in die Technik zu übertragen.[24]

Die Charakteristik „Deep" leitet sich aus dem Umstand ab, dass das Computersystem die Datenverarbeitung mittels einer (bildlich gesprochen) schichtweisen, tiefergehenden Hierarchie-Bildung vornimmt. Deep-Learning-Systeme lassen sich insofern als künstliche neuronale Netze skizzieren.[25] Die Daten durchlaufen die algorithmische Anwendung in mehreren Phasen. Ergebnisse der ersten Phase maschinellen Lernens werden hierarchisch gestalteten Wertungsprozessen zugeordnet, um dann abermals Prozessen des maschinellen Lernens zugeführt zu werden, in denen die Informationen des vorherigen Prozesses weiterverarbeitet werden. So trainiert der Algorithmus nicht nur auf Basis einer Datengrundlage (Ausgangsdaten), sondern evaluiert und justiert auch die internen Parameter, um auf dieser neuen Grundlage den Datenbestand abermals zu trainieren und so eine qualitative Verbesserung zu entwickeln. Durch das mehrschleifige Verfahren werden die Ausgangsmerkmale von Daten und Algorithmen zunehmend abstrakter.[26] Über die „Verselbstständi-

[16] Nach *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 13, verwendet schwache KI zur Problemlösung Verfahren aus der Mathematik, Informatik und Statistik.

[17] Nach *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 13, weist starke KI höhere intellektuelle Fähigkeiten wie „logisches Denken, strategisches Planen, Lernen, Entscheiden unter unsicherer Informationslage und die ständige flexible Kombination dieser Fähigkeiten" auf.

[18] *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 13, betonen, dass Maschinelles Lernen historisch als KI-Teilgebiet eingeordnet war, sich inzwischen jedoch zu einer ebenbürtigen Disziplin entwickelt hat.

[19] Zusammenfassend bei *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 15: „Maschinelles Lernen konzentriert sich auf generalisierende Lernverfahren, deren Modelle die Fähigkeit besitzen, für noch nie gesehene Inputs einen Output vorherzusagen." Vgl. auch *Beck*, ZIS 2020, 41 (44); *McClendon/Meghanathan*, MLAIJ 2 (2015), 1 (3); sowie ausführlich *Danks*, in: Frankish/Ramsey (Hrsg.), The Cambridge Handbook of Artificial Intelligence, 2012, S. 151.

[20] *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 16 f.; *McClendon/Meghanathan*, MLAIJ 2 (2015), 1 (4 ff.).

[21] *Možina/Žabkar/Bench-Capon/Bratko*, Artificial Intelligence and Law 13 (2006), 53; zum Forschungsschwerpunkt „explainable Artificial Intelligence" (erklärbare KI) vgl. statt vieler *Lepri/Oliver/Letouze/Pentland/Vinck*, Philosophy & Technology 2018, 611 ff. m.w.N.; *Cornelius*, ZIS 2020, 51 (56 f.); *Thieltges*, ZfP 67 (2020), 1 (7).

[22] *Cornelius*, ZIS 2020, 51 (63); *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 17 ff., mit weiteren Differenzierungen.

[23] *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 21.

[24] *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 20; *Cornelius*, ZIS 2020, 51 (52 m.w.N.).

[25] *Goel/Davies*, in: Sternberg/Kaufman (Hrsg.), The Cambridge Handbook of Intelligence, 2012, S. 468 (474); *Quarck*, ZIS 2020, 65.

[26] *LeCun/Bengio/Hinton*, Nature 2015, 436.

gung“ der Algorithmen im Deep-Learning-Prozess wird die Datenverarbeitung nicht mehr rekonstruierbar, die Entstehungsweise der maschinellen Entscheidung kann (wenn überhaupt) nur eingeschränkt validiert bzw. verifiziert werden.

III. Anwendungsfelder in der Strafrechtspflege

Mit Blick auf das vorher dargelegte technische Begriffsverständnis von Künstlicher Intelligenz, wonach es sich um maschinelle Selbstlernfähigkeiten handelt, die jenseits der Ausgangs- bzw. Trainingsdaten neue Problemfelder automatisch erfassen und diese einer Bewertung zuführen können,[27] wird klar, dass deren Anwendungsbereich nach dem heutigen Stand der Technik eng ist. Das zeigt die nachfolgende Analyse von digitalen Hilfsmitteln und Algorithmen für Strafverfolgungsbehörden.

Dazu wird ein systematischer Überblick zu den Einsatzfeldern von digitalen Anwendungen in der Strafrechtspflege gegeben. Die Systematik folgt den chronologischen Gesichtspunkten des Strafverfahrensrechts. Beginnend mit sicherheitspolizeilichen Aspekten (1.) werden Einsatzmöglichkeiten im Ermittlungs- (2.) und Hauptverfahren (3.) besprochen, um schließlich auf die im Schrifttum überaus umstrittene Anwendung im Bereich der Gefährlichkeitsprognose einzugehen (4.).

1. Kriminalitätsprognosen

Ein praktischer Schwerpunkt beim Einsatz digitaler Anwendungen in der Strafrechtspflege ist der Bereich von Kriminalitätsprognosen. In der Fachliteratur hat sich dazu der englische Terminus predictive policing (vorhersagebasierte Polizeiarbeit) durchgesetzt.[28] Im Wesentlichen geht es dabei um die Analyse von Detailinformationen aus Polizeiakten, namentlich Daten zu Personen und ihren Merkmalen (im Rahmen von personenbezogenen Verfahren) und/oder Daten zu Tatzeit und -ort (im Rahmen von raumbezogenen Verfahren). Ziel dieser Analysen ist es, die Steuerung von Polizeieinsätzen im Hinblick auf künftige Straftaten effizienter zu gestalten.[29] Den Algorithmen liegen dabei unterschiedliche kriminal-soziologische Annahmen zugrunde. Ein Beispiel für eine solche Annahme ist der Near-Repeat-Ansatz,[30] wonach bei bestimmten Verbrechensarten (z.B. Wohnungseinbruchsdiebstahl) berechenbare Muster vorhanden sind, weil professionelle Serientäter durch ihre rationale Opferauswahl und Kosten/Nutzen-Abwägung wiederkehrende Handlungsmuster aufweisen, weshalb durch hinreichende Auswertung personen- und ortsbezogener Daten eine Kriminalitätsprognose möglich wird.

Im Folgenden sollen zwei Beispiele der Kriminalitätsprognose dargelegt werden: Die Hot-Spot-Identifikation gehört zu den raumbezogenen Verfahren, das Crime-Linking basiert hingegen auf personenbezogenen Verfahren.

a) Hot-Spot-Identifikation

Die Sicherheitskräfte greifen auf Algorithmen zurück, um aus den zu Verfügung stehenden Daten geografische Muster zu erstellen, in denen statistisch gesehen eine höhere Kriminalitätsrate anfällt. Anhand von vergangenheitsbezogenen Daten sollen statistische Zusammenhänge erhoben werden, um etwaige Muster zu identifizieren.

Nach erfolgter Identifikation derartiger Kriminalitätsmuster kann auf Basis des Algorithmus nicht nur ein entsprechender Hot-Spot identifiziert werden, an welchen Orten zu gewissen Zeiten eine statistisch relevante Häufigkeit von Straftaten auftaucht. Vielmehr sind die Algorithmen in der Lage, eine zukünftige Kriminalitätsprognose im Sinne einer Echtzeitwarnung zu generieren, sodass Polizeikräfte die Auskunft darüber erhalten, dass zu einem bestimmten zukünftigen Zeitpunkt die Wahrscheinlichkeit des Begehens eines Verbrechens statistisch hoch ist. Dementsprechend können personelle Ressourcen der Polizei präventiv am „chronischen“ Kriminalitätsbrennpunkt platziert werden.[31]

Ein bereits etabliertes System, das in mehreren US-amerikanischen Städten praktische Anwendung findet, ist das System RTM (Risk Terrain Modeling).[32] Es handelt sich hierbei um einen Algorithmus, der durch Auswertung verschiedener, auf kriminalitätsbegünstigende Faktoren bezogener Daten geographisch Orte identifiziert, um so eine statistische Vorhersage über die Begehung von Drogendelikten zu

[27] Vgl. auch *Gaede* (Fn. 4), S. 19 f.

[28] Vgl. etwa *Gless*, in: Herzog/Schlothauer/Wohlers/Wolter (Hrsg.), Rechtsstaatlicher Strafprozess und Bürgerrechte: Gedächtnisschrift für Edda Weßlau, 2016, S. 165; *Rademacher*, AöR 142 (2017), 366 (368 ff., 391: der Autor beschreibt das Phänomen als „Verfahren automatisierter Gefahrverdachtserkennung“); *Singelnstein*, NStZ 2018, 1; *Rostalski*, GA 2019, 481 (482 f.); Bode/Seidensticker (Hrsg.), Predictive Policing, Eine Bestandsaufnahme für den deutschsprachigen Raum, 2020.

[29] Für einen Überblick über den Einsatz von Predictive Policing in der Schweiz siehe *Egbert*/*Krasmann*, Predictive Policing, Eine ethnographische Studie neuer Technologien zur Vorhersage von Straftaten und ihre Folgen für die Polizeiliche Praxis, 2019, S. 35, abrufbar unter: https://www.wiso.uni-hamburg.de/fachbereich-sowi/professuren/hentschel/forschung/predictive-policing/egbert-krasmann-2019-predictive-policing-projektabschlussbericht.pdf (4.4.2020); für Deutschland siehe die Studie auf S. 27 ff. sowie jene von *Seidensticker*/*Bode*/*Stoffel*, Predictive Policing in Germany, 2018, S. 2 m.w.N., abrufbar unter http://nbn-resolving.de/urn:nbn:de:bsz:352-2-14sbvox1ik0z06 (4.4.2020).

[30] Vgl. *Singelnstein*, NStZ 2018, 1 (2 Fn. 9) m.w.N.

[31] *Singelnstein*, NStZ 2018, 1 (2 ff.), sieht in Predicting Policing erhebliches Entwicklungspotential, das gerade auch im behördlichen Umgang mit Kriminalität fruchtbar zu machen ist. Gleichwohl macht der Autor auf verschiedene Risiken aufmerksam, etwa die Verdrängung und Verlagerung von Kriminalität oder das Ausnutzen von systembedingten Schwachstellen durch versierte Täter, aber auch die Verzerrung.

[32] http://www.riskterrainmodeling.com/ (4.4.2020).

geben (sog. crime mapping). Kern dieser Vorhersage sind die Daten, mit denen der Algorithmus arbeitet. Es handelt sich hierbei um Größen wie das Vorhandensein schlechter oder nicht-funktionierender Straßenbeleuchtung, die Umgebung von Geldautomaten, Wechselstuben, Parkplätzen, Schulen oder die räumliche Nähe von Nachtclubs, Haltestellen des ÖPNV, Bahnhöfen, Kreuzungen von Hauptverkehrsadern. Durch die Verbindung der Kriminalstatistik mit derartigen geografischen Daten wurde in einigen großen Städten eine Karte angelegt, um die Hot Spots des Drogenhandels zu identifizieren und eine statistische Aussage darüber zu erhalten, wo das Risiko des Drogenhandels hoch ist. Diese Information beeinflusst die kriminalpräventive Strategie der Polizei.

Ein ähnliches System, das aus einem Forschungsprojekt der UCLA und dem Los Angeles Police Department hervorging und bereits seit mehreren Jahren in den USA und in Großbritannien operativ ist,[33] ist PredPol (kurz für: Predictive Policing).[34] PredPol nutzt einen Algorithmus, der verschiedene kriminalstatistische Datensätze miteinander verknüpft, nämlich die Art, den Ort und den Zeitpunkt der Straftat. Auf dieser Grundlage errechnet der Algorithmus die statistische Wahrscheinlichkeit, in welchen örtlichen Gebieten die Begehung spezifischer Straftaten wahrscheinlich ist. Die Software, die nunmehr von einer privaten Firma in Kalifornien entwickelt wird, gilt nicht nur als wirtschaftliches Erfolgsmodell, sondern weist nach Angaben des Unternehmens durchaus beeindruckende kriminalistische Erfolge aus.[35]

Derartige Systeme finden inzwischen auch in Europa Anwendung. So wird in Italien seit kurzem die von der Polizeipräfektur Neapel (unter der Federführung von *Elia Lombardo*) in Kooperation mit zwei Universitäten entwickelte Software XLAW in zahlreichen Polizeidirektionen eingesetzt. Der Algorithmus basiert demnach auf der Annahme, dass bestimmte Verbrechen gegen das Vermögen örtlich und zeitlich wiederkehren, weshalb die Software Polizeiberichte nach kriminellen Mustern durchsucht, um auf diese Weise Prognosen zu erstellen, wo und wann eine Tatbegehung wahrscheinlich ist.[36]

In Deutschland, aber auch in einigen Gebieten der Schweiz, nutzen die Sicherheitskräfte die Software Precobs (kurz für: Pre Crime Observation System), welche durch das Institut für musterbasierte Prognosetechnik in Oberhausen entwickelt wurde. Die Grundidee basiert auch auf der Beobachtung, dass auf gewisse Straftaten (insbesondere Wohnungseinbrüche) in zeitlicher und räumlicher Nähe Folgetaten auftreten, oft durch dieselben Täter(gruppen). Auf Basis dieser Annahmen und unter Heranziehung entsprechender Polizeidaten erhebt die Software Daten über jene Orte, in denen das Begehen einer Folgetat wahrscheinlich ist.[37]

b) Crime Linking

Eine andere Art der Kriminalitätsprognose lässt sich unter dem Oberbegriff crime linking zusammenfassen. Hier geht es um die Verknüpfung von geschehenen Straftaten mit bestimmten (bereits identifizierten oder noch zu identifizierenden) Individuen mit dem Ziel der Vorhersage, wo und wann diese Individuen das nächste Verbrechen begehen. Im Zentrum dieser algorithmenbasierten Analyse stehen also Serientäter mit ihren wiederkehrenden Charakteristika bei der Tatbegehung.

In Großbritannien findet etwa das System HART (kurz für: Harm Assessment Risk Tool) Anwendung, welches von einer Forschungsgruppe an der Universität Cambridge entwickelt wurde.[38] Zweck der Software ist es, Personen nach verschiedenen Risiken für die zukünftige Straftatenbegehung zu klassifizieren. Dazu nutzt der Algorithmus von der Polizei in Durham gesammelte Daten, unter Berücksichtigung von Faktoren wie Vorstrafen, Fluchtrisiko und Schwere des aktuellen Verbrechens. Damit soll der Algorithmus den Sicherheitskräften bei der Entscheidung helfen, wie lange der Gewahrsam eines Verdächtigen dauert bzw. ob eine Freilassung auf Kaution vor oder erst nach der Anklageerhebung gewährt wird.

In Italien wurde durch die Mailänder Polizeidirektion (unter Leitung von Mario Venturi) eine Software namens KeyCrime entwickelt.[39] Der Grundidee des Algorithmus lag die Beobachtung zugrunde, dass eine Vielzahl von Raubüberfällen durch dieselben Täter begangen wird. Die Charakteristika dieser Taten werden individuellen Tätern zugeordnet, um auf diese Weise eine Prognose über Ort, Zeit und Art der zukünftigen Tatbegehung zu erstellen.[40]

2. Ermittlungsverfahren

Im Gegensatz zum polizeilichen Sicherheitsrecht stellt das Ermittlungsverfahren diametral andere Anforderungen an die digitale Anwendung „künstlich intelligenter" Systeme. Bei der präventiven Gefahrenprognose sind Datenmenge und -inhalte frei durch die Ermittler determinierbar. Oft werden daher bereits vorhandene Informationen in Prognosealgo-

[33] *Singelnstein*, NStZ 2018, 1 (2 f.).
[34] https://www.predpol.com/ (4.4.2020).
[35] Vgl die Angaben der Testimonials unter https://www.predpol.com/testimonials/ (4.4.2020).
[36] *Lombardo*, Sicurezza 4P, Lo Studio alla base del software XLAW per prevedere e pervenire i crimini, 2019; vgl. *Di Gennaro*/*Lombardo*/*Marselli*/*Spina*, in: Di Gennaro/Marselli (Hrsg.), Criminalità e sicurezza a Napoli. Secondo rapporto, 2017, S. 175 (200 ff.).
[37] Zur Funktionsweise von Precobs, aber auch zur Wirksamkeit des Programms siehe die Evaluationsstudie des Max-Planck-Instituts für ausländisches und internationales Strafrechts: *Gerstner*, Predictive Policing als Instrument zur Prävention von Wohnungseinbruchdiebstahl, Evaluationsergebnisse zum Baden-Württembergischen Pilotprojekt P4, 2017, S. 19 ff.; vgl. auch *Singelnstein*, NStZ 2018, 1 (1 f.); sowie ausführlich zu den deutschen Politprojekten *Rademacher*, AöR 142 (2017), 366 (369 ff.).
[38] Vgl. *Oswald/Grace/Urwin/Barnes*, Information & Communications Technology Law 27 (2018), 223 (227 f.) m.w.N.
[39] https://www.keycrime.com/ (4.4.2020).
[40] *Akhgar/Saathoff/Arabnia/Hill/Staniforth/Bayerl*, Application of Big Data for National Security, 2015, S. 246.

rithmen eingegeben. Im Ermittlungsverfahren hingegen zwingt der Legalitätsgrundsatz (§ 152 Abs. 2, § 160 StPO) die Behörden zur umfassenden Analyse des gesamten Untersuchungsmaterials, zur Bestimmung oder Wiederlegung des Anfangsverdachtes. Je nach Tatvorwurf ist es leicht vorstellbar, dass sich Staatsanwälte und Strafverteidiger enormen Datenbergen gegenübersehen, die es abzutragen gilt. Folglich ist das Aufgabenprofil einer digitalen Ermittlungssoftware im Bereich des Ermittlungsverfahrens nicht die Datengenerierung (wie bei der Prognose), sondern die Datenreduktion.

a) Datenreduktion

Wie eine solche Ermittlungssoftware eine sinnvolle Datenreduktion vornehmen kann, zeigt ein Forschungsprojekt des Justizministeriums Nordrhein-Westfalen, das gemeinsam mit der Zentral- und Ansprechstelle Cybercrime NRW (ZAC NRW), im Sommer 2019 vorgestellt wurde. Die Software soll mit Hilfe von „Künstlicher Intelligenz“[41] den Kampf gegen Kinderpornographie unterstützen.[42]

Hintergrund dieses Forschungsvorhabens ist der Umstand, dass bei der Ermittlung gegen die Verbreitung von kinderpornographischen Inhalten im Internet Untersuchungsmaterial in unvorstellbaren Mengen[43] anfällt, das derzeit noch manuell (d.h. von Ermittlern) gesichtet werden muss. Die strafrechtlich relevanten Daten gehen wie die Nadel im Heuhaufen harmloser Bilddateien unter. Vor dem Hintergrund der enormen psychischen Belastung, der Beamte bei der Sichtung und Auswertung der Bilder[44] ausgesetzt sind, leuchtet es schnell ein, dass die manuelle Ermittlung sowohl in zeitlicher[45] als auch in personeller[46] Hinsicht wenig zweckmäßig erscheint.

In Kooperation mit Experten aus verschiedenen Branchen wurde daher ein Algorithmus entwickelt, der harmlose Bilder von kinderpornografischem Material unterscheiden kann.[47] Durch die digitalisierte Vorermittlung erhoffen sich die Justizbehörden zum einen eine erhebliche Beschleunigung der Datenauswertung, zum anderen soll die Datenmenge auf die strafrechtlich relevanten Bilddateien reduziert werden. Der reduzierte Datensatz kann dann abschließend von menschlicher Hand analysiert werden. Diese automatisierte Vorerhebung gründet sich bereits auf einen entsprechenden Tatverdacht und fügt sich als zulässige beschleunigungsfördernde Unterstützung in die Tätigkeit der Justiz ein. Das Potential solcher digitalen Ermittlungsvorarbeiten ist, in einem Zeitalter der Digitalisierung mit immer zunehmenden Datenmassen evident, aber erscheint letztlich auch unumgänglich.

b) Folgefragen

Das volle Potential derartiger Systeme kann aber nur genutzt werden, wenn die Ermittlungsmaßnahme auf einem dafür vorbereiteten festen Rechtsfundament mit klaren Regeln für den Einsatz und die Verwendung aufbaut. Ist das – wie derzeit – nicht der Fall, droht die Arbeitsentlastung in einer zum Teil nicht absehbaren Flut von Folgeproblemen unterzugehen.

Das wird im folgenden Beispiel deutlich. Neben datenschutzrechtlichen Bedenken im Rahmen der Entwicklung der Bildbearbeitungssoftware aus NRW stellt sich die Frage, ob und wieweit die Weiterleitung der Daten zur Analyse an Drittunternehmen, als Verbreitung und Besitz kinderpornografischer Schriften unter die §§ 184b ff. StGB fällt. Gelöst wurde das Problem im konkreten Fall durch eine starke Komprimierung der Bilddateien, sodass die Bilder zwar für das Bilderkennungsprogramm verwertbar, für den Menschen aber nicht mehr erkennbar waren und die Rückbearbeitung in den Urzustand unmöglich gemacht wurde.[48]

3. Hauptverfahren

Über die tatsächliche Anwendung von Algorithmen im Hauptverfahren, nämlich zur Beweiserhebung oder zur richterlichen Entscheidung, gibt es aktuell keine gesicherten Informationen. Im Zuge dieser systematischen Darstellung ist gleichwohl auf die dortigen Herausforderungen und Probleme einzugehen, wie sie sich bei einer algorithmenbasierten Unterstützung der richterlichen Tätigkeit ergeben.

a) Beweiserhebung

Im Bereich der Hauptverhandlung stellt sich die Frage, wie die Ergebnisse des digitalisierten Ermittlungsverfahrens Ein-

[41] Wie genau die Software funktioniert, ist unklar. Davon hängt es aber ab, ob die Software in den Bereich digitale Ermittlung oder echte Künstliche Intelligenz fällt. Von letzterem wird man nur dann sprechen können, wenn die Software selbstständig neue Parameter zur Bilderkennung generiert und nicht vorgegebenen Input prüft.

[42] MMR-Aktuell 2019, 419673; weiterführend zu den Auswirkungen der Digitalisierung auf das Ermittlungsverfahren *Schneider*, ZIS 2020, 79 (80 ff.).

[43] Die Pressemitteilung spricht von einer Datenmenge von rund drei Petabyte, was 1.024 Terabyte bzw. 1.048.576 Gigabyte entspricht.

[44] welt.de v. 1.11.2019, abrufbar unter https://www.welt.de/regionales/nrw/article202822224/Seelsorger-ueber-Kindesmissbrauch-Tonspur-kaum-auszuhalten.html (4.4.2020).

[45] Als Beispiel: Wenn sich 100 Mitarbeiter hauptberuflich 40 Stunden pro Woche ohne Urlaub der Bilder widmen würden, würden sie für die Auswertung mehr als sechs Jahre benötigen.

[46] Hier geht es nicht nur um die organisatorischen personellen Kapazitäten. Vielmehr sind der enorme Stressfaktor und die psychische Belastung der Mitarbeiter durch den kinderpornographischen Inhalt auch mit zu bedenken.

[47] Vgl. Pressemitteilung von Microsoft Deutschland v. 5.8.2019, abrufbar unter https://news.microsoft.com/de-de/ki-im-einsatz-gegen-kinderpornografie/ (4.4.2020).

[48] *Holzki*, Handelsblatt v. 5.8.2019, abrufbar unter https://www.handelsblatt.com/technik/forschung-innovation/cyberkriminalitaet-mit-kuenstlicher-intelligenz-will-die-justiz-kinderpornografie-bekaempfen/24871714.html?ticket=ST-16416080-CiQMGjsjRIQQQBU6NSgU-ap4 (4.4.2020).

gang in ein rechtsförmiges Hauptverfahren finden. Für die Beantwortung dieser Frage ist vom Unmittelbarkeitsgrundsatz auszugehen. Der in der Gesamtschau der §§ 244 Abs. 2, 250 S. 2 und § 261 StPO normierte Unmittelbarkeitsgrundsatz verlangt, dass das richterliche Urteil auf Tatsachen gründet, die aus dem Inbegriff der Hauptverhandlung erhoben wurden. Dies setzt in seinem formellen Gehalt nicht nur voraus, dass der Richter über die Dauer der Hauptverhandlung physisch anwesend ist, sondern auch, dass der Richter die Beweise inhaltlich unmittelbar selbst wahrnehmen kann.[49]

Dies ist jedoch bei digitalen und elektronischen Beweismitteln problematisch, weil nicht das Beweismittel selbst – also etwa die Urkunde –, sondern dessen vermeintlicher Inhalt in Form einer elektronischen Kopie zum Beweisgegenstand wird. Seit der Neufassung des § 249 Abs. 1 S. 2 StPO sind zwar auch elektronische Dokumente „Urkunden", soweit sie verlesbar sind, sodass man meinen könnte, man wäre für die digitale Zukunft gut gerüstet. Das Problem bleibt indes für andere digitale Beweismittel, alles nicht Verlesbare, bestehen. Wie diese Beweismittel in das Hauptverfahren eingeführt werden können, ist daher noch unklar.

b) Richterliche Entscheidung

Aus einer ex-post-Perspektive generieren sich richterliche Entscheidungen in einem arithmetisch binären[50] und damit letztlich zwingenden Ergebnis. Denn die konkreten Tatbestandsvoraussetzungen liegen entweder vor oder nicht. Nicht von ungefähr sehen manche die Juristerei als ein großes Ordnungssystem, bei dem man „nun wirklich selten denken müsse", welches man „leicht durchrattern" könne und das schon jetzt durch Computersysteme zu vernünftigen Antworten auf juristische Fragen komme.[51]

Wer aber mit Montesquieu und Weber den Richter als Subsumtionsautomaten abtut, verkennt die Weisheit, dass man sich vor Gericht und auf Hoher See allein in Gottes Hand befindet. Mag das Ergebnis richterlicher Entscheidung ex post betrachtet noch so klar und eindeutig sein – die Tatbestandsvoraussetzungen selbst, im Strafprozess insbesondere das Maß der Schuld, die letztlich über die Sanktionsfolge entscheidet, sind gerade nicht mathematisch abgrenzbar und in einen binären Code überführbar, sondern an vielen Stellen Ausdruck richterlichen Ermessens, eines nicht greifbaren Judizes[52] im Einzelfall. Die richterliche Glaubwürdigkeitsbewertung von Beschuldigten-, Opfer- oder Zeugenaussagen verlangt ein Repertoire menschlicher Fähigkeiten und Einfühlungsvermögen, deren Erwerb durch ein KI-System aus heutiger Sicht nicht vorstellbar ist.[53] Die vollständige Ausgliederung (straf-)gerichtlicher Entscheidungen, mit all ihren politischen, soziologischen, philosophischen und wirtschaftlichen Zusammenhängen, auf KI-basierte Systeme[54], erscheint daher bislang nur im Film denkbar.[55] In jedem Fall aber fehlt dem Algorithmus das Verständnis für den Inhalt der Entscheidung. Auch wenn letztlich die Frage nach dem Ersatz des menschlichen Richters durch Künstliche Intelligenz[56] jedenfalls für kleinere unproblematische Sachverhalte nicht mehr eine rein hypothetische ist, wie etwa das Beispiel Estlands zeigt,[57] steckt diese Entwicklung zumindest in Europa[58] noch weithin in den Kinderschuhen.

Aus heutiger Sicht lassen sich nur sehr enge Bereiche der richterlichen Entscheidung vorstellen, in denen Algorithmen für eine Entscheidungsfindung zur Anwendung kommen können. Zentral ist dabei, dass sich die Informationen binär übersetzen lassen. Soweit ersichtlich, wurde dies lediglich bei der Rechtsprechung des Europäischen Gerichtshofs für Menschenrechte (EGMR) versucht, wo es Forschern gelang, entsprechende Gerichtsentscheidungen vorherzusagen.[59] Es liegt somit nahe, dass sich in diesem Bereich ein mögliches, wenngleich überaus enges Anwendungsfeld für Algorithmen eröffnet. Denkbar wäre etwa die algorithmenbasierte Berechnung der Entschädigungshöhe wegen überlanger Verfahrensdauer nach Art. 6 EMRK, da die relativen Parameter von Verfahrensart, Verfahrensdauer, Anwendung von Untersuchungshaft, Dauer der Freiheitsentziehung u.a. in Bezug auf die Entschädigungshöhe aufgrund der vorhandenen Datensät-

[49] Statt vieler vgl. *Kreicker*, in: Knauer/Kudlich/Schneider (Hrsg.), Münchener Kommentar zur Strafprozessordnung, 2016, Bd. 2, § 250 Rn. 1 ff.

[50] Also durch die Zahlenabfolge „1" oder „0" im Sinne von „ja" oder „nein".

[51] Zitat von David Precht unter Verweis auch auf Max-Weber, siehe *Gless*/*Wohlers*, in: Böse/Schumann/Toepel (Hrsg.), Festschrift für Urs Kindhäuser zum 70. Geburtstag, 2019, S. 147 (150 f.). Anwendungsbereich ist etwa die Geltendmachung von Ansprüchen wegen Flugverspätungen: *Frese*, NJW 2015, 2090 (2092).

[52] *Vogel*, Juristische Methodik, 1998, S. 175.

[53] Vgl. *Schwintowski*, NJOZ 2018, 1601 (1604); *Fateh-Moghadam*, ZStW 131 (2019), 863 (870).

[54] *Buchholtz*, JuS 2017, 955 (956 f.).

[55] *Gless/Wohlers* (Fn. 52), S. 149 f.

[56] Instruktiv *Gless/Wohlers* (Fn. 52), S. 147 ff.; sowie *Enders*, JA 2018, 721 (723 ff.).

[57] In Estland sollen demnächst kleine Rechtsstreitigkeiten mit einer Schadenssumme von maximal 7.000 EUR durch einen KI-basierten Algorithmus entscheiden, wobei das Urteil im Zuge des Rechtsmittelweges vor einem menschlichen Richter beanstandbar wäre, vgl. Der Standard v. 3.4.2019, https://www.derstandard.at/story/2000100613536/justiz-estland-will-richter-durch-kuenstliche-intelligenz-ersetzen (4.4.2020).

[58] So antwortete der Präsident des US Supreme Court John Roberts im Jahr 2017 auf die Frage, wann denn Maschinen mit Künstlicher Intelligenz bei der Suche nach Tatsachen im Gerichtssaal oder sogar bei gerichtlichen Entscheidungen helfen können, dass dieser Tag bereits hier ist und eine erhebliche Herausforderung für die Justiz darstellt: vgl. *Liptak*, The New York Times v. 1.5.2017, abrufbar unter https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html (06.12.2019).

[59] *Aletras/Tsarapatsanis/Preoţiuc-Pietro/Lampos*, PeeJ Computer Science 2016, 2 (8, 11); vgl. dazu *Rauchegger*, in: Burger/Palmsdorfer/Prickartz/Weber/Weiser/Weismann (Hrsg.), Recht und Sprache, 2019, 23 (31 f.).

ze durchaus als Binärcode erfassbar sind und sich damit algorithmisch verarbeiten ließen. Dadurch wären Standardverfahren überaus schnell und effizient zu erledigen und die EGMR-Arbeitsauslastung entsprechend reduzierbar.

4. Gefährlichkeitsprognosen

Ein anderes Anwendungsfeld von Algorithmen konzentriert sich auf datenbasierte Analysen zur Unterstützung einer gerichtlichen Entscheidung über die vorzeitige Freilassung von Strafgefangenen bzw. Freilassung auf Kaution.

In der deutschsprachigen Straf- und Maßregelvollzugspraxis, nämlich unter anderem im Strafvollzug des Kantons Zürich, kommt die online-basierte Applikation FOTRES (Forensisches Operationalisierte Therapie-Risiko-Evaluations-System) zum Einsatz.[60] Das Programm wurde vom forensischen Psychiater *Frank Urbaniok* entwickelt. Auf der Grundlage von über 80 Risikoeigenschaften gibt die Anwendung eine spezifische Fallhypothese, um einerseits eine kriminalprognostische Aussage zu treffen und andererseits auch eine kriminal-therapeutische Strategie zu verfolgen.[61]

Gerade in der US-amerikanischen Strafrechtspflege ist seit mehreren Jahren ein deutlicher Trend zur Nutzung von empirischer Risikobewertung bei der Frage der Freiheitsentziehung beobachtbar.[62] Hochkonjunktur haben dabei insbesondere Instrumente zur algorithmenbasierten Risikobewertung, welche feststellen, ob, bzw. unter welchen Bedingungen ein Angeklagter aus der Haft entlassen werden soll. Bereits 2007 empfahl das National Center for States Courts, auch in der Phase der Strafzumessung durch die Gerichte auf Risikobewertungen zurückzugreifen, die sich auf empirische Daten stützen.[63] Dementsprechend werden Gerichte in einigen staatlichen Vorschriften (etwa Kentucky[64], Ohio[65] oder Washington[66]) ausdrücklich dazu angehalten, Risikobewertungen in die Entscheidungsfindung miteinfließen zu lassen. Inzwischen findet sich dieser Ansatz auch im Model Penal Code.[67] Überaus bekannt sind in den USA zwei KI-basierte Systeme, nämlich PSA und COMPAS.

Die Entwicklung des Programms PSA (Public Safety Assessment) wurde unter Federführung einer gemeinnützigen Organisation[68] mit dem Ziel entwickelt, den Richtern im konkreten Einzelfall eine wissenschaftsbasierte (und damit unparteiische) kriminalistische Prognose für die Wahrscheinlichkeit eines Bewährungsrückfalls an die Hand zu geben.[69] Der Algorithmus gleicht bestimmte Risikofaktoren der zu untersuchenden Person mit einer Datenbank ab, die über 1,5 Mio. Fälle aus 300 US-Gerichtsdistrikten beinhaltet. Die Analyse erstreckt sich auf neun Risikofaktoren, einschließlich Alter, Vorstrafenregister, frühere gerichtliche Vorladungen oder Anzeigen, wobei gleichzeitig weder geografische Herkunft noch Rasse[70] als Risikofaktor einfließt. Nach Durchführung der Risikoanalyse wird das Ergebnis auf einer sechsteiligen Skala mitgeteilt. Eine empirische Evaluation des Einsatzes von PSA hat etwa im Bundesstaat New Jersey gezeigt, dass die Entlassung auf Bewährung wesentlich zugenommen hat, wobei laut Statistik insbesondere jene Personen in den Genuss vorzeitiger Entlassungen kamen, die nicht wohlhabend waren.[71] Studien aus anderen Bundesstaaten zeigen ähnliche Tendenzen, nämlich den Rückgang der Kautionsent-

[60] *Urbaniok*, FOTRES, Diagnostik, Risikobeurteilung und Risikomanagement bei Straftätern, 3. Aufl. 2016.

[61] *Adams/Kury*, Forum Strafvollzug 2010, 81.

[62] Nach dem Monitoring der National Conference of State Legislatures wurden zwischen 2012 und 2017 zahlreiche Gesetzesinitiativen geschaffen, die sich mit der Verwendung von Risikobewertungen im Zuge des Ermittlungsverfahrens auseinandersetzen: für den Zeitraum von 2012 bis 2014 siehe *Widgery*, Trends in Pretrial Release: State legislation, National Conference of State Legislatures, März 2015, S. 1, abrufbar unter https://www.ncsl.org/portals/1/ImageLibrary/WebImages/Criminal%20Justice/NCSL%20pretrialTrends_v05.pdf; für den Zeitraum 2015 bis 2017 siehe den ergänzenden Bericht von April 2018, abrufbar unter http://www.ncsl.org/portals/1/ImageLibrary/WebImages/Criminal%20Justice/pretrialEnactments_2017_v03.pdf (4.4.2020).

[63] *Garret*/*Monahan*, University of Virginia School of Law, Public Law and Legal Theory Research Paper Series 2018-44, Juli 2018, S. 1 (12).

[64] Kentucky Revised Statutes, § 532.007 Abs. 3 lit. a: „[Sentencing Judges shall consider:] … the results of a defendant's risk and needs assessment including in the presentence investigation".

[65] Ohio Revised Code, § 5120.114, Abs. A: „The department of rehabilitation and correction shall select a single validated risk assessment tool for adult offenders." Dieses Tool soll laut Gesetzestext verschiedenen Institutionen, darunter auch den unicipal courts, common pleas courts, county courts in der Phase des Sentencing zur Verfügung stehen.

[66] Washington Revised Code RCW 9.94A.500, insbesondere Abs. 1 UAbs. 3: „The court shall consider the risk assessment report and presentence reports, if any, including any victim impact statement and criminal history".

[67] *American Law Institute*, Model Penal Code: Sentencing, Proposed Final Draft v. 10.4.2017, S. 171 f.: „The [sentencing] commission shall develop actuarial instruments or processes to identify offenders who present an unusually low risk to public safety, but who are subject to a presumptive or mandatory sentence of imprisonment under the laws or guidelines of the state."

[68] Laura and John Arnold Foundation, https://www.arnoldventures.org/work/criminal-justice (4.4.2020).

[69] https://www.psapretrial.org/ (4.4.2020).

[70] Der Begriff Rasse wird hier aus dem englischen Sprachgebrach „race" übernommen, der dort – untechnisch – auf rein äußerliche Erscheinungsmerkmale abstellt und versteht sich nicht als biologische Determination mehrerer menschlichen Rassen, die es nach letzterem Verständnis gerade nicht gibt.

[71] Vgl. den Evaluationsbericht von 2018, abrufbar unter https://www.njcourts.gov/courts/assets/criminal/2018cjrannual.pdf?c=95Y (4.4.2020).

lassungen und den Ausbau von Entlassungen auf Bewährung.[72]

Der wohl berühmteste Algorithmus im Bereich der US-Strafrechtspflege ist COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), der von einem privaten Unternehmen bereits im Jahr 1998 entwickelt wurde und in der Strafpraxis weit verbreitet ist.[73] Zweck des Programms ist eine Risikobewertung, um die richterliche Entscheidung über die Unterbringung, Überwachung und Behandlung von Straftätern zu unterstützen. Die Risikobewertung fußt dabei (in seiner Grundkonfiguration) auf den Antworten zu 137 Fragen zu verschiedenen Themengebieten, etwa Vorstrafen und Vergehen, Gewaltvergangenheit, Umgang mit Kriminellen, Drogenmissbrauch, wirtschaftliche Schwierigkeiten, Schwierigkeiten in der allgemeinen und beruflichen Bildung, kriminelles familiäres Umfeld, sozialer Kontext, Freizeitgestaltung, Wohnsituation, soziales Umfeld, Persönlichkeitsfaktoren. Der Fragebogen wird entweder von der betreffenden Person selbst, oder anhand der Datenbanken von Polizei und Staatsanwaltschaft beantwortet. Der Fragebogen[74] erfasst nicht nur Vorstrafen des Betreffenden, sondern auch weitreichende Informationen über Familienangehörige (deren Vorstrafen bzw. Erkenntnisse über Alkohol- oder Drogenmissbrauch) und Umgang mit suspekten Freunden und Bekannten, Häufigkeit von Wohnort- und Beschäftigungswechseln, Aggressions- und Wuttendenzen, Schulden und enthält schließlich sogar allgemein gehaltene Fragen wie „Hat eine hungrige Person das Recht zu stehlen?“, die der Betreffende mittels Zustimmung oder Ablehnung beantworten muss.[75]

COMPAS verdankt seine Bekanntheit wohl zuvorderst den kritischen Stimmen aus dem Schrifttum zum Fall Loomis gegen Wisconsin (2016).[76] Der Angeklagte beanstandete vor dem Obersten Gerichtshof in Wisconsin die Höhe der gegen ihn ausgesprochenen Strafe, deren Berechnung auf der Anwendung des COMPAS-Algorithmus basierte. Dabei beanstandete Loomis die mangelnde Transparenz zur Funktion des Algorithmus, zumal COMPAS bereits in der Vergangenheit wegen algorithmeninhärenten Diskriminierungen nach Geschlecht oder „Rasse“ in der Kritik stand. Der Oberste Gerichtshof sprach in seinem Urteil zwar eine Warnung zur zukünftigen Verwendung von COMPAS aus, weil die Funktionsmechanismen von COMPAS unter das Berufsgeheimnis fielen, die Bewertungen des Programms nicht individuell, sondern auf Basis von Gruppenzugehörigkeiten bestimmt würden und daher ein Risiko bestehe, dass die Wahrscheinlichkeit der Straftatenbegehung von bestimmten ethnischen Minderheiten zu sehr gewichtet werde. Gleichwohl wurde das Rechtsmittel von Loomis deshalb verworfen, weil die Bewertungen von COMPAS für die Strafbemessung nicht entscheidend waren, da sie einer Überprüfung und Evaluierung durch einen menschlichen Richter zugeführt worden seien.

IV. Ausblick

Nach diesem Überblick über die bestehenden Einsatzfelder von Algorithmen wird klar, dass es sich hierbei nicht um ein Modethema handelt, sondern der Einsatz von KI möglicherweise sogar als Paradigmenwechsel in der Strafrechtspflege wahrzunehmen ist. Für die wissenschaftliche Auseinandersetzung mit KI ist allerdings entscheidend zu hinterfragen, ob die jeweilige Anwendung tatsächlich der KI-Technologie zuzurechnen ist und nicht mit dem allgemeineren Begriff der Digitalisierung verwechselt wird. Vor diesem Hintergrund gilt es, die technischen Entwicklungen mit einer prinzipienorientierten Strafrechtswissenschaft konstruktiv-kritisch zu begleiten. Dazu gibt es bereits verschiedene Impulse auf europäischer Ebene (unter 1.). Diese sind allerdings um einige technische Überlegungen zu erweitern (unter 2.). Zum Schluss wird am Beispiel des Grundsatzes der Waffengleichheit kurz aufgezeigt (unter 3.), dass die Digitalisierung Neuorientierungen in der Dogmatik des Strafverfahrensrechts einleiten wird.

1. Prinzipienorientierte Entwicklungsimpulse auf europäischer Ebene

Auf europäischer Ebene gibt es gewichtige Impulse für Ansätze einer prinzipienorientierten Begleitung der technischen Entwicklungs- und Anwendungsmöglichkeiten von Künstlicher Intelligenz. So hat 2002 das Ministerkomitee des Europarats die Europäische Kommission für die Effizienz der Justiz (CEPEJ) eingesetzt, die die Qualität der Rechtssysteme der Vertragsstaaten überwacht und prüft. Nachdem der Einsatz von KI in der Strafrechtspflege in der jüngeren Vergangenheit zugenommen hat und sich durchaus gewichtige Risiken für die Menschenrechte ergeben,[77] hat die CEPEJ am

[72] Vgl. die Auswahl an wissenschaftlichen Evaluationen zu PSA unter https://www.psapretrial.org/about/research (4.4.2020).

[73] Vgl. *Sommerer*, NK 2017, 147 (148 f.); *Steege*, MMR 2019, 715 (716), mit konkreten Anwendungsbeispielen von COMPAS.

[74] Der Fragebogen findet sich unter https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html (4.4.2020).

[75] Vgl. das Handbuch von COMPAS unter http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core-_031915.pdf (4.4.2020).

[76] Wisconsin S.C., State v. Loomis, 881, N.W.2d 749 (2016) = Harvard Law Review 2017, 1530; vgl. dazu das Schrifttum in Fn. 74; aus dem englischsprachigen Schrifttum etwa *De Miguel Beriain*, Law, Probability and Risk 2018, 45 (46 ff.); *Donohue*, Harvard Journal of Law & Technology 32 (2019), 657 (662 f., 664, 678); *Freeman*, North Carolina Journal of Law & Technology 2016, 75 (79 ff., 83 ff.); *Liu/Lin/Chen*, International Journal of Law and Information Technology 2019, 122 (126 ff., 129 ff.).

[77] Bereits im März 2018 wurde vom Europarat eine Studie unter dem Titel „Algorithms and Human Rights“ veröffentlicht, vgl. DGI (2017) 12, abrufbar unter https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5 (4.4.2020). Dort wurden unter dem Kapitel fair trial und due process bestimmte Bedenken vorwegge-

4. Dezember 2018 die Ethische Charta für den Einsatz von Künstlicher Intelligenz in Justizsystemen veröffentlicht.[78] Auf diese Weise sollen grundsatzgeleitete Rahmenbedingungen aufgezeigt werden, die der Politik, dem Gesetzgeber und der Justiz als Orientierungshilfe für den Einsatz von KI in nationalen Gerichtsverfahren dienen. Dabei geht CEPEJ davon aus, dass KI zwar zur Verbesserung der Effizienz und Qualität der Justizsysteme beitragen kann. Allerdings müssen menschenrechtliche Standards und Grundrechte, insbesondere jene der EMRK und des Europarats-Datenschutzübereinkommens[79], eingehalten werden. Die Charta ist zwar nicht rechtsverbindlich, doch enthält sie wichtige Impulse für eine prinzipiengeleitete Entwicklung für die Gesetzgeber, welche den normativen Regelungsrahmen für den Einsatz von KI-Instrumenten vorgeben.

Vor diesem Hintergrund enthält die Charta, der eine breite Studie zum bisherigen Einsatz von KI in den Rechtssystemen angehängt ist,[80] fünf Grundprinzipien, die für den Einsatz von KI im Justizwesen einzuhalten sind:

- Grundsatz der Achtung der Menschenrechte
- Grundsatz der Nichtdiskriminierung
- Grundsatz der Qualität und Sicherheit
- Grundsatz der Transparenz, Unabhängigkeit und Fairness
- Grundsatz der Gewährleistung menschlicher Intervention

Die meisten der fünf Grundsätze sind den Rechtsordnungen bereits länger bekannt, und doch erhalten sie durch den Einsatz von KI im Justizwesen eine besondere Bedeutung, wie dies in den Begleittexten zur Charta ersichtlich wird. Im Einzelnen soll der Verweis auf die Achtung der Menschenrechte sicherstellen, dass einerseits die Hersteller von Algorithmen ihre Produkte insbesondere im Lichte der Errungenschaften zur EMRK und zum Datenschutzübereinkommen des Europarats[81] trainieren. Umgekehrt wird angeregt, dass die Justizsysteme ihre Entscheidungen im Rahmen der open data policy zur Verfügung stellen, um letztlich Kooperationsmöglichkeiten zwischen technischen Entwicklungen und rechtlichen Einsatzgebieten zu forcieren.[82] Der Nichtdiskriminierungsgrundsatz nimmt die bisherigen Erfahrungen zum Einsatz von KI in den Fokus, in denen sich durch die Dateneingabe und -analyse Benachteiligungen von Gruppen oder Einzelpersonen ergeben haben. Damit betont die Charta die Bedeutung und Wichtigkeit einer Qualitätskontrolle des Datenursprungs. Der dritte Grundsatz von Qualität und Sicherheit weist inhaltliche Überschneidungen mit dem vorherigen Prinzip auf, weil er auf die Qualität der Algorithmen und der Ursprungsdaten abzielt. So muss für eine Prävention des Diskriminierungsrisikos nicht nur eine akkurate Datenauswahl erfolgen, sondern der Entscheidungsprozess auch im Nachhinein überprüfbar sein, um eine Qualitätssicherung zu gewährleisten. Doch darüber hinaus erfordert auch die Sensibilität justizieller Daten, dass entsprechende Sicherheitsmechanismen bei der Datenverwaltung beachtet werden, um Datenintegrität, aber auch den Zugriff auf die Daten im Sinne der rechtlich eingeräumten Möglichkeiten zu gewährleisten.[83] Der vierte Grundsatz zu Transparenz, Unabhängigkeit und Fairness zielt auf überaus wichtige Eckpfeiler des Strafverfahrens ab. Demnach sollen algorithmenunterstützte Verfahren zur Analyse von rechtlichen Daten zugänglich, verständlich und überprüfbar gestaltet werden. Die Charta fokussiert hier den Bereich der technischen Transparenz, nämlich die Übersetzung technischer Mechanismen, in eine allgemein verständliche Sprache, um den maschinellen Entscheidungsfindungsprozess nachvollziehbar zu machen. Damit klammert die Charta das für den Rechtsanwender herausfordernde Spannungsfeld von Transparenz zwischen legitimen Geheimhaltungsinteressen betrieblicher, geschäftlicher oder ermittlungstechnischer Natur und den Interessen auf Zugang zu automatisierten Entscheidungsfindungsprozessen aufseiten des Betroffenen aus. Von beträchtlichem Innovationspotential ist der Vorschlag der Charta, nicht näher konkretisierte unabhängige Behörden oder Experten institutionell zu verorten, welche a priori und wiederkehrend die technischen Anwendungen und Dienstleistungen im Justizbereich überprüfen, um den Grundsätzen Unabhängigkeit und Fairness entsprechend Rechnung zu tragen.[84] Beim letzten Grundsatz gerät das KI-System selbst in den Mittelpunkt, um sicherzustellen, dass die Rechtsanwender nicht durch KI-basierte Systeme bevormundet werden, sondern als informierte Personen die volle Kontrolle über ihre Entscheidungen haben. Die Computersysteme sollen demnach den unabhängigen Entscheidungsprozess des Anwenders derartiger Systeme nicht eingrenzen, sondern bereichern. Dieser Grundsatz enthält eine Vielfalt an Implikationen, die hier nur beispielhaft aufgezählt werden können. Um dieses Prinzip zur Anwendung zu bringen, bedarf es etwa nicht nur Schulungen der Strafrechtspflege zum Umgang mit derartigen Programmen, sondern auch eines wissenschaftlichen Austauschs innerhalb der Rechtswissenschaften bzw. eines interdisziplinären Dialogs zwischen Technik- und Rechtswissenschaften. So gilt es im Lichte der neuen Technologien zu erheben und interdisziplinär zu vermitteln, welchen Wert und welche Reichweite rechtskräftige Entscheidungen für ein neues Verfahren mit ähnlicher Sachlage haben. Nur wenn hier das Basiswissen gründlich erarbeitet und entsprechend vermittelt wird, kann ein an Rechtsstaatlichkeit, Demokratie und Menschenrechten orientierter und prinzipiengetreuer Einsatz neuer Techniksys-

nommen, auf die die Charta des CEPJ vom 4.12.2018 eine Antwort zu geben versucht.

[78] CEPEJ, European ethical Charter of the use of Artificial Intelligence in judicial systems and their environment, 2019, abrufbar unter https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c (4.4.2020); vgl. dazu *Cornelius*, ZIS 2020, 51 (55).

[79] Übereinkommen zum Schutz des Menschen bei der automatischen Verarbeitung personenbezogener Daten (European Treaty Series Nr. 108).

[80] Vgl. CEPEJ (Fn. 79), S. 48 ff.

[81] Vgl. Fn. 80.

[82] CEPEJ (Fn. 79), Appendix I, S. 19 ff.

[83] Vgl. CEPEJ (Fn. 79), Anm. zu Prinzip Nr. 3, S. 10.

[84] CEPEJ (Fn. 79), S. 11.

teme in der Strafrechtspflege gelingen. Ferner muss der Einsatz neuer Technologien bei Strafprozessen insoweit transparent gemacht werden, damit die Verteidigung auf Basis solider Informationen entsprechende Prozesshandlungen und -strategien vornehmen kann.

Die Grundsätze der Charta geben zwar richtungsweisende Impulse, doch große Fragen zur Ausgestaltung des Einsatzes von KI-Systemen sind noch grundlegend zu durchdringen.[85] Gleichwohl mangelt es nicht an der Impulsgeberschaft durch die europäischen Institutionen. So nahm im September 2019 das vom Europarat lancierte Ad-Hoc-Komitee über Künstliche Intelligenz (Ad Hoc Committee on Artificial Intelligence, kurz: CAHAI) seine Arbeit auf, das im Rahmen seines Mandats bis 2021 einen Rechtsrahmen für die Entwicklung, Gestaltung und Anwendung Künstlicher Intelligenz auf der Grundlage der Standards des Europarats zu Menschenrechten, Demokratie und Rechtsstaatlichkeit prüft.[86] Auch auf EU-Ebene nimmt die Rechtsentwicklung an Fahrt auf. Im April 2018 hat die Kommission eine Europäische Strategie der „menschzentrierten KI" veröffentlicht,[87] wonach der Mensch im Mittelpunkt der KI-Entwicklung stehen soll und insofern der Einsatz dieser leistungsfähigen Technik in den größten Herausforderungen für die Welt gefördert werden soll. Im Dezember 2018 informierte die Kommission über ihren „koordinierten Plan für KI", dass sie nicht nur die Bekämpfung von Krankheiten und Klimawandel durch KI fördern möchte, sondern auch die Kriminalitätsbekämpfung.[88] Für die neue EU-Kommission unter der Führung von Ursula von der Leyen stellt der Einsatz von Künstlicher Intelligenz im Strafverfahren einen thematischen Schwerpunkt dar.[89] Erste Leitlinien lassen sich dem im Februar 2020 veröffentlichten Weißbuch der EU-Kommission entnehmen.[90]

2. Technische Überlegungen

Nach dieser Skizze einer prinzipienorientierten Entwicklung gilt das Interesse nun den Algorithmen selbst. Denn die Datenverarbeitungsmechanismen sind nur dann operativ, wenn ihnen ein solider Datenstock zu Grunde liegt. Hier stellen sich wichtige Grundsatzfragen von juristischer Relevanz, die sich um den Datenbegriff drehen. An anderer Stelle wurde dargelegt, dass Daten in der heutigen Zeit einen wichtigen Rohstoff darstellen.[91] Für den Einsatz von KI in der Strafrechtspflege gelten deshalb einige Besonderheiten, weil Algorithmen nur so schlau sind wie die Daten, auf die sie zugreifen.

a) Datenursprung

Zunächst stellt sich die Frage, welche Daten den KI-Systemen zugrunde liegen, d.h. wo der Datenursprung liegt. Das betrifft nicht nur die legale Beweiserhebung, d.h. allfällige unerlaubte Grundrechteeingriffe, um an Daten zu kommen.[92] Das betrifft insbesondere auch den Grundsatz der Verhältnismäßigkeit bei der Datenbeschaffung. Die Rechtspraxis zeigt, dass die Strafverfolgungsbehörden möglichst großflächig Daten „absaugen" dürfen, eine Kontrolle über die Relevanz der Daten erfolgt erst ex post.[93] Die Datenflut stellt damit nicht nur das Datenschutzrecht für Strafverfolgungsbehörden, sondern auch den Verhältnismäßigkeitsgrundsatz vor gewichtige Herausforderungen praktischer, aber auch dogmatischer Natur.[94]

Ein großer, bislang wenig beachteter Problempunkt ist jene Datengrundlage, die nicht spezifisch für das Training von Algorithmen und KI-Systemen entwickelt wurde, aber dennoch Verwendung findet.[95] Können und dürfen Informationen aus öffentlichen behördlichen bzw. gerichtlichen Pressemitteilungen für eine Datenauswertung herangezogen werden, auch wenn diese Informationsquelle möglicherweise durch das Zielpublikum eine weniger rechtstechnische und dafür mehr allgemeinverständliche Sprache nutzt? Werden alle öffentlichen Informationen der Strafverfolgungsbehörden herangezogen, d.h. auch solche mit vorläufig geltenden Informationen, die später möglicherweise widerrufen oder aktu-

[85] Vgl. die Studie der ETH Zürich: *Jobin/Ienca/Vayena*, Nature Machine Intelligence 2019, 389. Die Studie analysiert 84 weltweit veröffentliche Dokumente, die Ethik-Richtlinien mit Prinzipien propagieren, die bei der Verwendung von Künstlicher Intelligenz zur Anwendung kommen sollen. Dabei wurden nach Häufigkeit genannt: Transparenz (74/84), Gerechtigkeit und Fairness (jeweils 68/84), Schadensvermeidung und Verantwortung (jeweils 60/84), Privatsphäre (47/84), Fürsorge (41/84), Freiheit und Autonomie (34/84), Vertrauen (28/84), Nachhaltigkeit (14/84), Würde (13/84) und Solidarität (6/84). Befremdlich erscheint die numerisch geringe Nennung von Begriffen wie Vertrauen, Nachhaltigkeit und (Menschen-)Würde, die für ethische Herausforderungen im digitalen Zeitalter elementar erscheinen.

[86] Ausführlich zum Mandat des CAHAI unter https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016809737a1 (4.4.2020); vgl. jüngst die Empfehlungen des Europarates über die Auswirkungen algorithmischer Systeme auf die Menschenrechte, CMRec(2020)1 v. 8.4.2020.

[87] „Künstliche Intelligenz für Europa", KOM (2018) 237 endg.

[88] „Koordinierter Plan für künstliche Intelligenz", KOM (2018) 795 endg., S. 1.

[89] *Csonka*, eucrim 2/2019, 77 (78).

[90] White Paper On Artificial Intelligence – A European approach to excellence and trust, COM(2020)65 final; weiterführend aus EU-Perspektive etwa *Frischhut*, in: Austrian Council for Research and Technology Development (ed.), Ethische Herausforderungen im Zeitalter des Digitalen Wandels, 2020 (im Erscheinen).

[91] *Staffler*, NZWiSt 2018, 269 (270 m.w.N.).

[92] Vgl. *Gless*, StV 2018, 671 (673 ff.)

[93] Vgl. *Gless*, StV 2018, 671 (676 ff.).

[94] Vgl. *Anders*, ZIS 2020, 70.

[95] Herausforderungen im Umgang mit Daten, die in engen Bereichen des Privat- und Familienlebens durch digitale Assistenten gesammelt und gespeichert werden, beschreibt *Gless*, StV 2018, 671.

alisiert werden? Der Fundus an solchen Öffentlichkeitspapieren ist groß und durch Online-Archivierung relativ einfach zu beschaffen, weist allerdings in rechtstechnischer Terminologie möglicherweise Abstriche auf, die zu Ungenauigkeiten im KI-Training oder gar zu Verzerrungen des Systems führen.

b) Daten-Bias

Eine empfindliche Schwachstelle in der Anwendung von Algorithmen[96] ist das sog. „bias"[97], also das Auftreten von systematischen Verzerrungen, die auf das ungewollte und im Zweifel unentdeckte Antrainieren von Vorurteilen durch die Künstliche Intelligenz zurückzuführen sind. Um ein KI-System zu trainieren, muss man diesem eine Datenmenge zuführen. Auf der Grundlage dieser Datenmenge kann der Algorithmus ein generalisierendes Ergebnis auswerfen. Problematisch wird die generalisierte Aussage dann, wenn als Trainingsdatei ein sehr spezifischer Datensatz verwendet wird: Lässt man einen Algorithmus beispielsweise enge ortsbezogene Kriminalitätsphänomene aus einem festgelegten Zeitraum trainieren, so wird der Algorithmus aufgrund dieser Datengrundlage in seinen Prognosemodellen stets diesen spezifischen Ortsbezug ausweisen und stärker gewichten, auch wenn das Programm selbst für eine breite Anwendung vorgesehen ist.

Die Gefahr eines „bias" ist real, wie das oben beschriebene Beispiel von COMPAS darlegt. Durch die statistische Auswertung von entsprechenden Datensätzen – deren Nachvollziehbarkeit und ex-post-Kontrolle durch die strenge Akzentuierung des gewerblichen Rechtsschutzes nicht möglich ist – tritt ein rassistisches „bias" auf, wonach für dunkelhäutige Straftäter ein höheres Risiko der Straftatenbegehung errechnet wird als für hellhäutige Täter, selbst wenn letztgenannter einen qualitativ schwereren Vorstrafenhintergrund aufweist.[98] Das Problem lässt sich hypothetisch auch auf andere, oben skizzierte Anwendungsfelder übertragen. Wenn raumbezogene Verfahren dazu führen, dass die Polizei vermehrt an einem als Hot Spot identifizierten Ort Präsenz zeigt, werden möglicherweise (überproportional) mehr Straftäter aufgegriffen. Diese Daten fließen dann wieder zurück in die Datengrundlage des Algorithmus, der den Hot Spot dann noch stärker betont – und damit das Ergebnis zu einem Zirkelschluss verzerrt.

c) Technikvertrauen

Insgesamt spitzen sich die technischen Überlegungen auf die Frage zu, wie zuverlässig die Technik tatsächlich ist. Die technischen Fortschritte geben zwar Hoffnung für einen zuverlässigen Einsatz der entwickelten Systeme,[99] doch nach dem heutigen state of the art erscheint es angebracht, eine gesunde Technikskepsis an den Tag zu legen.[100] Um diesen Aufruf zu belegen, kann auf zwei jüngere Meldungen in den Massenmedien verwiesen werden: So hat das Gesichtserkennungsprogramm eines führenden US-Technikkonzerns 28 US-Kongressmitglieder mit festgenommenen Verdächtigen verwechselt.[101] Der KI-basierte Chatbot eines anderen führenden US-Technikkonzerns wurde innerhalb kurzer Zeit durch Nutzer rassistisch trainiert, was wenige Stunden nach seinem Start zur Abschaltung des Chatbots führte.[102] Da beide Konzerne zu den weltweit führenden Technologie-Unternehmen zählen, wird ersichtlich, dass Technikfortschritt keineswegs von Fehlerfreiheit begleitet ist.

Im Bereich der Strafrechtspflege kann der Einsatz von KI dazu führen, dass bereits auf Ermittlungsebene richtungsweisende Entscheidungen für Ermittlungshypothesen forciert werden. Diesbezüglich hat *Singelnstein* treffend darauf hingewiesen, dass Strafverfolgungsorgane algorithmische Auswertungen als das auffassen müssen, was sie tatsächlich sind: Wahrscheinlichkeitsaussagen mit einer gewissen Fehlerquote. Ein unreflektiertes Technikvertrauen birgt die Gefahr, dass Ermittlungen in die falsche Richtung geleitet werden.[103]

Doch gerade auch im Bereich der richterlichen Tätigkeit gilt es, vor einem technischen Perseveranzeffekt zu warnen.[104] Perseveranz beschreibt die Situation, in der der Beobachter (hier: der Richter) durch ursprüngliche Informationen (hier: durch die technischen Programme) nachhaltig geprägt bzw. beeinflusst wird, sodass sich das bereits vorgefertigte Meinungsbild des Entscheiders durch spätere Informationen nur schwer verändern lässt. Im Kontext von KI-Anwendungen ist damit auf die Gefahr hinzuweisen, dass die Unterstützung durch technische, algorithmenbasierte Anwendungen als Auslagerung richterlicher Tätigkeit auf die Maschine verstanden wird. Es braucht die Bewusstseinsbildung darüber, dass KI-Systeme ein Unterstützungsmechanismus für richterliche Entscheidungen sein können – darüber hinaus darf allerdings eine individuelle Betrachtung des Einzelfalls und Einzelschicksals nicht vernachlässigt werden. Den Rich-

[96] Ausführlich bei *Burgstaller/Hermann/Lampesberger* (Fn. 2), S. 25 ff.

[97] Englisch für Befangen- oder Voreingenommenheit.

[98] Grundlegend zu diesem Problem bei der Anwendung von COMPAS: *Angwin/Larson/Mattu/Kirchner*, Pro Publica v. 23.5.2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (4.4.2020).

[99] Zustimmend auch *Buchholtz*, JuS 2017, 955 (959 f.), der zu einem kritischen Umgang mit Legal Tech mahnt.

[100] So auch *Buchholtz*, JuS 2017, 955 (959 f.); *Fateh-Moghadam*, ZStW 131 (2019), 863 (884); vgl. *Rademacher*, AöR 142 (2017), 366 (400), mit seinem Vorschlag, in diesem Zusammenhang Misstrauen als Rechtsmaßstab zu etablieren.

[101] *Breithut*, SPIEGEL v. 27.7.2018, abrufbar unter https://www.spiegel.de/netzwelt/web/amazons-gesichtserkennung-macht-us-politiker-zu-verdaechtigen-a-1220492.html (4.4.2020).

[102] *Perez*, techcrunch v. 24.3.2016, abrufbar unter https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/ (4.4.2020).

[103] Instruktiv *Singelnstein*, StV 2016, 830.

[104] Instruktiv zum Perseveranzeffekt *Schünemann*, StV 2000, 159; zum kulturellen Perseveranzeffekt siehe *Staffler*, ZStW 131 (2019), 173 (215 f.).

ter von der maschinenbasierten Anwendung zu trennen und KI-basierte Unterstützung im richterlichen Begründungsprozess erst nach dessen Arbeit i.S.e. Verifizierung einzubauen, mag möglicherweise für die Qualität richterlichen Entscheidens durchaus seine Berechtigung haben. Allerdings scheint darin nicht nur eine zu starke Bevormundung i.S.e. Misstrauens zur Technikaffinität von Richtern durchzuschlagen, sondern auch der Sinn einer algorithmischen Unterstützung bei der richterlichen Arbeit leerzulaufen, wenn sich deren Anwendungsfeld erst nach der menschlichen Tätigkeit eröffnet. KI-Systeme können durchaus sinnvoll die richterliche Arbeit ergänzen und auch beschleunigen. Die diesbezügliche Qualitätssicherung richterlicher „Handwerkskunst" ist allerdings nicht über die Bevormundung im Rahmen einer ex-post-Verifikation, sondern nur durch eine grundlegende Bewusstseinsbildung der Richterschaft zu erreichen, die bei der IT-Schulung von Richtern notwendigerweise mit einzubauen ist.[105]

3. Neuorientierung bei strafverfahrensrechtlichen Grundsätzen?

Der Einsatz von digitalen Technologien wie Künstliche Intelligenz lässt neue Fragestellungen,[106] aber auch grundlegende Neuorientierungen bei strafverfahrensrechtlichen Grundsätzen erwarten.[107] Letzteres soll anhand des Beispiels der strafprozessualen Waffengleichheit illustriert werden.

Waffengleichheit gehört zum Kern der Verfahrensfairness (Art. 6 Abs. 3 lit. d EMRK) und soll sicherstellen, dass die Prozessparteien ihre Interessen in einer Weise effektiv vertreten können, die unter gleichwertigen Bedingungen erfolgt, sodass zwischen den Parteien ein prozessuales Gleichgewicht herrscht.[108] Jede Prozesspartei soll hinreichend Gelegenheit haben, ihre eigenen Positionen sowohl zu Sach- als auch zu Rechtsfragen zu präsentieren, ohne gegenüber dem Prozessgegner wesentlich benachteiligt sein.[109] Dies impliziert, dass die Prozessparteien hinreichend vergleichbaren Zugang zu den relevanten Dokumenten und Beweismitteln und insofern vergleichbare Informationsrechte haben. Diesbezüglich spielt die Offenlegung von Beweismitteln eine nicht zu unterschätzende Rolle, da ein einseitiger Wissensvorsprung (verbunden mit der Möglichkeit zur besseren Verfahrensgestaltung) dem Waffengleichheitsgebot zuwiderläuft.

Die Waffengleichheit wird im digitalen Zeitalter durchaus mit erheblichen Herausforderungen konfrontiert, da die gegenwärtigen Regelungen zum Strafverfahren weitestgehend für eine analoge Welt gedacht und konzipiert sind. Welche Friktionen auftreten können, lässt sich mit zwei einfachen Beispielen intuitiv erfassen.

In einer Welt, in der Daten zunehmend auch für das Strafrecht relevant werden, erfordert der Grundsatz der Waffengleichheit, dass die Verteidigung auf dieselben Daten Zugriff hat wie die Staatsanwaltschaft, denn nur dann kann erstere die Wertungen der Anklagebehörde nachvollziehen und eine sinnvolle Verteidigungsstrategie entwerfen. Dafür ist nicht nur Akteneinsicht erforderlich, sondern die Aushändigung von Kopien der elektronischen Datei. Nach dem österreichischen Gerichtsgebührengesetz[110] wird aber für jede kopierte elektronische Datei eine Gebühr von 66 Cent pro Datei (!) in Rechnung gestellt.[111] Mit Blick auf die großen Datenmengen, die von jedem Menschen produziert und die großzügig von Staatsanwaltschaften abgeschöpft werden, offenbart sich hier das Groteske einer solchen Regelung. Denn im Lichte der Unschuldsvermutung und des Grundrechts auf Verteidigung im Strafverfahren ist ein solcher Gebührensatz evident unverhältnismäßig und gleichheitswidrig, weil das vollinhaltliche Akteneinsichtsrecht letztlich nur vermögenden Beschuldigten zusteht, die sich die Entrichtung einer solchen Gebühr leisten können.

Waffengleichheit erfordert allerdings auch, dass bei großen Datenmengen die involvierten Parteien zumindest den Zugang zu denselben Datenanalyse-Programmen haben sollten und die Daten diesen Analysetools zuführen dürfen. Herausforderungen für die Waffengleichheit gibt es dabei sowohl auf Seiten der Staatsanwaltschaft, als auch auf Seiten der Strafverteidigung. Während Staatsanwaltschaften in Österreich beispielsweise die Datenmengen mittels Analysetools durchleuchten dürfen, gewähren sie der Verteidigung im

[105] Vgl. ferner *Beck*, ZIS 2020, 41 (46).

[106] Etwa zum Einsatz von künstlich intelligenten Systemen bei der Unterstützung fremdsprachiger Beschuldigter im Strafverfahren, vgl. *Staffler*, ZStrR 2020, 21 (47 f.). Zum Recht auf Akteneinsicht bei digitalen Beweismitteln vgl. EGMR, Urt. v. 25.7.2019 – 1586/15 (Rook v. Germany) m. Anm. *Meyer/Staffler*, forumpoenale 4/2020.

[107] Etwa zum Thema E-Evidence und der internationalen Rechtshilfe, vgl. *Burchard*, ZIS 2018, 190 ff., 249 ff.; *Tsilikis*, in: Meier/Zurkinden/Staffler (Hrsg.), Recht und Innovation, 2020, S. 163 ff.

[108] Für Österreich vgl. statt vieler *Khakzadeh-Leiler*, Die Grundrechte in der Judikatur des Obersten Gerichtshofs, 2011; aus dem deutschen Schrifttum hingegen *Gaede*, in: Knauer/Kudlich/Schneider (Hrsg.), Münchener Kommentar zur Strafprozessordnung, 1. Aufl. 2018, Bd. 3-2, EMRK Art. 6 Rn. 302; *Meyer*, in: Wolter (Hrsg.), Systematischer Kommentar zur Strafprozessordnung und zum Gerichtsverfassungsrecht, Bd. 10, 5. Aufl. 2019, EMRK Art. 6 Rn. 155, jeweils m.w.N.

[109] *Grabenwarter*/*Pabel*, Europäische Menschenrechtskonvention, 6. Aufl. 2016, § 24 Rn. 67.

[110] Auf dieses Beispiel wies *Weratschnig* im Referat zu den 5. Unternehmensstrafrechtlichen Tagen in Zürich (16.11. 2019) hin.

[111] Vgl. Tarifpost 15, Anmerkung 6, letzter Satz, österreichisches Gerichtsgebührengesetz: „Für unbeglaubigte Aktenabschriften oder -ablichtungen und sonstige Kopien sowie Ausdrucke ist eine Gebühr in Höhe von 66 Cent für jede Seite zu entrichten, werden sie von der Partei unter Inanspruchnahme gerichtlicher Infrastruktur zur Herstellung solcher Abschriften, Ablichtungen, Kopien oder Ausdrucke selbst angefertigt, eine Gebühr in Höhe von 34 Cent für jede Seite. Dies gilt für die einer Partei ausgestellte Kopie einer elektronischen Datei – unter der Voraussetzung, dass die Datei nicht auf Betreiben der Partei erstellt wurde – mit der Maßgabe, dass die Datei einer Seite gleichzuhalten ist."

Rahmen der Akteneinsicht nur den Zugang zu den Daten mittels standardisierten PDF-Readern, während die Verwendung der behördlichen Analyseprogramme ebenso unterbunden wird wie etwa die Anwendung von Analysetools, die der Verteidiger selbst mitbringt.[112] Das Akteneinsichtsrecht läuft damit faktisch leer, wenn die Verteidigung zu einer de facto manuellen Auswertung riesiger Datenmengen gezwungen wird. Doch auch die Verteidigung kann erhebliche Vorteile ausschöpfen, wenn sie tatsächlich im Besitz der betreffenden Daten ist und finanzielle bzw. personelle Kapazitäten hat, die Daten entsprechend auszuwerten, während eine vergleichbare Kapazität auf Seiten der staatlichen Behörden oftmals fehlt. Im Sinne der Waffengleichheit wäre es verfehlt, bei solchen Ungleichgewichten die jeweils andere Seite von Datenquellen oder Anwendungen völlig auszuschließen. Insofern wird Waffengleichheit im digitalen Zeitalter möglicherweise die Rolle der Strafverteidigung als Organ der Rechtspflege neu ausrichten, um ihr die effektive Ausübung ihrer Rechte tatsächlich zu gewährleisten, andererseits neue Kooperationen zwischen Staatsanwaltschaft und Verteidigung initiieren, um eine technische Benachteiligung i.S.d. Waffengleichheit gering zu halten. Ein praktischer Ansatzpunkt könnte die wechselseitige Einräumung von Remote-Access-Möglichkeiten sein, um den Parteien zumindest Zugang zu Daten und Softwareanwendungen zu geben, damit beide Seiten des Strafverfahrens ihre Interessen i.S.d. Rechtspflege effektiv ausüben können.[113]

Die beiden Beispiele zeigen, dass nicht nur Datafizierung und Digitalisierung Herausforderungen für die Dogmatik bereithalten.[114] Mit dem technischen Fortschritt auf dem Gebiet der KI scheint es letztlich unvermeidlich, die tradierten Verfahrensgrundsätze aus dem Strafverfahren der analogen Welt einer Weiterentwicklung zu unterziehen, um deren Bestehen in der digitalisierten Welt sicherzustellen. Die Einsatzmöglichkeiten neuer Technologien sind oft (erschreckend) weitreichend. Damit die Implementierung technischer Möglichkeiten in der Strafrechtspflege nicht zur Aushöhlung rechtsstaatlicher Fundamentalgarantien führt, bedarf es einer aufmerksamen Strafrechtswissenschaft.

[112] So *Weratschnig* (Fn. 110) unter Hinweis auf OLG Wien, Urt. v. 22.1.2015 – 19 Bs 160/14i (unveröffentlicht).
[113] So der Vorschlag von *Romerio* im Referat zu den 5. Unternehmensstrafrechtlichen Tagen in Zürich (16.11.2019), der in der Züricher Anwaltspraxis offenbar bereits z.T. praktiziert wird.
[114] Vgl. *Fateh-Moghadam*, ZStW 131 (2019), 863 (874).