

„Junges Publizieren“

Seminararbeit von

Malin Ebersbach

Big Data, Algorithmen und Bewährungsentscheidungen

Inhaltsverzeichnis

I. Einleitung	27
II. Bewährungsentscheidungen in Deutschland	27
1. <i>Rechtliche Rahmenbedingungen</i>	27
2. <i>Kriminalprognose</i>	29
III. Möglicher Einsatzbereich von Big Data und Algorithmen bei Bewährungsentscheidungen in Deutschland	31
IV. Chancen und Risiken des Einsatzes von Big Data und Algorithmen bei Bewährungsentscheidungen in Deutschland	31
1. <i>Leistungsfähigkeit, Effektivität und Kostenreduzierung</i>	31
2. <i>Chancengleichheit und Gleichheit der Prognosemethode</i>	32
3. <i>Begrenzte Aussagekraft statistischer Methoden</i>	33
4. <i>Weitere grundgesetzliche Fragestellungen</i>	35
V. Schlussbetrachtung	37

I. Einleitung

Neben dem Verhängen von Geldstrafen sind zur Bewährung ausgesetzte Freiheitsstrafen die häufigste Sanktion, die im Rahmen des allgemeinen Strafrechts in Deutschland verhängt wird.¹ Obwohl die Möglichkeit des Verhängens einer Bewährungsstrafe bereits seit 1953 im StGB verankert ist² und Freiheitsstrafen mehrheitlich (ca. 70%) zur Bewährung ausgesetzt werden,³ ist die der Aussetzung zugrunde liegende Prognosepraxis weder einheitlich noch wissenschaftlich fundiert.⁴ Nach empirischen Untersuchungen mangelt es insbesondere an der Sorgfalt, bisweilen auch Kompetenz und Objektivität der herangezogenen Sachverständigen.⁵ Einerseits werden die gesetzlichen Vorschriften als dynamische Verweisung auf den aktuellen Stand der Forschung gelesen, andererseits ist die Methode der Kriminalprognose in den verschiedenen Bezugswissenschaften umstritten, wird teilweise sogar als unmöglich angesehen.⁶ Hinzu kommt, dass davon ausgegangen wird, dass Rückfallraten, insbesondere hinsichtlich erheblicher Delikte, regelmäßig zu hoch eingeschätzt werden.⁷

Insoweit überrascht es nicht, dass vermehrt der Einsatz neuerer technischer Entwicklungen, die den Einsatz von Big Data und Algorithmen im Rahmen von Bewährungsentscheidungen ermöglichen, gefordert wird. Während sich in anderen Ländern, insbesondere in den USA, die Verwendung entsprechender Programme bereits durchgesetzt hat, ist dies in Deutschland bisher nicht umfassend denkbar.⁸ Der vorliegende Aufsatz untersucht, inwieweit die bestehenden strafrechtlichen Regelungen den Einsatz von Big Data gestützten Algorithmen ermöglichen und ob ein solcher das Potential birgt aufgezeigte Missstände der Prognosepraxis nachhaltig zu beheben. Zunächst werden die Bewährungsvorschriften des StGB vorgestellt und der mögliche Einsatz von Algorithmen aufgezeigt. Anschließend werden mit dem Einsatz solcher Systeme verbundene Chancen und Risiken diskutiert und abgewogen.

II. Bewährungsentscheidungen in Deutschland

1. Rechtliche Rahmenbedingungen

Eine Strafaussetzung ist grundsätzlich nur im Rahmen einer verhängten Freiheitsstrafe i.S.d. § 38 StGB möglich. Verurteilte sollen durch die Aussetzung der Strafe die Möglichkeit erhalten durch positives Legalverhalten die Vollstreckung abzuwenden. So können Haftkosten reduziert, entsozialisierenden Wirkungen des Freiheitsentzuges entgegengewirkt oder diese vollständig vermieden werden.⁹ Die rechtlichen Voraussetzungen für Bewährungsentscheidungen finden sich in den §§ 56 ff. StGB. Zu unterscheiden ist im Wesentlichen zwischen der Aussetzung der gesamten Strafe gemäß §§ 56-58 StGB und der Strafrestaussatzung zur Bewährung, geregelt in den §§ 57, 57a

¹ Grube, Jura 2010, 759 (759).

² Grube, Jura 2010, 759 (759).

³ Braasch, in: HK-StGB, 4. Auflage (2017), § 56 Rn.1; Kinzig, in: Schönke/Schröder, StGB, 30. Auflage (2019), § 56 Rn. 1; Ostendorf, in: NK-StGB, 5. Auflage (2017), vor § 56 Rn. 5.

⁴ Grube, Jura 2010, 759 (761); Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 19.

⁵ Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (38); Schöch, in: FS-Widmaier, 2008, S. 769 (969).

⁶ Brettel, Tatverleugnung und Strafrestaussatzung: ein Beitrag zur Praxis der Kriminalprognose, 2007, S. 19/23; Kury/Adams, Forum Strafvollzug 2010, 81 (81).

⁷ Angwin/Larson/Mattu/Kirchner, ProPublica 2016 machine Bias – There's software used across the country to predict future criminals. And it's biased against blacks. Abrufbar unter: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>, (zuletzt abgerufen am 17.10.2019); Braasch, in: HK-StGB, § 56 Rn. 8; Kury/Adams, Forum Strafvollzug 2010, 81 (86).

⁸ Cukier/Mayer-Schönberger, Wirtschaftswoche 2013, 94 (95); Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (46); Härtel, LKV 2019, 49 (53).

⁹ Braasch, in: HK-StGB, § 56 Rn. 1; Dünkel, in: NK-StGB, § 57 Rn. 1; Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 1; Ostendorf, in: NK-StGB, § 56 Rn. 4; Groß, in: MüKo-StGB, 3. Auflage (2019), vor § 56 Rn. 1; Grube, Jura 2010, 759 (760).

StGB.¹⁰ Die vollständige Aussetzung der Freiheitsstrafe ist nur bis zu einer Strafhöhe von maximal zwei Jahren möglich, § 56 StGB.¹¹ Daraus folgt, dass das Gericht zunächst die tat- und schuldangemessene Strafe gemäß § 46 Abs. 1 S. 2 StGB bestimmen muss, bevor eine Aussetzung der Strafe in Betracht gezogen werden kann.¹² Dogmatisch ist die Strafaussetzung zur Bewährung daher eine Modifikation der verhängten Freiheitsstrafe.¹³

Je nach Dauer der Gesamtstrafe hat das Gericht unterschiedliche Voraussetzungen zu prüfen. Bei einer Freiheitsstrafe von unter sechs Monaten ist die Aussetzung der Bewährung bei günstiger Kriminalprognose zwingend.¹⁴ Gegenstand der Legalprognose ist das zukünftige strafrechtskonforme Verhalten des/der Angeklagten.¹⁵ Bagatelldelikten sowie sog. „Jedermannsdelikte“ (z.B. Steuerhinterziehung) werden angesichts fehlender Prognostizierbarkeit nicht in die Verhaltensvorhersage einbezogen. Ebenso spielt generell sozialkonformes oder sozialerwünschtes Verhalten keine Rolle, soweit die Grenzen der Strafbarkeit nicht überschritten werden.¹⁶ Insofern ist der teilweise synonym verwendete Begriff der Sozialprognose irreführend.¹⁷ Die Legalprognose ist dann günstig, wenn das straffreie Verhalten des/der Angeklagten wahrscheinlicher ist als die Erwartung weiterer Straftaten.¹⁸ Diesbezüglich gilt der Grundsatz „in dubio pro reo“ nicht. Zweifel des Gerichts gehen bei hinreichend bewiesener Tatsachengrundlage zu Lasten des/r Angeklagten.¹⁹

Bei einer Freiheitsstrafe zwischen sechs und zwölf Monaten ist neben einer günstigen Legalprognose erforderlich, dass die Verteidigung der Rechtsordnung die Vollstreckung der Strafe nicht gebietet, § 56 Abs. 2 StGB. Dies ist nach der Rechtsprechung der Fall, wenn „eine Aussetzung der Vollstreckung im Hinblick auf schwerwiegende Besonderheiten des Einzelfalls für das allgemeine Rechtsempfinden schlechthin unverständlich erscheinen müsste und das Vertrauen der Bevölkerung in die Unverbrüchlichkeit des Rechts den Schutz der Rechtsordnung vor kriminellen Angriffen erschüttern könnte“.²⁰ Die Verteidigung der Rechtsordnung ist demnach insbesondere dann geboten, wenn die begangene Tat einen hohen Schaden verursacht hat, erhebliche kriminelle Intensität aufweist, hartnäckiges rechtsmissachtendes Verhalten des/r Täters*in festgestellt wird, eine besondere Sozialschädlichkeit der Tat besteht oder der/die Täter*in die Tat aus einer besonderen beruflichen Stellung heraus begeht und dabei das ihm/ihr entgegengebrachte Vertrauen grob missbraucht.²¹

Bei einer Strafdauer zwischen einem und zwei Jahren sind neben den Voraussetzungen des § 56 Abs. 2 StGB besondere Umstände nachzuweisen, § 56 Abs. 3 StGB.²² Dabei sind vor allem gewichtige mildernde Gründe, die in der Persönlichkeit des Täters liegen zu berücksichtigen.²³ Eine Verwertung dieser Kriterien im Rahmen der Strafzumessung, schließt eine nochmalige Berücksichtigung in der Strafaussetzungsentscheidung nicht aus.²⁴

¹⁰ Fischer, StGB, 66. Auflage (2019), § 56 Rn. 2; Groß, in: MüKo-StGB, vor § 56 Rn. 2; Ostendorf, in: NK-StGB, § 56 Rn. 1.

¹¹ Fischer, StGB, § 56 Rn. 2a; Groß, in: MüKo-StGB, vor § 56 StGB, Rn. 2; Grube, Jura 2010, 759 (760).

¹² Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 6; Grube, Jura 2010, 759 (760/763).

¹³ Braasch, HK-StGB, § 56 Rn. 1; Dünkel, in: NK-StGB, § 57 Rn. 4; Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 4; a.A.: Ostendorf, in: NK-StGB, vor § 56 Rn. 1.

¹⁴ Groß, in: MüKo-StGB, vor § 56 Rn. 14; Grube, Jura 2010, 759 (760).

¹⁵ Braasch, HK-StGB, § 56 Rn. 5; Grube, Jura 2010, 759 (761).

¹⁶ Groß, in: MüKo-StGB, vor § 56 Rn. 18; Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 16; Ostendorf, in: NK-StGB, § 56 Rn. 5.

¹⁷ Fischer, StGB, § 56 Rn. 3.

¹⁸ BGH, NStZ-RR 2005, 38 (38); BGH, NStZ 1997, 594 (595); Braasch, in: HK-StGB, § 56 Rn. 6; Fischer, StGB, § 56 Rn. 4a; Groß, in: MüKo-StGB, vor § 56 Rn. 16/24; Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 17.

¹⁹ BGH, StV 92, 106 (107); Braasch, in: HK-StGB, § 56 Rn. 10; Fischer, StGB, § 56 Rn. 4a; Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 17; Ostendorf, in: NK-StGB, § 56 Rn. 24; Grube, Jura 2010, 759 (761).

²⁰ BGHSt 24, 40 (46).

²¹ Groß, in: MüKo-StGB, vor § 56 Rn. 37; Grube, Jura 2010, 759 (763).

²² Groß, in: MüKo-StGB, vor § 56 Rn. 43; Grube, Jura 2010, 759 (760).

²³ Grube, Jura 2010, 759 (763).

²⁴ Grube, Jura 2010, 759 (763).

Gemäß §§ 57, 57a StGB kann eine Freiheitsstrafe bei günstiger Legalprognose und Einwilligung des Verurteilten auch nach teilweise vollstreckter Haft zur Bewährung ausgesetzt werden. § 57 Abs. 1 StGB regelt die sog. Zwei-Drittel-Aussetzung.²⁵ Demnach kann der Strafstrest zur Bewährung ausgesetzt werden, wenn zwei Drittel der verhängten Freiheitsstrafe, mindestens jedoch zwei Monate vollstreckt worden sind.²⁶ Im Unterschied zur primären Strafaussetzung gemäß § 56 StGB ist die Legalprognose dann als günstig anzusehen, wenn die Abwägung zwischen den zu erwartenden Wirkungen des bereits erlittenen Vollzuges und den Sicherheitsinteressen der Gemeinschaft eine weitere Vollstreckung nicht erforderlich erscheinen lässt. Maßgebliches Abwägungskriterium ist daher die Bedeutung der bei einem etwaigen Rückfall gefährdeten Rechtsgüter.²⁷

In Ausnahmefällen kommt auch eine sog. Halbstrafen-Aussetzung gemäß § 57 Abs. 2 StGB in Betracht. Demnach kann die Haftstrafe bereits nach hälftiger Verbüßung, mindestens jedoch nach sechs Monaten, zur Bewährung ausgesetzt werden, wenn es sich um eine Erstverbüßung handelt oder besondere Umstände eine Aussetzung ermöglichen.²⁸ Besondere Umstände können sich auch hier aus einer Gesamtwürdigung von Tat, Persönlichkeit des/der Verurteilten und seiner/ihrer Entwicklung während des Strafvollzuges ergeben, müssen allerdings, anders als i.R.d § 56 Abs. 2 StGB, überdurchschnittlich gewichtig sein.²⁹

§ 57a StGB trifft eine Sonderregelung für die Aussetzung der Vollstreckung einer lebenslangen Freiheitsstrafe. Maßgebliche Gesichtspunkte sind neben der Kriminalprognose auch die Schuldschwere der Anlasstat und Aspekte der Generalprävention.³⁰ Die Aussetzung ist erstmalig nach 15 Jahren verbüßter Freiheitsstrafe denkbar. Zudem darf die besondere Schwere der Schuld die weitere Vollstreckung nicht gebieten.³¹

2. Kriminalprognose

Allen dargestellten Aussetzungsentscheidungen gemein ist die Erforderlichkeit einer günstigen Kriminalprognose. Das Begehen weiterer erheblicher Straftaten durch den/die Täter*in muss demnach einer Wahrscheinlichkeitsprüfung unterzogen werden.³² Zur Erstellung dieser Prognose sind die Persönlichkeit der/des Verurteilten, ihr/sein Vorleben, die Umstände der Tat, Ihr/sein Verhalten nach der Tat, ihre/seine Lebensverhältnisse, der soziale Empfangsraum, die Wirkungen die von der Aussetzung für ihn/sie zu erwarten sind, sowie sonstige Umstände die Rückschlüsse auf das zukünftige Verhalten der/des Angeklagten zulassen, heranzuziehen. Dabei ist auch zu beachten, inwieweit positive Auswirkungen durch, während der Bewährungszeit zu erfüllende, Auflagen und Weisungen zu erwarten sind.³³ Präzisiert werden diese Anforderungen durch die Rechtsprechung des *BVerfG*. Demnach muss im Rahmen der Begutachtung „deutlich werden, in welchem Zusammenhang Ausgangsdelikt und frühere Delinquenz mit der Persönlichkeit stehen [...] und ob deliktsspezifische Persönlichkeitszüge persistieren oder nicht. Dabei muss die prognostische Relevanz der Vortaten und der Anlasstat in die Gesamtpersönlichkeit

²⁵ Fischer, StGB, § 57 Rn. 12.

²⁶ Dünkel, in: NK-StGB, § 57 Rn. 4.

²⁷ Braasch, in: HK-StGB, § 56 Rn. 8; Fischer, StGB, § 57 Rn. 12; Boetticher/Kröber/Müller-Isberner/Böhm/Müller-Metz/Wolf, NStZ 2006, 537 (538).

²⁸ Dünkel, in: NK-StGB, § 57 Rn. 25, 26, 47; Braasch, in: HK-StGB, § 57 Rn. 16, 17.

²⁹ Braasch, in: HK-StGB, § 57 Rn. 24.

³⁰ Fischer, StGB, § 57 Rn. 14.

³¹ Dünkel, in: NK-StGB, § 57a Rn. 7; Fischer, StGB, § 57a Rn. 4-6.

³² Groß, in: MüKo-StGB, vor § 56 Rn. 3; Grube, Jura 2010, 759 (760).

³³ BGH, StV 1992, 62 (64); 1987, 63 (63); Dünkel, in: NK-StGB, § 57 Rn. 31; Groß, in: MüKo-StGB, vor § 56 Rn. 26/27; Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 23.

des Betroffenen nachvollziehbar abgeleitet werden.³⁴ Welche konkrete Methode das erkennende Gericht jedoch zur Erstellung der Legalprognose anwendet, bleibt grundsätzlich diesem überlassen.³⁵ Zu unterscheiden ist zwischen der intuitiven (subjektive Überzeugung des Gerichts), statistischen (nomothetischen/aktuarischen), und klinischen (idiografischen) Prognosemethode.

Die intuitive Methode stützt sich im Wesentlichen auf die subjektive Überzeugung des Gerichts.³⁶ Obwohl diese nur eine geringe Validität aufweist und diese Unzulänglichkeit bereits seit über 100 Jahren bekannt ist, findet sie in der gerichtlichen Praxis überwiegend Anwendung.³⁷

Die klinische Methode umfasst ein Sachverständigengutachten, welches anhand einer Anamnese, eines umfassenden Aktenstudiums und weiterer Instrumente unterschiedlicher Herkunft, insbesondere psychodiagnostischer Tests, erstellt wird.³⁸ Auf Grundlage der Ermittlung von kriminogenen und protektiven Faktoren wird eine individuelle Kriminalitätstheorie entwickelt, die Aussagen über künftiges Verhalten des/der Probanden*in ermöglichen soll.³⁹ Auch im Rahmen der klinischen Methode obliegt die Beurteilung und Einordnung der ermittelten Daten jedoch letztlich der Erfahrung und der Expertise des/r Gutachters*in, sodass die Objektivität, Reliabilität und Validität der grundsätzlich auf wissenschaftlichen Erkenntnissen beruhenden Prognosemethode geschwächt wird.⁴⁰

Bei der statistischen Prognosemethode wird der/die Täter*in anhand von bestimmten Merkmalen einer Vergleichsgruppe zugeordnet. Gedanklicher Ausgangspunkt ist die Erkenntnis, dass bestimmte Merkmale in bestimmten Gruppen statistisch mit Straffälligkeit korrelieren. Kann der/die Proband*in einer bestimmten Gruppe zugeordnet werden und weist besonders viele oder wenige kriminogene/protektive Merkmale auf, die nach der empirischen Datengrundlage mit der Begehung von Straftaten in Zusammenhang stehen, kann eine Wahrscheinlichkeitsaussage über das zukünftige Legalverhalten getroffen werden.⁴¹ Bei einigen Erhebungsinstrumenten werden Punktwerte ausschließlich nach dem Bestehen oder Nichtbestehen bestimmter Merkmale vergeben, bei anderen ist eine zusätzliche Gewichtung möglich, sodass die individuelle Merkmalsausprägung erfasst wird.⁴²

Neuere Instrumente kombinieren statistische und klinische Prognosemethoden.⁴³ Ob dies einen Qualitätszuwachs verspricht, ist bislang nicht eindeutig geklärt.⁴⁴ Während in Deutschland statistische Prognosemethoden jedenfalls nicht ausschließlich zur Erstellung von Kriminalprognosen verwendet werden, gehört die Nutzung dieser Methode

³⁴ BVerfG, NJW 2004, 739 (743).

³⁵ Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 17; Grube, Jura 2010, 759 (761).

³⁶ Budde, Bewährungshilfe 2014, 161 (162); Grube, Jura 2010, 759 (761); Rettenberger, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 28 (29).

³⁷ Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 19; Brettel, S. 34; Urbaniok, Validität von Risikokalkulationen bei Straftätern – Kritik an einer methodischen Grundannahme und zukünftige Perspektiven, 2004, S. 261; Budde, Bewährungshilfe 2014, 161 (162); Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39); Grube, Jura 2010, 759 (761); Rettenberger, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 28 (30).

³⁸ Brettel, S. 34.

³⁹ Dünkel, in: NK-StGB, § 57 Rn. 86; Kinzig, in: Schönke/Schröder, § 56 Rn. 20; Brettel, S. 34; Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (38); Schneider, StV 2006, 99 (100).

⁴⁰ Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39).

⁴¹ Brettel, S. 30; Budde, Bewährungshilfe 2014, 161 (162); Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39).

⁴² Brettel, S. 31; Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (38); Schneider, StV 2006, 99 (102); Rosegger/Laubacher/Moskvitin/Villmar/Palermo/Endrass, International Journal of Offender Therapy and Comparativ Criminology, 716 (716).

⁴³ Braasch, in: HK-StGB, § 56 Rn. 11; Döbele, Standardisierte Prognoseinstrumente zur Vorhersage des Rückfallrisikos von Straftätern – Eine kritische Betrachtung des Einsatzes in der Strafrechtspflege aus juristischer Sicht, 2014, S. 18; Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (37); Schneider, StV 2006, 99 (100).

⁴⁴ Dafür: Dünkel, in: NK-StGB, § 57 Rn. 107; Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 22; Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (37); Dagegen: Boetticher/Dittmann/Nedopil/Nowara/Wolf, NSTZ 2009, 478 (580); Rettenberger, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 28 (34); Bock, StV 2007, 269 (273).

im anglo-amerikanischen Raum zu den führenden Methoden.⁴⁵

Die Rechtsprechung stellt auf Grundlage der dargelegten gesetzlichen Prognoseindizien hohe Anforderungen an die Individualisierung der Verhaltensvorhersage, sodass letztlich eine rein statistische Prognosestellung, die ausschließlich Aussagen über eine gruppenbezogene Wahrscheinlichkeit treffen kann, in Deutschland nicht möglich ist und allenfalls als Ergänzung der idiografischen Methode eingesetzt werden kann.⁴⁶ Im Rahmen einer interdisziplinären Arbeitsgruppe wurden, vor dem Hintergrund der defizitären Praxis, Mindestanforderungen für Prognosegutachten entwickelt, die diesen Befund bestätigen: „[Es] verbietet sich eine abstrakte, allein auf statistische Wahrscheinlichkeiten gestützte Prognose.“⁴⁷

III. Möglicher Einsatzbereich von Big Data und Algorithmen bei Bewährungsentscheidungen in Deutschland

Letztlich ist lediglich die statistische Prognosemethode der strukturierten Erfassung durch Algorithmen, welche die Auswertung von großen Datenbeständen aus sämtlichen Lebensbereichen (Big Data)⁴⁸ ermöglichen, zugänglich.⁴⁹ Denkbar wäre daher gegebenenfalls der unterstützende Einsatz von Tools, wie sie in den USA oder in der Schweiz bereits verwendet werden. Das amerikanische System COMPAS berechnet auf der Grundlage von insgesamt 137 Kriterien einen individuellen Risk-Score auf einer Skala von eins bis zehn.⁵⁰ Die erforderlichen Daten werden zum Teil über die Abfrage automatisierter Datenbanken, zum Teil über Erhebungen durch Justizangestellte gewonnen⁵¹ und basieren auf Fragebogeninterviews mit den Delinquenten und der Erfassung verschiedener Umweltfaktoren, wie der Historie und der persönlichen Lage des/der Angeklagten, einschließlich soziologischer, psychologischer und biographischer Faktoren.⁵² In der Schweiz kommt das Programm FOTRES zur Anwendung, welches mehrere hundert Kriterien erfasst und Informationen zur Rückfallwahrscheinlichkeit, Behandelbarkeit des Probanden und der Bewertung von Therapieerfolgen generiert.⁵³ Im Folgenden soll kritisch diskutiert werden, welche Chancen und Risiken der Einsatz derartiger Prognoseinstrumente in Deutschland birgt.

IV. Chancen und Risiken des Einsatzes von Big Data und Algorithmen bei Bewährungsentscheidungen in Deutschland

1. Leistungsfähigkeit, Effektivität und Kostenreduzierung

Im gerichtlichen Alltag ist eine wissenschaftlich basierte Prognose, aufgrund der immensen Kosten sowie des personellen Aufwands, nicht in jedem Fall möglich.⁵⁴ Insbesondere im Rahmen von Aussetzungsentscheidungen gemäß § 56 StGB sind in der Regel über die intuitive Prognose hinaus keine weiteren Möglichkeiten gegeben; der

⁴⁵ Kinzig, in: Schönke/Schröder, StGB§ 56 Rn. 20; Urbaniok, S. 261.

⁴⁶ BVerfGE 109, 242 (251); Kinzig, in: Schönke/Schröder, StGB, § 56 Rn. 22; Brettel, S. 32; Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (37, 41-48); Rettenberger, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 28 (34).

⁴⁷ Boetticher/Kröber/Müller-Isberner/Böhm/Müller-Metz/Wolf, NSTZ 2006, 537 (539).

⁴⁸ Bilski/Schmid, NJOZ 2019, 657 (657).

⁴⁹ Brettel, S. 31; Ernst, JZ 2017, 1026 (1028); Rettenberger, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 28 (31).

⁵⁰ Shadowen, Ethics and Bias in Machine Learning: A Technical Study of What Makes Us „Good“, 2017, S. 7, 12; Ziegler, c't 2017, 68 (68).

⁵¹ Ziegler, c't 2017, 68 (68).

⁵² Müller/Pöschhacker, in: Algorithmic Risk Assessment als Medium des Rechts – Medientechnische Entwicklungen und institutionelle Verschiebungen aus Sicht einer Techniksoziologie des Rechts, 2019, S. 157 (163).

⁵³ Adams/Kury, Forum Strafvollzug 2010, 81 (81).

⁵⁴ Budde, Bewährungshilfe 2014, 161 (162); Ostendorf, in: NK-StGB, § 56 Rn. 20.

erforderlichen intensiven Datenerfassung steht schon der Grundsatz der Verhältnismäßigkeit entgegen, sodass auch eine algorithmenbasierte Entscheidung verbunden mit der hohen Eingriffsintensität grundsätzlich ausscheidet.⁵⁵

Ohne wissenschaftliche Rückkopplung der Prognosemethode steigt jedoch das Risiko einer Fehlprognose eklatant.⁵⁶ Der Rückgriff auf Big Data Systeme könnte die für eine solche Analyse erforderliche umfassende Datenerhebung und Datenanalyse um ein Vielfaches erleichtern, den Prozess beschleunigen und die Kosten des Gerichts wesentlich reduzieren.⁵⁷ Zudem wäre sichergestellt, dass jede/r Angeklagte auf Grundlage einer wissenschaftlich fundierten Wahrscheinlichkeitsaussage und nicht bloß einer subjektiven Eingabe des/r entscheidenden Richters/in zu einer bedingten/unbedingten Freiheitsstrafe verurteilt würde. Weiterhin steigt die Vorhersagegüte mit der Masse an generierten Daten. Von einem Algorithmus könnte diese Erhebung innerhalb kürzester Zeit übernommen werden.⁵⁸ Teilweise wird durch eine derart strukturierte und umfassende Auswertung auch ein genereller Erkenntnisgewinn über die Ursachen von Kriminalität erhofft, der in Folge dessen auch dem/der einzelnen Probanden*in wiederum zu Gute käme. Individuelle Fördermöglichkeiten und Unterstützungsangebote könnten bedarfsgerecht gestaltet werden und das Rückfallrisiko des/r Einzelnen wesentlich minimieren.⁵⁹ Gleichzeitig wird durch eine algorithmenbasierte Wahrscheinlichkeitsprognose die Möglichkeit eröffnet, die der standardisierten Prognose zugrundeliegenden Kriterien offenzulegen, mithin auch für den/die Verurteilte/n transparent, nachvollziehbar, sowie gegebenenfalls angreifbar zu machen.⁶⁰ Ein Algorithmus ist auch nicht tagesformabhängig oder lässt sich von individuellen Sympathien/Antipathien leiten. So konnte beispielweise gezeigt werden, dass Richter*innen mit fortschreitender Tagesstunde in ihren Aussetzungsentscheidungen zunehmend weniger risikofreundlich agieren.⁶¹ Ein algorithmenbasiertes Entscheidungssystem könnte also auch die Rationalität der getroffenen Entscheidungen wesentlich erhöhen.⁶²

2. Chancengleichheit und Gleichheit der Prognosemethode

Untersuchungen zeigen, dass klinische Prognosen von Psychiater*innen grundsätzlich eine geringe Reliabilität und damit auch eine geringe Validität aufweisen, während statistische Methoden unabhängig von den zur Verfügung stehenden Daten und der Berufserfahrung des Prognostizierenden besser abschneiden.⁶³ Mithin kann einzig die statistische Prognosemethode jedem einzelnen Probanden gleiche Chancen auf Grundlage der gleichen Methode und gleicher Kriterien gewähren und kommt so den Anforderungen des Art. 3 GG am nächsten. Insbeson-

⁵⁵ Albrecht, in: FS-Frisch, 2013, S. 1063 (1064).

⁵⁶ Ostendorf, in: NK-StGB, § 56 Rn. 17/19; Boetticher/Dittmann/Nedopil/Nowara/Wolf, NSTZ 2009, 478 (579); Gless, in: GS-Weßlau, 2016, S. 167 (172).

⁵⁷ Kury/Adams, Forum Strafvollzug 2010, 81 (81); Rettenberger, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 28 (31).

⁵⁸ Brettel, S. 88; Martini, Blackbox Algorithmus – Grundfragen einer Regulierung künstlicher Intelligenz, 2019, S. 27; Zweig, Wo Maschinen irren können – Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung, 2018, S. 24; Bilski/Schmid, NJOZ 2019, 657 (657); Egbert, APuZ 2017, 17 (21); Ernst, JZ 2017, 1026 (1028); Gless, in: FS-Weßlau, S. 167 (167); Krüger/Lischka, in: Mohabbat Kar/Thapa/Prycek, (Un-)berechenbar? – Algorithmen und Automatisierung in Staat und Gesellschaft, 2019, S. 441; Singelstein, NSTZ 2018, 1 (4).

⁵⁹ Belina, Monatsschrift für Kriminologie und Strafrecht 2016, 85 (95); Bilski/Schmid, NJOZ 2019, 657 (657); Martini, DVBl 2014, 1481 (1482); Singelstein, NSTZ 2018, 1 (3).

⁶⁰ Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39); Kury/Adams, Forum Strafvollzug 2010, 81 (81).

⁶¹ Zweig/Krafft, in: Mohabbat Kar/Thapa/Prycek, S. 205.

⁶² Bundesverband Verbraucherzentrale, Algorithmenbasierte Entscheidungsprozesse – Thesenpapier des Bundesverbands Verbraucherzentrale, 2017, S. 7; Zweig, S. 5, 15; Cukier/Mayer-Schönberger, Wirtschaftswoche 2013, 94 (94); Krüger/Lischka, in: Mohabbat Kar/Thapa/Prycek, S. 441; Zweig/Krafft, in: Mohabbat Kar/Thapa/Prycek, S. 205.

⁶³ Rettenberger, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 28 (32).

dere vor dem Hintergrund des Bestehens einer hohen Falsch-Positiv-Rate, könnte eine algorithmenbasierte Entscheidungsfindung zu einer wesentlichen Erhöhung der Strafaussetzungen beitragen.⁶⁴ Ebenso sind fundierte Aussagen über die Irrtumswahrscheinlichkeiten des jeweiligen angewandten Algorithmus möglich.⁶⁵

3. Begrenzte Aussagekraft statistischer Methoden

Andererseits wird vielfach die Aussagekraft nomothetischer Prognosemethoden bezweifelt. So werden von diesen vorrangig Merkmale erfasst, die in der Vergangenheit der/s Proband*in liegen, auf welche er/sie jedoch zum Zeitpunkt der Urteilsfindung keinen Einfluss hat, gegebenenfalls nie hatte,⁶⁶ obwohl die neuere Prognoseforschung gezeigt hat, dass früh in der Biografie liegende Merkmale mit zunehmenden zeitlichem Abstand an Bedeutung verlieren.⁶⁷

Auch ist davon auszugehen, dass verschiedenen Kriterien, je nach Einzelfall, eine unterschiedliche Gewichtung zukommt. Wechselseitige Beziehungen und das Zusammenwirken der unterschiedlichen Faktoren können in einem algorithmenbasierten System nicht hinreichend abgebildet werden.⁶⁸ Auch für den Einzelfall bedeutsame Faktoren, die bisher nicht in die Ermittlung der Verhaltensprognose aufgenommen worden sind, können von dem System nicht ohne weiteres erfasst werden.⁶⁹ Da grundsätzlich jedoch davon ausgegangen werden muss, dass kriminelles Verhalten multifaktoriell bedingt ist und nicht anhand von statischen Kontrollfaktoren zugerechnet werden kann, stellt dies einen erheblichen Mangel in dem Prognosesystem dar.⁷⁰ Eine allgemeingültige Kriminalitätstheorie konnte zudem bislang nicht aufgestellt werden.⁷¹

Des Weiteren ist das Verhalten des/r Probanden*in beispielsweise im Strafvollzug kritisch zu hinterfragen und nicht vorschnell durch den Einsatz eines Algorithmus einer Kategorie zuzuordnen. So sind Anpassungsleistungen von Strafgefangenen während der Dauer des Vollzuges bekannt (Gefängnis als totale Institution), ermöglichen gegebenenfalls aber nur einen geringen Aussagegehalt über das Verhalten des/r Probanden*in in Freiheit.⁷² Sobald ein Manual jedoch eine individuelle Gewichtung der einzelnen Faktoren durch den/die Prognostizierende*n ermöglicht, werden Reliabilität und Validität wiederum beeinträchtigt.⁷³

Weiterhin ergibt sich bei Anwendung rein statistischer Methoden das sogenannte „Mittelfeldproblem“ auf Grund der Gauß'schen Normalverteilung. Der Algorithmus neigt überproportional dazu, Probanden in mittlere Risikokategorien um die 50% einzuordnen, sodass das Gericht letztlich erneut eine intuitive Prognose treffen muss oder auf ein weiteres Sachverständigengutachten angewiesen ist. Nur in Extremfällen kann auf Grundlage der algorithmenbasierten Wahrscheinlichkeitsvorhersage eine eindeutige Entscheidung getroffen werden.⁷⁴

⁶⁴ Rettenberger, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 28 (32).

⁶⁵ Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39).

⁶⁶ Dünkel, in: NK-StGB, § 57 Rn. 25; Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39); Schneider, StV 2006, 99 (103).

⁶⁷ BVerfG, NSTZ 2000, 109 (110); Dünkel, in: NK-StGB, § 57 Rn. 25; Bock, StV 2007, 269 (273).

⁶⁸ Ostendorf, in: NK-StGB, § 56 Rn. 21.

⁶⁹ Braasch, in: HK-StGB, § 56 Rn. 21; Ostendorf, in: NK-StGB, § 56 Rn. 21; Brettel, S. 83; Bock, StV 2007, 269 (272); Ernst, JZ 2017, 1026 (1027).

⁷⁰ Egbert, APuZ 2017, 17 (21).

⁷¹ Brettel, S. 83.

⁷² Dünkel, in: NK-StGB, § 57 Rn. 30; Grube, Jura 2010, 759 (762).

⁷³ Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39).

⁷⁴ Braasch, in: HK-StGB, § 56 Rn. 8; Dünkel, in: NK-StGB, § 57 Rn. 114; Ostendorf, in: NK-StGB, § 56 Rn. 21; Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39).

Ein weiterer Kritikpunkt folgt aus der Funktionsweise statistisch basierter Prognoseinstrumente, die letztlich eine empirisch bestehende Rückfallwahrscheinlichkeit einer bestimmten Gruppe dem einzelnen Probanden zuschreibt.⁷⁵ Die individuell durch das Programm errechnete Wahrscheinlichkeit suggeriert dabei eine für den Einzelfall in der Realität nicht bestehende Aussagesicherheit,⁷⁶ und enthält notwendiger Weise umfassende Generalisierungen, sodass Fehlprognosen unvermeidlich bleiben.⁷⁷ Gesetzlich vorgesehen ist jedoch ausdrücklich eine individuelle Wahrscheinlichkeitsprognose, nicht die Bestimmung eines Kollektivrisikos (s.o.).⁷⁸

Die Aussagekraft von algorithmenbasierten Bewährungsentscheidungen ist ferner nur schwer überprüfbar. Da die Aussetzungsentscheidung bei Straftäter*innen, die als gefährlich eingestuft werden, i.d.R. negativ ausfällt, ist die Ermittlung der Anzahl falsch-positiver Prognosen faktisch kaum möglich.⁷⁹ Neben dem Individuum und der Gesellschaft erhält auch das Programm keine Rückmeldung über etwaige Fehlentscheidungen (Feedbackasymetrie), sodass selbstverstärkende Feedback-Schleifen entstehen.⁸⁰ Dieser Effekt wird durch eine anzunehmende wesentliche Verzerrung der empirischen Daten auf Grund von Hell- und Dunkelfeldeffekten noch verstärkt.⁸¹ Zudem ist die Treffsicherheit eines Prognoseinstrumentes maßgeblich von der Basisrate⁸² abhängig. Ist diese wesentlich niedriger als die Treffsicherheit des Systems wird statistisch gesehen die gesamte Gefangenenpopulation benachteiligt, sofern Individualprognosen erstellt werden. Ein besseres Ergebnis könnte erzielt werden, wenn alle Proband*innen in Freiheit entlassen werden. Im umgekehrten Fall, die Treffsicherheit des Instruments erreicht nicht die Höhe der Basisrate, würde man statistisch ein besseres Ergebnis erzielen, wenn alle Aussetzungsentscheidungen negativ beschieden würden.⁸³

Weiterhin folgt aus der Etikettierung eines Individuums mit einem bestimmten Risiko-Score in der Regel ein erheblicher Stigmatisierungseffekt. Strafgefangene verhalten sich auf Grund des Labels mit hoher Wahrscheinlichkeit entsprechend der Vorhersage des Algorithmus. Ob der Algorithmus jedoch im Einzelfall die richtige Vorhersage getroffen hat oder die Prognose lediglich im Sinne einer sich selbst erfüllenden Prophezeiung als zutreffend angesehen werden muss, bleibt offen.⁸⁴ Selbst wenn ein Rückfall mit 75 prozentiger Wahrscheinlichkeit vorhergesagt werden könnte, bedeutet dies, dass in einem Viertel der Fälle die Bewährungsentscheidung die dem Algorithmus folgt, falsch ausfiele.⁸⁵

Nicht zuletzt ist darauf hinzuweisen, dass keinesfalls ein allgemeingültiges Prognosemanual für alle Gruppen von Straftäter*innen existiert. So ermöglicht beispielsweise das auch in Deutschland verwendete Level of Service Inventory-Revised, im Rahmen der Prognose von Gewaltstraftäter*innen eine vergleichsweise verlässliche Prog-

⁷⁵ BGH, StV 2010, 484 (485); Dünkel, NK-StGB, § 57 Rn. 123; Belina, Monatsschrift für Kriminologie und Strafrecht 2016, 85 (87); Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (39); Schöch, in: FS-Widmaier, S. 967 (958); Singelstein, NStZ 2018, 1 (8); Wisser, American Criminal Law Review 2019, 1811 (1815).

⁷⁶ Ostendorf, NK-StGB, § 56 Rn. 21; Zweig, S. 25; Ernst, JZ 2017, 1026 (1028); Schneider, StV 2006, 99 (102).

⁷⁷ Bilski/Schmid, NJOZ 2019, 657 (658); Boetticher/Dittmann/Nedopil/Nowara/Wolf, NStZ 2009, 478 (579).

⁷⁸ Schneider, StV 2006, 99 (101).

⁷⁹ Adams/Kury, Forum Strafvollzug 2010, 81 (86); Cukier/Mayer-Schönberger, Wirtschaftswoche 2013, 94 (97); Grube, Jura 2010, 759 (762).

⁸⁰ Zweig, S. 26; Zweig/Krafft, in: Mohabbat Kar/Thapa/Prycek, S. 221.

⁸¹ Egbert, APuZ 2017, 17 (21).

⁸² Anteil der Rückfälligen an einer bestimmten Population.

⁸³ Dünkel, NK-StGB, § 57 Rn. 112.

⁸⁴ Bock, StV 2007, 269 (272); Cukier/Mayer-Schönberger, Wirtschaftswoche 2013, 94 (96); Egbert, APuZ 2017, 17 (21); Fröhlich/Spiecker, Können Algorithmen diskriminieren?, Verfassungsblog 2018, abrufbar unter: <https://verfassungsblog.de/koennen-algorithmen-diskriminieren/> (zuletzt abgerufen am 18.10.2019); Krüger/Lischka, in: Mohabbat Kar/Thapa/Prycek, S. 440 (444).

⁸⁵ Cukier/Mayer-Schönberger, Wirtschaftswoche 2013, 94 (96); Egbert, APuZ 2017, 17 (19); Singelstein, NStZ 2018, 1 (3).

nose. Schon bei der Anwendung auf muslimisch geprägte Gewalttäter*innen, ist eine überzufällige Prognose jedoch nicht mehr möglich.⁸⁶ Ein entsprechender Algorithmus müsste demnach zunächst die Kulturangehörigkeit des/r Probanden*in ermitteln, und erst danach eine Prognose anhand von kulturspezifisch ausgewählten Merkmalen erstellen, und somit je nach Kulturzugehörigkeit etc. unterschiedliche Voraussetzungen an eine positive Legalprognose stellen. Die Erstellung einer zuverlässigen Prognose würde daher zwangsläufig zu einer erheblichen Diskriminierung führen.

4. Weitere grundgesetzliche Fragestellungen

Die Kriminalprognose wird anhand einer Vielzahl von persönlichen Daten des Probanden erstellt.⁸⁷ So erfasst beispielsweise das System COMPAS nicht nur die Vorstrafen des/r Angeklagten, sondern auch Vorstrafen von Familienangehörigen, Erkenntnisse über Alkohol- oder Drogenmissbrauch in der Familie, soziale Bindungen des/r Probanden*in, der Umgang mit sog. „anti-sozialen“ Freunden und Bekannten, Schulden, häufige Wohnort- oder Beschäftigungswechsel, Tendenzen zu Wut oder Aggression sowie eine kognitive Verhaltensprognose.⁸⁸ Insofern wird eine massive Beeinträchtigung des Allgemeinen Persönlichkeitsrechts sowie des Rechts auf informationelle Selbstbestimmung (Art. 2 Abs. 1 i.V.m. Art. 1 Abs. 1 GG) und das Recht auf Schutz von persönlichen Daten und auf Vergessen (Art. 8 GRCh) befürchtet.⁸⁹ Allerdings ist bei der Bestimmung der Eingriffsintensität zu berücksichtigen, dass der/die Angeklagte selbst den Anlass zur Erhebung und Auswertung der Daten geschaffen hat.⁹⁰ Zugleich handelt es sich bei den erfassten Daten jedoch um besonders sensible Informationen, die durch die digitale Verknüpfung im Rahmen eines algorithmischen Systems den Nutzer des Systems in die Lage versetzen spezifische Persönlichkeits- und Verhaltensprofile der Proband*innen zu erstellen.⁹¹ Zumindest erscheint daher die Schaffung einer gesetzlichen Grundlage, die den legitimen Zweck für den Einsatz derartiger Systeme hinreichend konkretisiert, mithin die umfassende und unbedachte Verwendung limitiert, bestimmt, welche Daten in welchem Umfang verarbeitet werden, eine Löschung der Daten nach einem bestimmten Zeitraum zwingend vorsieht und Daten weitgehend anonymisiert, zwingend erforderlich.⁹²

Auch das oben angeführte Argument der besseren Nachvollziehbarkeit und Transparenz ist vor dem Hintergrund neuerer technischer Entwicklungen nicht tragfähig. Die Entscheidungsfindungen komplex arbeitender Algorithmen, insbesondere sofern Deep Learning Systeme auf Basis neuronaler Netze eingesetzt werden, ist für den Menschen, zumindest nach dem derzeitigen wissenschaftlichen Kenntnisstand, nicht nachvollziehbar (Black-Box Phänomen). Das Ergebnis des Algorithmus ist weder überprüfbar, noch für die Proband*innen verständlich.⁹³ Auch wenn es in Deutschland nicht vorstellbar ist, dass der dem Algorithmus zugrunde liegende Code nicht öffentlich zugänglich gemacht wird, wie im Fall von COMPAS,⁹⁴ schafft ein extrem komplexer und sich zudem ständig

⁸⁶ Dahle/Schmidt, Forensik/Psychiatrie/Psychologie/Kriminologie 2014, 104 (109); ähnlich: Ostendorf, in: NK-StGB, § 56 Rn. 21; Bock, StV 2007, 269 (272).

⁸⁷ Dahle/Lehmann, Forensik/Psychiatrie/Psychologie/Kriminologie 2018, 37 (38).

⁸⁸ Ziegler, c't 2017, 68 (68).

⁸⁹ Härtel, LKV 2019, 49 (52, 53); Martini, DVBl 2014, 1481 (1483/1484); Singelstein, NStZ 2018, 1 (6); Werner, NJOZ 2019, 1041 (1042, 1043).

⁹⁰ Härtel, LKV 2019, 49 (55); Singelstein, NStZ 2018, 1 (6).

⁹¹ Härtel, LKV 2019, 49 (55); Martini, DVBl 2014, 1481 (1483).

⁹² Härtel, LKV 2019, 49 (55); Werner, NJOZ 2019, 1041 (1042, 1043).

⁹³ Martini, S. 27; Bilski/Schmid, NJOZ 2019, 657 (659); Ernst, JZ 2017, 1026 (1036); Herberger, NJW 2018, 2825 (2828); Gless, in: GS-Weßlau, S. 167 (171); Müller/Pöchhacker S. 158; Wisser, American Criminal Law Review 2019, 1811 (1812).

⁹⁴ Müller/Pöchhacker, S. 164.

verändernder, da lernender Algorithmus mithin keine Nachvollziehbarkeit. Damit ist die Entscheidung nicht hinreichend rechtlich überprüfbar und mithin nicht mit dem Rechtsstaatsprinzip Art. 20 Abs. 3 GG und dem Recht auf effektiven Rechtsschutz Art. 19 Abs. 4 GG vereinbar.⁹⁵

Vor dem Hintergrund der fehlenden Überprüfbarkeit ergibt sich noch ein weiterer Problemkreis: Diskriminierungen auf Grund von Programmierungsfehlern in dem zugrundeliegenden Code oder einer unzureichenden Auswahl der Trainingsdaten sind nur schwer sichtbar zu machen und werden vielfach jahrelang nicht bemerkt. Insoweit besteht insbesondere die Gefahr der Reproduktion von praktischen Relevanzen, impliziten oder expliziten Vorurteilen, sowie die Verzerrung der Datenanalyse (Algorithmic Bias).⁹⁶ Dieser Effekt wird noch dadurch verstärkt, dass einem Algorithmus keine ethische Orientierung immanent ist, sodass in der empirischen Datenbasis bestehende Korrelationen unreflektiert in den Code aufgenommen werden.⁹⁷ Ein Algorithmus ist nicht in der Lage zwischen Korrelation und Kausalität zu unterscheiden (Cum-hoc-ergo-propter-hoc Fehlschluss).⁹⁸ So steht beispielweise mittlerweile fest, dass COMPAS Angeklagte mit dunkler Haut systematisch diskriminiert, obwohl die Hautfarbe oder Ethnie des Individuums nie in das System eingegeben worden ist.⁹⁹ Derartige diskriminierende Entscheidungen sind nicht zuletzt vor dem Hintergrund des Art. 3 GG selbstverständlich nicht hinnehmbar.¹⁰⁰

Die Reproduktion von Vorurteilen kann weiterhin dazu führen, dass bestimmte Personen kategorisch durch das System benachteiligt werden. Dies ist zwar grundsätzlich auch bei einer vorurteilsbehafteten Person, die eine Entscheidung trifft denkbar, allerdings erlangt ein algorithmenbasiertes System einen weit größeren Einflussbereich als ein/e einzelne/r Richter*in, sodass nicht nur einmalig die Chance des/r Probanden/in auf eine positive Aussetzungsentscheidung beeinträchtigt wird, sondern er/sie gänzlich von der Chance einer positiven Prognose ausgeschlossen wird.¹⁰¹ Die bei der Anwendung des Systems COMPAS zu Tage getretenen diskriminierenden Entscheidungen beruhen zudem auf einem statistischen Dilemma. Einem Algorithmus ist es mathematisch nicht möglich, objektive Fairness einzuhalten, wenn bestimmte Eigenschaften (Geschlecht/Hautfarbe) in den durch das Manual erfassten Gruppen unterschiedlich oft vorkommen.¹⁰²

Probleme ergeben sich auch im Hinblick auf die richterliche Unabhängigkeit Art. 97 GG. Der/Die Richter*in muss die Möglichkeit der Wahrscheinlichkeitsprognose überprüfen und die vorgenommenen Wertungen nachvollziehen können. Dies ist in komplexen und intransparenten Systemen gerade nicht der Fall.¹⁰³ Zumindest müsste der/die einzelne Richter*in über fundiertes Wissen zur Bewertung der statistisch generierten Prognosewerte verfügen. Dies ist bislang nicht die Regel und würde umfassende Fortbildungsmaßnahmen notwendig machen.¹⁰⁴

⁹⁵ Martini, S. 40; Martini, DVBl 2014, 1481 (1489); Singelstein, NStZ 2018, 1 (7); Tinnefeld, ZD 2019, 333 (333); Werner, NJOZ 2019, 1041 (1043).

⁹⁶ Ernst, JZ 2017, 1026 (1029); Gless, in: GS-Weßlau, S. 176 (172); Härtel, LKV 2019, 49 (55); Müller/Pöchhacker, S. 166; Werner, NJOZ 2019, 1041 (1042).

⁹⁷ Zweig, S. 26; Bilski/Schmid, NJOZ 2019, 657 (658); Werner, NJOZ 2019, 1041 (1042).

⁹⁸ Martini, S. 60; Belina, Monatsschrift für Kriminologie und Strafrecht 2016, 85 (96).

⁹⁹ Angwin/Larson/Mattu/Kirchner, ProPublica 2016 machine Bias – There's software used across the country to predict future criminals. And it's biased against blacks; Zweig, S. 31; Härtel, LKV 2019, 49 (55); Ziegler, c't 2017, 68 (69).

¹⁰⁰ Tinnefeld, ZD 2019, 333 (333); Werner, NJOZ 2019, 657 (659).

¹⁰¹ Martini, S. 89; Zweig, S. 28; Zweig/Krafft, in: Mohabbat Kar/Thapa/Prycek, S. 204 (208).

¹⁰² Martini, S. 55, 56; Zweig/Krafft, in: Mohabbat Kar/Thapa/Prycek, S. 214.

¹⁰³ BVerfGE 109, 130 (164); Boetticher/Kröber/Müller-Isberner/Böhm/Müller-Metz/Wolf, NStZ 2006, 537 (539); Boetticher, Forensische Psychiatrie und Psychotherapie 2009, 3 (4).

¹⁰⁴ Boetticher/Dittmann/Nedopil/Nowara/Wolf, NStZ 2009, 478 (579); Gless, in: GS-Weßlau, S. 176 (171); Schneider, StV 2006, 99 (102).

V. Schlussbetrachtung

Bislang finden algorithmenbasierte Entscheidungssysteme bei Bewährungsentscheidungen in der Praxis deutscher Gerichte noch keine Anwendung. Die Vorteile liegen jedoch auf der Hand. Die Anwendung einer wissenschaftlich basierten Prognosemethode, die allen Betroffenen die gleiche Chance auf eine positive Aussetzungsentscheidung ermöglicht, ist erstrebenswert. Zudem ermöglicht das Generieren eines prozentualen Wahrscheinlichkeitswertes die vergleichsweise einfache Subsumtion unter das gesetzlich bestimmte Tatbestandsmerkmal. Fest steht allerdings auch, dass bei einer Verwendung algorithmenbasierter Prognosemanuale eine neue Form der Expertise insbesondere bei Richter*innen gewährleistet sein muss. Nur so können die ermittelten Werte gewinnbringend nachvollzogen und für eine individuelle Prognose, wie vom BVerfG und BGH gefordert, nutzbar gemacht werden.¹⁰⁵ Gleichzeitig ergeben sich vielfältige Probleme und Risiken bei der Nutzung derartiger Prognosetools. Zunächst ist festzuhalten, dass algorithmenbasierte Instrumente nur im Rahmen von Extremfällen eindeutige Ergebnisse produzieren können. Der Einsatz dieser ist daher bereits systemimmanent beschränkt. Auch das Spannungsfeld zwischen dem Eingriff in die datenschutzrechtlichen Grundrechte des Betroffenen und der Erforderlichkeit hinreichend aussagekräftiger Datenmengen zur Erzeugung verwertbarer Prognosen, lässt sich nicht auflösen. Solange jedoch die komplexen Strukturen der Prognoseinstrumente nicht hinreichend transparent gemacht werden können, diese sich jeder Nachprüfung entziehen und gegebenenfalls jahrelang unbemerkt diskriminierende Entscheidungen treffen, kann ein umfassender Einsatz solcher Instrumente vor dem Hintergrund der grundgesetzlichen Wertungen nicht verantwortet werden.¹⁰⁶

Allerdings ermöglichen auch herkömmliche Prognosemethoden, wie erörtert, keine vorzugswürdige Alternative. Deshalb sollte ein alternatives Regelungsmodell zumindest in Betracht gezogen werden. So wird teilweise vorgeschlagen, vor dem Hintergrund hoher Falsch-Positiv-Raten und der Erkenntnis, dass durch den Freiheitsentzug im besten Fall eine Nichtwirkung, in der Regel aber ein kontraproduktiver Effekt erwartet werden muss,¹⁰⁷ gegebenenfalls gänzlich von einer Wahrscheinlichkeitsprognose abzusehen und vorrangig normative Kriterien für Aussetzungsentscheidungen gesetzlich zu verankern.¹⁰⁸

Die Kriminalpolitische Zeitschrift (KriPoZ) darf dieses Werk unter den Bedingungen der Digital Peer Publishing Lizenz (DPPL) elektronisch übermitteln und zum Download bereitstellen. Der Lizenztext ist im Internet abrufbar unter der Adresse <http://nbn-resolving.de/urn:nbn:de:0009-dppl-v3-de0>.

¹⁰⁵ Müller/Pöchhacker, S. 175.

¹⁰⁶ so auch: Urbaniok, S. 263; *Wisser*, American Criminal Law Review 2019, 1811 (1815); *Zweig/Krafft*, in: Mohabbat Kar/Thapa/Prycek, S. 204 (219).

¹⁰⁷ *Jasch*, KJ 2014, 237 (238).

¹⁰⁸ *Dünkel*, in: NK-StGB, § 57 Rn. 28, 112, 114.