

Von Dr. Kim Philip Linoh, M.Mel., Halle (Saale)**

Die Strafzumessung steht regelmäßig am Ende des Strafverfahrens. Durch sie kommt die Rechtsfolge für eine Straftat überhaupt erst in die Welt. Die rechtsstaatlich notwendige Begründung der Strafzumessungsentscheidung ist aber nicht trivial und theoretisch wie praktisch mit Problemen behaftet. Diese Probleme, die sich nicht gänzlich beseitigen lassen, liegen in der Natur des Strafzumessungsvorgangs als hochkomplexem Abwägungsprozess. Sie haben aber auch Auswirkungen für den Einsatz von KI bei der Strafzumessung: Ungenügende Trainingsdaten und mangelnde Nachvollziehbarkeit sowie generelle rechtsstaatliche Erwägungen, die mit der Begründung der Strafzumessung zusammenhängen, verbieten (jedenfalls derzeit) einen Einsatz von KI bei der Strafzumessung.

Sentencing usually marks the endpoint of criminal proceedings. It is through this act that the legal consequences of a criminal offense come into being in the first place. However, the constitutionally necessary justification of the sentencing decision is not trivial and is fraught with problems in both theory and practice. These problems, which cannot be completely eliminated, lie in the nature of the sentencing act as a highly complex balancing process. However, they also have implications for the use of AI in sentencing: insufficient training data and a lack of transparency, as well as general constitutional considerations relating to the justification of sentencing, prohibit the use of AI in sentencing (at least for now).

I. Einführung: Bedeutung der Strafzumessung

Die Strafzumessung stellt den Kulminationspunkt dar, auf den der Strafprozess hinsteuert.¹ In dieser Entscheidung komplettiert sich nach einer verurteilenden Entscheidung über die Schuldfrage der Tenor und damit das Urteil. Gleichzeitig wird aus dem abstrakten Recht eine konkrete Rechtsfolge, die Kontingenz des Verfahrens reduziert sich auf eine konkrete Zahl aus Jahren und Monaten oder Tagessätzen und Euro. Erst durch diesen Akt kommt die Rechtsfolge des Straftatbestandes und damit die Konsequenz der Tat für den Täter in die Welt.

Dabei hat die Strafzumessung Bedeutung für alle Verfahrensbeteiligten. Für die Verurteilten entscheidet sie über das

* Dieser Beitrag basiert auf einem gleichnamigen Vortrag des Verf. auf der Tagung „Der Rechtsstaat und das Straf- und Strafverfahrensrecht“ am 27.9.2024 in Halle (Saale). Die vorliegende Fassung ist gegenüber der Vortragsfassung sprachlich angepasst und leicht erweitert worden.

** Der Verf. ist Wissenschaftlicher Mitarbeiter und Habilitand am Lehrstuhl für Strafrecht, Strafprozessrecht und Medizinrecht der Martin-Luther-Universität Halle-Wittenberg (Prof. Dr. Henning Rosenau).

¹ Bruns/Güntge, Das Recht der Strafzumessung, 3. Aufl. 2019, Kap. 2 Rn. 9, bezeichnen die Strafzumessung zurecht als „den allerwichtigsten Teil des Strafrechts“.

weitere Leben, darüber, welche Eingriffe und Konsequenzen getragen werden müssen und welche Implikationen und Stigmata dies für die Zukunft hat.² Für Richter bedeutet die Zumessung der Strafe den eigentlichen Akt des Richtens, der gleichzeitig auch ein schwieriger ist,³ er bleibt immer eine individuelle Entscheidung über einen konkreten Menschen und seine strafbare Handlung. Für die Verteidigung kann der Strafausspruch Gradmesser für den Erfolg (und gegebenenfalls auch die Qualität) der Verteidigungsstrategie und -tätigkeit sein sowie entscheidend für die Beurteilung der Frage, ob die Entscheidung im Wege des Rechtsmittels angegriffen werden soll oder nicht.⁴ Denn die konkrete Strafe für den Mandanten ist immer ein Faktor, der – trotz aller anderen rechtlichen Gesichtspunkte – die Entscheidung maßgeblich beeinflusst. Letztendlich kommt es darauf an, welche konkrete Strafe und damit welche persönlichen Folgen akzeptiert werden (können) oder nicht. Für die Staatsanwaltschaft geht es beim Strafausspruch vor allem darum, ob dem staatlichen Strafanspruch genüge getan wurde, ob die Strafe aus ihrer Sicht angemessen im Verhältnis zur Tat ist. Schließlich hat die Strafzumessung auch für die Allgemeinheit aller Rechtsunterworfenen Bedeutung:⁵ Die Strafhöhe und deren Bewertung als „gerecht“ formt das Bild der Gesellschaft von einem funktionierenden, gerechten und effektiven Strafrecht und einer wirkungsvollen Strafverfolgung. Die maßgebliche Wahrnehmung der Bevölkerung ist nicht die Verurteilung an sich, sondern die ausgeworfene Strafe. Sie kann Unverständnis und Protest ob (scheinbar) zu milder oder harter Strafen hervorrufen und das Vertrauen in die Geltung der Rechtsordnung fördern oder erschüttern. Kurz gesagt: Dem Strafzumessungsakt kommt eine entscheidende Bedeutung zu, weil sein Ergebnis nicht nur für die Verfahrensbeteiligten, allen voran für den Verurteilten, von Bedeutung ist, sondern die öffentliche und gesellschaftliche Wahrnehmung des Strafrechts von diesen Ergebnissen – jedenfalls teilweise – abhängig ist.

II. Die Begründung der Strafzumessungsentscheidung und ihre rechtsstaatliche Dimension

Besondere Bedeutung neben der eigentlichen quantifizierten Strafe erlangt aber auch die Begründung der Strafzumessungsentscheidung. Diese dient verschiedenen Zwecken: Sie soll einerseits – wie jede Begründung eines Urteils – die Verfahrensbeteiligten von der Richtigkeit der Entscheidung überzeugen oder diese Richtigkeit jedenfalls plausibel machen.⁶ Gleichzeitig ist sie auch Dokumentation der Genese

² Bruns/Güntge (Fn. 1), Kap. 2 Rn. 3.

³ Bruns/Güntge (Fn. 1), Kap. 2 Rn. 5; Kahl, DJZ 1906, 895.

⁴ Bartel, in: Barthe/Gericke (Hrsg.), Karlsruher Kommentar zur Strafprozessordnung, 9. Aufl. 2023, § 267 Rn. 1.

⁵ Vgl. beispielsweise die Untersuchungen zur Wahrnehmung von Strafe bei Hoven, KriPoZ 2018, 276 (278 ff.) und Bögelein, Deutungsmuster von Strafe, 2016, S. 60 ff.

⁶ Bartel (Fn. 4), § 267 Rn. 1.

der Strafzumessung und der Berücksichtigung des Vorbringens und der Persönlichkeit des Angeklagten. Damit wird durch die Begründung auch eine Eigenkontrolle des Gerichts, sowie eine Fremdkontrolle durch die Verfahrensbeteiligten und das Rechtsmittelgericht ermöglicht.⁷ Denn wirksame Kontrolle kann nur dort stattfinden, wo auch eine Begründung gegeben werden muss; eine Kontrolle ohne Kenntnis der Gründe für die Entscheidung ist kaum möglich. Daher ist es auch notwendig, dass die Begründung im Recht – anders als im Alltag – immer auch eine transsubjektive Begründung sein muss, die auf die Rechtfertigung der Entscheidung zielt.⁸ Schließlich soll so auch – und darauf wird später noch zurückzukommen sein – dargelegt werden, dass das Urteil gerade auf rationalen Gründen beruht.

Diese Zwecke sind aber untrennbar mit den rechtsstaatlichen Anforderungen an gerechtes Strafen verbunden. Blickt man auf die Merkmale, die dem Rechtsstaatsprinzip des Art. 20 Abs. 3 GG inhärent sind, werden Schlagworte wie die Rechtsbindung, die Rechtssicherheit, der Rechtsschutz, die Verhältnismäßigkeit und die Rechtsgleichheit bemüht.⁹ In allen diesen Dimensionen wird die Begründung der Strafzumessungsentscheidung rechtsstaatlich relevant. Effektiver Rechtsschutz gegen die Strafzumessung kann nur dort erlangt werden, wo die Entscheidung auch begründet ist, weil ohne die Angabe von Gründen gar nicht erkannt werden kann, ob die Entscheidung fehlerbehaftet ist oder nicht. Die Begründung versetzt also den Verurteilten (und seinen Verteidiger) erst in die Lage, Angriffspunkte gegen die Strafzumessungsentscheidung zu finden.¹⁰ Nur durch eine Begründung kann der Verurteilte das Urteil nachvollziehen und erfahren, aus welchen Gründen seine Strafe in dieser Höhe verhängt wurde. Gleiches gilt auch für das kontrollierende Rechtsmittelgericht: Ohne Begründung lässt sich die Entscheidung nicht überprüfen. Hält man sich nochmals vor Augen, welche Bedeutung die hier in Rede stehende Strafzumessung hat – im Kern geht es um das Ausmaß der Strafe – spricht schon der Gesichtspunkt des effektiven Rechtsschutzes für eine Begründungspflicht hinsichtlich der Strafzumessung.¹¹ Aber auch weitere Dimensionen der Rechtsstaatlichkeit sind betroffen. Ein strafrechtliches Urteil muss sich ebenfalls am Verhältnismäßigkeitsprinzip messen lassen – gerade die ausgeworfene Strafe muss jeweils dem Über- und Untermaßverbot genügen.¹² Zudem sind der Aspekt der Rechtsgleichheit

und der richtigen Anwendung des Rechts (Vorbehalt des Gesetzes) von Bedeutung: Nicht nur die Einhaltung der gesetzlichen Strafrahmen, die Freiheit von diskriminierenden Einflussfaktoren (jedenfalls *idealiter*) und insbesondere die methodisch und sachlich richtige Zumessung der konkreten Strafe im Einzelfall anhand der in § 46 StGB vorgegebenen Kriterien, kann nur durch eine entsprechende Begründung nachvollzogen und überprüft werden. Damit zeigt sich, dass Nexus der rechtsstaatlichen Dimensionen die Nachvollziehbarkeit der richterlichen Entscheidung ist; sie ist gleichsam Grundbedingung eines rechtsstaatlichen Strafens. Nur ein nachvollziehbares Urteil kann überhaupt an rechtsstaatlichen Kriterien gemessen und diesen gerecht werden.¹³

Hinzu kommt ein Weiteres: Der Zwang zur Begründung und die Abfassung der Begründung der Strafzumessung fördert die Einhaltung der materiellen Prinzipien und dient damit nicht nur der Nachvollziehbarkeit. Denn das erkennende Gericht ist durch die Angabe der Gründe für die Entscheidung gezwungen, diese Erwägungen – und insbesondere alle strafzumessungsrelevanten Gesichtspunkte – selbst zu reflektieren, sodass, verbunden mit der (drohenden) Kontrolle im Rechtsmittelverfahren, die Begründung Disziplinierungsfunktion hat und so auch Rechtssicherheit hergestellt werden kann.¹⁴ Der schlichte technische Akt der Begründung der Strafzumessungsentscheidung hat also eine erhebliche rechtsstaatliche Relevanz. Das ist aber auch richtig, vergegenwärtigt man sich noch einmal, was für den Verurteilten auf dem Spiel steht.

Dieses rechtsstaatliche Erfordernis setzt auch das einfache Recht um. § 267 Abs. 3 StPO schreibt die Begründungspflicht für die Strafzumessungsentscheidung vor. Diese ist aber nicht nur eine rein prozessrechtliche Pflicht, sondern folgt nach der Rechtsprechung des BGH auch aus dem materiellen Recht, sodass die Strafzumessung im Regelfall auf die Sachrüge hin überprüft werden kann.¹⁵ Diese materiellrechtliche Begründung folgt daraus, dass § 46 StGB eine schuldangemessene Strafe und eine umfassende Abwägung aller strafzumessungsrelevanten Gesichtspunkte fordert. Die Einhaltung dieser Vorgaben kann aber nur dann überprüft und nachvollzogen werden, wenn die Strafzumessung begründet werden muss.¹⁶ Von der Begründungspflicht kennt

⁷ *Bartel* (Fn. 4), § 267 Rn. 1 m.w.N.

⁸ Vgl. zur Anforderung an transsubjektiv gerechtfertigte Argumente – wenn auch in anderem Kontext – *Linoh*, Der rechtfertigende Notstand im Medizinrecht, 2024, S. 111.

⁹ Vgl. *Grzeszick*, in: Dürig/Herzog/Scholz, Grundgesetz, Kommentar, 97. Lfg., Stand: Januar 2022, Art. 20 Rn. 26 ff., 50 ff.; *Jarass*, in: Jarass/Pieroth, Grundgesetz, Kommentar, 18. Aufl. 2024, Art. 20 Rn. 41 ff.

¹⁰ *Bartel* (Fn. 4), § 267 Rn. 1;

¹¹ *Roxin/Schünemann*, Strafverfahrensrecht, 30. Aufl. 2022, § 50 Rn. 4; *Bartel* (Fn. 4), § 267 Rn. 1.

¹² Vgl. hierzu die Spielraumtheorie der Rechtsprechung, bei der Über- und Untermaßverbot die Ober- und Untergrenzen des konkreten Strafrahmens bestimmen: BGHSt 7, 28 (32);

20, 264 (266 f.); 24, 132 (133 f.); *Bußmann*, in: Matt/Renzikowski (Hrsg.), Strafgesetzbuch, Kommentar, 2. Aufl. 2020, § 46 Rn. 7; *Streng*, in: Kindhäuser/Neumann/Paeffgen/Saliger (Hrsg.), Nomos Kommentar, Strafgesetzbuch, Bd. 2, 6. Aufl. 2023, § 46 Rn. 97 ff. jeweils m.w.N.

¹³ Vgl. die Zweckerwägungen bei *Bartel* (Fn. 4), § 267 Rn. 1 m.w.N.

¹⁴ *Wenske*, in: Knauer/Kudlich/Schneider (Hrsg.), Münchener Kommentar zur Strafprozessordnung, Bd. 2, 2. Aufl. 2024, § 267 Rn. 311; *Bartel* (Fn. 4), § 267 Rn. 1.

¹⁵ *van Gemmeren*, in: Schäfer/Sander/van Gemmeren, Praxis der Strafzumessung, 7. Aufl. 2024, Rn. 1359 f.; *Bartel* (Fn. 4), § 267 Rn. 58; vgl. auch *Wenske* (Fn. 14), § 267 Rn. 313 f.

¹⁶ BGHSt 24, 268; BGH NSZ-RR 2003, 4; *van Gemmeren* (Fn. 15), Rn. 58.

das Gesetz aber auch Ausnahmen: So können die Urteilsgründe gem. § 267 Abs. 4 StPO abgekürzt werden, wenn das Urteil hinsichtlich aller Verfahrensbeteiligten rechtskräftig ist. Hier müssen Strafzumessungsgründe nicht angegeben werden, sondern stehen vielmehr im Ermessen des Gerichts, das „den weiteren Inhalt der Urteilsgründe“ – abseits der erwiesenen Tatsachen und des angewendeten Strafgesetzes, die auch bei abgekürzten Urteilen angegeben werden müssen – unter Berücksichtigung der Umstände des Einzelfalls, bestimmt (§ 267 Abs. 3 StPO).¹⁷ Das Fehlen von Ausführungen zum Rechtsfolgenausspruch ist aber angesichts der erwähnten rechtsstaatlichen Dimensionen der Begründung und der Selbstkontrolle des Gerichts kritisch zu sehen. Das gilt selbst dann, wenn in diesem Fall das Urteil und damit auch die Strafe von allen Beteiligten akzeptiert werden. Hier ließe sich freilich argumentieren, dass bei einem allseits akzeptierten Urteil auch die Strafe keiner weiteren Begründung bedürfe, weil diese für die Beteiligten aufgrund der fehlenden Rechtsmittel irrelevant sei.¹⁸ Nach dem bereits Gesagten, greift das aber zu kurz, besonders eingedenk der (selbst)disziplinierenden Funktion der Begründung.

III. Begründungsprobleme der Strafzumessung

Die geforderte Begründung der Strafzumessungsentscheidung steht aber vor erheblichen theoretischen wie praktischen Problemen. Diese stellen das erkennende und begründende Gericht vor erhebliche Herausforderungen und führen in der Praxis nicht selten zu unzureichend begründeten Urteilen.

1. Begründungsprobleme

Im Zuge der Einführung der gesetzlich gebundenen Strafzumessung gab es immer wieder Bedenken gegen eine Rationalisierung der Strafzumessung und eine rationale Bindung des Richters.¹⁹ Auch wenn diese Argumentation heutzutage nicht mehr von großer Bedeutung ist – die gesetzliche Regelung des § 46 StGB hat sich bewährt – lassen sich aus diesen Ansichten grundlegende Probleme bei der Darstellung der Strafzumessung im Urteil ableiten, die auch aktuell Gültigkeit beanspruchen. Im Folgenden sollen drei Problemkreise dargestellt und näher betrachtet werden.

a) Einzelfallcharakter der Strafzumessung

In der Betonung des einzigartigen Charakters der Strafzumessungsentscheidung liegt der Kern einer Argumentation, die davon ausgeht, dass eine Systematisierung und Verrechtlichung der Strafzumessung unmöglich ist.²⁰ Der Grundgedanke ist, dass jede Entscheidung über einen konkreten Fall eine einzigartige Entscheidung ist, die so nicht wiederholt werden kann, also nur ein einziges Mal fallen kann. Die Mannigfaltigkeit des realen Lebens steht daher einer Rationalisierung entgegen, weil jeder Fall so und genau so nur ein einziges

Mal auftritt und zu beurteilen ist.²¹ Treffend umschreiben lässt sich dies mit einem berühmten japanischen Vier-Zeichen-Idiom: 一期一会 (ichi-go ichi-e).²² Wörtlich „einmalige Begegnung, einzigartiger Moment“, so drückt dies den Kern dessen aus, was die Argumentation ausmacht. Wie bei der Heraklit zugeschriebenen Flusslehre und dem berühmten Ausspruch, man könne nicht zweimal in denselben Fluss steigen,²³ wird die Vergänglichkeit und Nichtreproduzierbarkeit jeder einzelnen Begegnung und Situation betont. Im Strafverfahren betrifft das nicht nur die Fallgestaltung selbst (Täter, Tatzeit, Tatort, Tatausführung, Verletzter etc.), sondern auch die Einzigartigkeit des Ergebnisses des durchgeführten Verfahrens. Die Art und Weise der Verfahrensdurchführung, der konkrete Verfahrensgang sowie die Ergebnisse des Verfahrens hängen eben auch von irrationalen und unwägbar Faktoren ab – beispielsweise der Wortwahl von Zeugen, der Körpersprache des Angeklagten oder der Aufmerksamkeit oder den Lichtverhältnissen bei einer Inaugenscheinnahme –, was dazu führt, dass (jedenfalls ontologisch) jedes Ergebnis einmalig ist. Das mag für die Schuldfrage nicht in dem Maße gelten, wie es für die Frage der Strafzumessung gilt. Während bei dieser auch kleine Nuancen zu großen Unterschieden führen können, toleriert jene eher kleinere Abweichungen.²⁴ Die Strafzumessung erfolgt unter Einbeziehung aller Eindrücke des Gerichts im jeweiligen Verfahren und ist deshalb einzigartig und individuell. Das führt letztlich zu der Forderung von weiten Strafrahmen ohne weitere Einschränkungen oder Bindungen des Richters, weil ein weiter Strafrahmen das Einzige sei, was man dem Richter auf dem Weg zum gerechten Urteil mitgeben könne.²⁵

b) Strafzumessung als „irrational bedingter schöpferischer Gestaltungsakt“

Andere Stimmen betonen, dass der Strafzumessungsakt ein „irrational bedingter schöpferischer Gestaltungsakt“ sei.²⁶ Dieser Vorgang sei ein irrationaler, weil die Zumessung der Strafe vielmehr intuitiv erfolge als rational reflektiert. Entscheidend – und hier zeigen sich Schnittmengen zum eben dargestellten Konzept der Einzigartigkeit des Verfahrens –

²¹ Bruns/Güntge (Fn. 1), Kap. 2 Rn. 12.

²² Das Sprichwort soll auf den Begründer des japanischen Teezeremonie *Sen no Rikyū* (千利休, 1522-1591) zurückgehen, dessen Schüler *Yamanoue Sōji* (山上宗二, 1544-1590) es überlieferte, vgl. *Varley*, in: *Varley/Isao* (Hrsg.), *Tea in Japan: Essays on the history of Chanoyu*, 1989, S. 161 (187).

²³ „Ποταμὸς οὐκ ἔστιν ἐμβῆναι δις τῷ αὐτῷ“ nach *Diels*, *Die Fragmente der Vorsokratiker*, Bd. 1, 3. Aufl. 1912, Kap. 12, B 91 (S. 96).

²⁴ Während kleinere Abweichungen die grundsätzliche Beurteilung der Schuldfrage nicht in Frage stellen, ist der Einfluss kleinster Akte auf die Strafzumessung aufgrund ihrer teilweise unbewussten Beeinflussung naheliegend. *Bruns/Güntge*, (Fn. 1), Kap. 2 Rn. 5, bezeichnen die Schuldfrage als „rational lösbar“.

²⁵ *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 11.

²⁶ *Würtenberger*, *Kriminalpolitik im sozialen Rechtsstaat*, 1970, S. 157, 175 ff.; *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 14.

¹⁷ Vgl. hierzu instruktiv *van Gemmeren* (Fn. 15), Rn. 1350 f.; *Wenske* (Fn. 14), § 267 Rn. 465.

¹⁸ *Wenske* (Fn. 14), § 267 Rn. 465.

¹⁹ Vgl. hierzu *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 11 ff.

²⁰ *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 11.

sei das „persönliche Momentum“ im Strafprozess; dieses sei viel leitender für die Entscheidung über die konkrete Strafe als die Rechtsnorm. Damit manifestiert sich, so wird behauptet, im Strafzumessungsvorgang der Vorrang des Irrationalen, der Tradition und Erfahrung.²⁷

Während empirische Untersuchungen zu Einflussfaktoren wie Tageszeit oder Hunger umstritten sind,²⁸ sind Faktoren wie Orientierung an Amtsvorgängern, Kollegen oder regionalen Üblichkeiten, Gepflogenheiten oder Strafmaßlisten²⁹ nicht von der Hand zu weisen. Die Entscheidung eines Richters ist – auch im Bereich der Strafzumessung – eine Entscheidung, die nicht frei von subjektiven und irrationalen Aspekten ist. Die rechtsstaatlichen Probleme der Irrationalität liegen auf der Hand – allein das in Deutschland traditionell bestehende Nord-Süd-Gefälle bei den Strafhöhen³⁰ gerät in große Friktionen mit dem Grundsatz der Rechtsgleichheit, auch wenn eine individuelle Berufung auf Art. 3 GG angesichts der „Einzelfallentscheidung“ schwerfällt. Da sich aber ein irrationaler Einfluss auf die Strafe nicht ausblenden lässt, solange wir Menschen als Richter einsetzen, muss für diese Problematik nach anderen Lösungen gesucht werden, um eine gleichmäßige Rechtsanwendung sicherzustellen: Vorgesprochen werden beispielsweise Strafzumessungsdatenbanken, Strafzumessungskommissionen oder Sentencing Guidelines.³¹ Diese Vorschläge betreffen aber nur einheitlichere Strafaussprüche, nicht aber die Frage der Begründung der Strafzumessung und können an dieser Stelle nicht weiter verfolgt werden. Akzeptiert man aber einen (zumindest gewissen) irrationalen Einfluss auf die Strafe, der den Richter wohl oft unbewusst beeinflusst, so dient die Begründung des Rechtsfolgenausspruchs im Urteil – so das Argument – letztlich nur der Plausibilisierung des vorher gefundenen Ergebnisses.³² Kurzum: Es wird rückwirkend begründet und rational unterlegt, was der Richter bereits vorher – irrational beeinflusst – festgelegt hat. Die Begründung gerät damit zum Feigenblatt, das die subjektiven und irrationalen Einflüsse verschleiern soll.

²⁷ *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 12.

²⁸ Zum sog. hungry judge effect *Danziger/Levav/Avnaim-Pesso*, PNAS 108 (2017), 6889 ff.; krit. aber spätere Analysen von *Glöckner*, Judgment and Decision Making 11 (2016), 601 ff.; *Kohn*, Künstliche Intelligenz und Strafzumessung, 2021, S. 114 f.

²⁹ So hat beispielsweise der *Verf.* im Rahmen des juristischen Vorbereitungsdienstes regelmäßig mit solchen Tabellen bzw. Listen gearbeitet, die interne Vorgaben für die zu beantragende Strafe bei bestimmten Deliktsmerkmalen vorsahen.

³⁰ *Grundies*, in: Hermann/Horten/Pöge (Hrsg.), Kriminologie, Handbuch für Wissenschaft und Praxis, 2. Aufl. 2024, S. 295 ff.; *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 23; *Kohn* (Fn. 28), S. 98 mit umfassenden Nachweisen.

³¹ Hierzu instruktiv *Kaspar*, in: Hoven/Weigend (Hrsg.), Auf dem Weg zu rationaler und konsistenter Strafzumessung, 2024, S. 217 (220 ff.); *ders.*, KriPoZ 2023, 1 (3 ff.); *Hoven*, in: Hoven/Weigend (a.a.O.), S. 229 (233 ff.).

³² *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 12.

c) Notwendige Intransparenz und Unmöglichkeit der sprachlichen Darstellung

Schließlich erscheint das Problem der Intransparenz und der sprachlichen Darstellung offensichtlich zu sein. Diese beiden Probleme haben große Überschneidungsbereiche und sollen daher gemeinsam besprochen werden. Die These der notwendigen Intransparenz geht davon aus, dass viele Faktoren Einfluss auf die Strafzumessung haben, die auch unbewusst wirken und diese gar nicht transparent gemacht werden können, weil sie der entscheidenden Person eben nicht bewusst sind.³³ Was aber nicht als bewusster Entscheidungsfaktor wahrgenommen wird, kann seinen Weg nicht in die Urteilsgründe finden. Diese sind daher notwendigerweise kein genaues Abbild aller wesentlichen Faktoren für die Strafzumessung, sondern bilden nur einige, nämlich die bewussten Faktoren ab.³⁴

Auf ein ähnliches Problem weist die These von der Unmöglichkeit der sprachlichen Darstellung des Strafzumessungsvorgangs. Mit der Intransparenz hat sie gemeinsam, dass es notwendigerweise zu unvollständigen und intransparenten Entscheidungsgründen führt, wenn eine sprachliche Darstellung der entsprechenden Gründe überhaupt nicht möglich ist. *Raimund Hassemer* spricht davon, dass Richter den komplexen Vorgang der Abwägungsentscheidung gar nicht erschöpfend darstellen können.³⁵ Der Abwägungsvorgang ist ein hochkomplexer Vorgang, der sich der Operationalisierung und sprachlichen Fassung widersetzt. Diejenigen, die die Abwägung vornehmen, können den Vorgang im Regelfall gar nicht näher beschreiben, als die entsprechenden Faktoren auf beiden Seiten der Waagschale zu benennen; den eigentlichen Vorgang der Abwägung zu verbalisieren und niederzulegen scheidet hingegen. Vielmehr ist dieser Vorgang so individuell und komplex, dass er sich der sprachlichen Darstellung entzieht. Dabei soll es bereits daran scheitern, alle abwägungsrelevanten Faktoren zu benennen und den Vorgang der Abwägung selbst zu beschreiben. Letztlich bedeutet dies – akzeptierte man die Unmöglichkeit der sprachlichen Darstellung –, dass die Urteilsgründe die wirkliche Genese der Entscheidung nicht abbilden können und nur dazu dienen, eine Entscheidung anhand der Kriterien des Revisionsgerichts revisionsfest zu machen.³⁶ Auch hier wird nur ein Teil der wahren Strafzumessungsgründe dargestellt, weil die wahre, sprachlich unmögliche Darstellung durch eine konsenterte Darstellung der Strafzumessung unter Nutzung der Benennung von sprachlich fassbaren Teilen ersetzt wird, die aber notwendigerweise unvollständig ist.³⁷

d) Gegenpositionen

Die aufgezeigten Probleme lassen sich aber auch grundsätzlich hinterfragen. Wie *Bruns* und *Güntge* richtig bemerken, steht die Tatsache, dass es sich bei der Strafzumessung um

³³ Vgl. *Kohn* (Fn. 28), S. 110 ff.

³⁴ Hierzu schon die Untersuchung von *R. Hassemer*, MSchKrim 1983, 26 ff.; *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 19.

³⁵ *R. Hassemer*, MSchKrim 1983, 26.

³⁶ *Bruns/Güntge* (Fn. 1), Kap. 2 Rn. 19.

³⁷ Vgl. *Kohn* (Fn. 28), S. 120 ff.

Einzelfallentscheidungen handelt, einer methodischen Rechtsanwendung und auch deren Begründung nicht im Wege.³⁸ Allein die Charakteristik als einmaliges Ereignis macht eine Begründung nicht unmöglich, sondern lediglich entsprechend komplex. Der Gesetzgeber hat in § 46 StGB eine Regelung getroffen und die Begründung in § 367 StPO vorgeschrieben. Die richterliche Entscheidung ist daher keine freie Entscheidung, sondern ein rechtlich gebundenes Ermessen, innerhalb des gesetzlichen Strafrahmens eine angemessene Lösung zu finden.³⁹

Die vorgenommene Mystifizierung der Irrationalität der Strafzumessung begegnet erheblichen rechtsstaatlichen Problemen. Eine gerichtliche Entscheidung in einem (modernen) Rechtsstaat erfordert es, dass sich eine rational und überindividuell nachvollziehbare Begründung für eben jene Entscheidung angeben lässt.⁴⁰ Nur so, nämlich durch eine Obliegenheit zur rationalen Begründung, kann eine Entscheidung überprüfbar sein; eine nicht überprüfbare Entscheidung steht aber immer in der Gefahr, Willkür hervorzubringen.⁴¹ Dieser kann also durch das Erfordernis, transsubjektiv vertretbare Entscheidungsgründe anzugeben, vorgebeugt werden. Wo aber ein Rest irrationaler Einflüsse verbleibt, dort gilt es, diese Irrationalität zu reflektieren und Möglichkeiten zur Vermeidung unbilliger Einflüsse zu entwickeln. Ein noch verbleibender „irrationaler Rest“ bei der Strafzumessung spricht jedenfalls nicht dagegen, soweit, wie es möglich ist, rationale Begründungen anzugeben. Die Anerkennung eines Restes von irrationalen Einflussfaktoren heißt dann aber auch, ein verbleibendes Begründungsdefizit anzuerkennen.

Schließlich spricht auch die Komplexität des Abwägungsvorgangs und die Schwierigkeit sprachlicher Fassbarkeit aller Strafzumessungsgründe nicht für ein generelles Begründungsproblem. Denn § 367 StPO fordert nicht die Angabe aller Faktoren der Strafzumessung, sondern nur der tragenden Gründe für die Entscheidung.⁴² Diese lassen sich aber in der Regel auch verbalisieren und darlegen. Damit geht das Gesetz einen Kompromiss ein und reduziert die Komplexität der Begründung. Gleichzeitig folgt daraus aber auch ein Verbot des Verschweigens von tragenden Zumessungsgründen, beispielsweise weil andere Gründe leichter darzustellen oder sozial oder politisch opportun sind.⁴³ Der komplexe Abwägungsvorgang muss also nicht in allen Facetten, aber die tragenden Punkte der Abwägung, insbesondere das Abwägungsmaterial und dessen Bewertung müssen angegeben werden. Nicht zu verschweigen ist freilich, dass damit trotzdem noch Schwierigkeiten bei der Formulierung

und Darstellung der Begründung verbleiben, weil auch die tragenden Gründe sprachlich dargestellt werden müssen. Dies lässt sich aber durchaus bewerkstelligen, ist dies doch deutlich weniger komplex als eine umfassende Darstellung des gesamten Abwägungsvorgangs.

e) Zwischenergebnis: Beherrschbare Begründungsprobleme

Die Analyse der Probleme der Begründung der Strafzumessungsentscheidung im Urteil hat gezeigt, dass die aufgeworfenen Probleme schon auf einer theoretischen Ebene nicht in dem Maße bestehen, wie es prima facie erschien. Irrationale Anteile der Entscheidung müssen aber ebenso anerkannt und reflektiert werden, wie sprachliche Darstellungsprobleme. Die verbleibenden Problemstellungen gilt es zu analysieren und im Rahmen der Entwicklung einer systematischen und operationalisierbaren Dogmatik der Strafzumessung zu bearbeiten. Die Begründungsprobleme erscheinen auf einer theoretischen Ebene jedenfalls durch entsprechende Forschung und dogmatische Weiterentwicklung beherrschbar.

2. Empirische Befunde

Die Begründungsprobleme der Strafzumessung lassen sich aber auch in der Praxis feststellen. Dass dies nicht nur eine vage These von der Unzulänglichkeit der Begründung ist, lässt sich auch anhand empirischer Befunde – jedenfalls schlaglichtartig – darlegen. Für diesen Beitrag wurden insgesamt 41 amtsgerichtliche und 59 landgerichtliche Urteile, die in der Datenbank „juris“ verfügbar waren, analysiert. Dabei wurde sich notwendigerweise auf diejenigen Urteile beschränkt, die Ausführungen zur Strafzumessung enthielten, womit freisprechende Urteile sowie solche, die allein Maßregeln der Besserung und Sicherung betrafen, ebenso ausgeschlossen wurden, wie solche, die sich allein mit Kostenfragen beschäftigten. Bei der Analyse zeigte sich, dass die Strafzumessungsbegründung im Vergleich zu den sonstigen Ausführungen im Urteil deutlich weniger Raum einnehmen. Die Erörterungen zum Sachverhalt, zur Beweiswürdigung und gegebenenfalls materiell-rechtliche Erwägungen nehmen einen deutlich größeren Umfang ein als die Strafzumessung. Im Rahmen der Strafzumessungsbegründung finden sich in der Regel ausreichende und umfassende Erörterungen zum gesetzlichen Strafrahmen und zu seinen Verschiebungen. Auch die für und gegen den Angeklagten sprechenden Umstände werden oft ausführlich aufgelistet, aber eben auch nur das. Es fehlt hingegen in nahezu allen betrachteten Urteilen an umfassenden Abwägungsbegründungen: Es erfolgt meist weder eine Bewertung der einzelnen Strafzumessungstatsachen und auch keine Valorisierung und Inbezugsetzung zu den anderen Zumessungsfaktoren. Vielmehr finden sich oft nur pauschale und oberflächliche Formulierungen, die kaum Erkenntniswert haben und eine Begründung der Abwägung vermissen lassen. Exemplarisch sollen hier einige dieser Formulierungen genannt werden:

„Nach Abwägung aller für und gegen den Angeklagten sprechenden Umstände und dem jeweiligen Gesamtbild der Taten hat die Kammer hinsichtlich der Taten 1 und 2

³⁸ Bruns/Güntge (Fn. 1), Kap. 2 Rn. 25 ff.

³⁹ Kinzig, in: Schönke/Schröder, Strafgesetzbuch, Kommentar, 30. Aufl. 2019, § 46 Rn. 68; Bußmann (Fn. 12), § 46 Rn. 6 f.; näher Schäfer/Sander/van Gemmeren, in: Schäfer/Sander/van Gemmeren (Fn. 15), Rn. 1505 ff.

⁴⁰ Bruns/Güntge (Fn. 1), Kap. 2 Rn. 18.

⁴¹ Bruns/Güntge (Fn. 1), Kap. 19 Rn. 16 f.

⁴² Bartel (Fn. 4), § 267 Rn. 52; van Gemmeren (Fn. 15), Rn. 1361; vgl. auch die Entscheidungen BGHSt 27, 3; 24, 268.

⁴³ Bruns/Güntge (Fn. 1), Kap. 18 Rn. 2 ff., Kap. 19 Rn. 16 f.

jeweils eine Geldstrafe in Höhe von 30 Tagessätzen für tat- und schuldangemessen erachtet.“⁴⁴

„Die Kammer hat im Rahmen der konkreten Strafzumessung erneut jeweils alle tat- und täterbezogenen Gesichtspunkte (siehe hierzu Ziffer F. II.) umfassend abgewogen und hält nach einer Gesamtbetrachtung all dieser tat- und täterbezogenen Gesichtspunkte eine Freiheitsstrafe von 7 Jahren für tat- und schuldangemessen.“⁴⁵

„Ausgehend von dem gesetzlichen Strafrahmen waren nach Abwägung aller für und gegen den Angeklagten sprechenden Umstände, insbesondere des mildernden Geständnisses, der Schadenswiedergutmachung und fehlender Voreintragungen eine Geldstrafe von 40 Tagessätzen zu je 60,00 € tat- und schuldangemessen.“⁴⁶

„Tat- und schuldangemessen erachtet das Gericht bezüglich dem Angeklagten T eine Geldstrafe von 40 Tagessätzen zu je 25 Euro. Hinsichtlich dem Angeklagten A ist eine Geldstrafe von 40 Tagessätzen zu je 80 Euro tat- und schuldangemessen. Aufgrund vorgenannter Umstände hat das Gericht gegen den Angeklagten B eine Geldstrafe von 50 Tagessätzen zu je 20 Euro festgesetzt.“⁴⁷

Dieser Ausschnitt an Befunden deckt sich nicht nur mit anekdotischer Erfahrung, sondern auch mit umfassenden empirischen Studien: So zeigt die Untersuchung von *Obert*, in der Urteile zum Wohnungseinbruchsdiebstahl analysiert wurden, dass nur in 19 % der Fälle eine echte Abwägung in Urteilen zu finden ist.⁴⁸ Sie bescheinigt auch eine quantitative Knappheit und inhaltliche Oberflächlichkeit der meisten Entscheidungen:

„Darüber hinaus legt die Urteilsanalyse teilweise gravierende Defizite in den Strafzumessungsbegründungen der ausgewerteten Urteile offen. Die Begründungen sind teilweise nicht einmal im Ansatz geeignet, die Strafzumessungsentscheidungen transparent zu machen. Regelmäßig sind die Strafzumessungserwägungen quantitativ knapp und inhaltlich oberflächlich ausgestaltet. Überwiegend werden die strafscharfenden und strafmildernden Strafzumessungserwägungen lediglich blockartig gegenübergestellt; eine echte Abwägung der einzelnen Erwägungen findet dabei meist nicht statt. Regelmäßig wird weder die Bedeutung und das Gewicht der einzelnen Erwägungen im konkreten Fall noch die Einordnung der einzelnen Erwägungen als schuld- oder präventionsrelevant transparent gemacht. Zudem muss konstatiert werden, dass die in den Urteilen explizit angeführten Erwägungen keine ver-

lässlichen Rückschlüsse darüber erlauben, welche Erwägungen tatsächlich entscheidungsrelevant waren.“⁴⁹

3. Strafzumessungsbegründung: Theorie-Praxis-Lücke

Empirische Befunde haben eine hochgradig defizitäre Lage der Begründung der Strafzumessung ergeben. Dafür mag es einerseits theoretische Gründe geben, diese sind aber – wie schon dargestellt – beherrschbar. In der Praxis werden vor allem andere, außerrechtliche Faktoren der Grund für unzureichende Begründungen sein: Zeit- und Kostendruck, Personalknappheit und die Arbeitsbelastung der Richter sind jedenfalls plausible Aspekte, führen sie doch zu mangelnder Zeit und damit zu knappen Urteilen. Plastisch führt hier *Wolting* aus:

„Wir haben die Zeit für die Darstellung ausführlicherer Abwägungen nicht. Eine Einzelrichterin beim Amtsgericht hat ein monatliches Pensum von etwa 40 Straf- und 20 OWi-Sachen. Ordnungswidrigkeiten sind in ihrer Komplexität den Strafsachen bei Einzelrichter/innen inzwischen mindestens vergleichbar. Also 60 Sachen im Monat bei 20 Arbeitstagen, das sind drei Sachen am Tag.“⁵⁰

Hierbei liegt es nahe, die Bereiche knapp auszugestalten, deren Begründung ohnehin schwierig ist, wie beispielsweise die Strafzumessung. Stattdessen erfolgt ein Rückzug auf bewährte Formulierungen, die aber oftmals zu oberflächlich und wenig aussagekräftig sind.⁵¹ Die rechtsstaatlich berechnete und notwendige Forderung nach einer besseren Begründung der Strafzumessung muss daher nicht nur theoretisch begleitet, sondern kann auch nur mit ausreichenden Ressourcen in der Justiz bewältigt werden. Solche scheinen zum aktuellen Stand aber weniger zur Verfügung zu stehen, weshalb der Einsatz künstlicher Intelligenz bei der Strafzumessung als ein möglicher Ausweg erscheint.

IV. Einsatz von sog. Künstlicher Intelligenz bei der Strafzumessung

Das führt nun zur entscheidenden Frage, ob die Technologie der sog. Künstlichen Intelligenz (KI) eine Lösung für die Begründungsproblematik bei der Strafzumessung liefern kann oder ob sich ihr Einsatz aufgrund der Begründungsprobleme verbietet. Der Einsatz von KI bei der Strafzumessung ist in verschiedenen Einsatzbereichen denkbar: So könnte KI dem Richter einen Entscheidungsvorschlag – entweder mit einem Strafkorridor oder mit einer Punktstrafe – unterbreiten, Strafbefehle automatisiert erstellen und dabei konkret Strafen festlegen, große Mengen von Entscheidungen aufbereiten und analysieren, um die Entscheidungsfindung des Gerichts zu unterstützen oder aber die Erstellung der Begründung für die Strafzumessungsentscheidung übertragen, nachdem sie das Gericht gefällt hat. Bevor der Einsatz von KI im Rahmen der Strafzumessung weiterverfolgt werden soll, sollen zunächst

⁴⁴ LG Braunschweig, Urt. v. 31.5.2022 – 9 Ks 15/21.

⁴⁵ LG Augsburg, Urt. v. 28.4.2023 – 3 KLS 201 Js 109552/22.

⁴⁶ AG Dortmund, Urt. v. 2.11.2023 – 729 Ds 114/23.

⁴⁷ AG Stuttgart-Bad Cannstatt, Urt. v. 30.1.2023 – 3 Cs 244 Js 98266/22.

⁴⁸ *Obert*, in: Hoven/Weigend (Fn. 31), S. 87 (101).

⁴⁹ *Obert* (Fn. 48), S. 104.

⁵⁰ *Wolting*, in: Hoven/Weigend (Fn. 31), S. 163 (167).

⁵¹ Vgl. *Wolting* (Fn. 50), S. 166.

einige technische und rechtliche Grundlagen dargestellt werden.

1. Deep Learning

Während eine allgemein konsentrierte Definition von KI noch aussteht, insbesondere weil der Begriff der Intelligenz noch nicht allgemein definiert ist,⁵² soll für die Zwecke dieser Untersuchung die Definition aus Art. 3 Nr. 1 der KI-Verordnung⁵³ zugrunde gelegt werden, nach der ein KI-System „ein maschinengestütztes System [...], das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können[, ist].“

Die Entwicklung von KI hat in den Jahren 2006 bis 2016 in rasanter Geschwindigkeit zugenommen und durch den Einsatz von Deep Learning einen Durchbruch erlangt, sodass KI-Systeme heute in vielen Bereichen eingesetzt werden (können).⁵⁴ Deep Learning als ein Unterfall des maschinellen Lernens ist der Kern der heute einsatzfähigen Systeme der generativen KI. Mit dieser Technologie gelingt das Training von KI-Systemen mittels tiefen, also mehrschichtigen neuronalen Netzen, die damit eine hohe Genauigkeit erlangen. Zusammengefasst und deutlich vereinfacht liegt ein mehrschichtiges selbstlernendes neuronales Netzwerk vor, welches von einem Eingangswert mittels mathematischer Formeln einen Ausgangswert produziert. Diese Operation erfolgt über mehrere Schichten von (künstlichen) Neuronen, die alle miteinander durch mathematische Operationen verbunden

sind. Beim Training dieses Netzwerks wird der Ausgabewert mit den bekannten (richtigen) Werten des jeweiligen Eingangs-Trainings-Datums verglichen und dann anhand dieses Ergebnisses selbstständig die einzelnen mathematischen Operationen des gesamten Netzwerks angepasst und gleichsam nachjustiert, um die Wahrscheinlichkeit eines korrekten Ergebnisses zu erhöhen. Mittels großer Datenmengen lassen sich so recht akkurate Netzwerke für gegebene Aufgaben trainieren.⁵⁵ Voraussetzung für das erfolgreiche Training ist also eine hinreichend große Menge an validierten Trainingsdaten,⁵⁶ die noch dazu möglichst „sauber“ sein müssen, weil eine fehlerhafte, diskriminierende oder voreingenommene Datengrundlage auch entsprechend anfällige Algorithmen produziert.⁵⁷ In der Kombination mit der grundsätzlichen Intransparenz der „Entscheidungen“ solcher Netzwerke, ergeben sich hier bereits Probleme, die auch für die Strafzumessung relevant werden. Die angesprochene Intransparenz basiert auf der Komplexität des Netzwerks, welches durch seine multiplen Schichten und Rückkoppelungen realistischere Weise von Menschen kaum nachvollzogen werden kann.⁵⁸ Die Rede ist von vom „black box problem“ der KI: Beobachtet werden können grundsätzlich nur die Inputs und Outputs der Algorithmen, nicht aber die komplexen und nichtlinearen inneren Vorgänge.⁵⁹

Auch wenn Ansätze einer „Explainable AI“ (XAI), also nachvollziehbarer und erklärbarer Künstlicher Intelligenz bereits entwickelt wurden und werden,⁶⁰ bleiben auch diese Ansätze aktuell unzureichend, weil auch die Erklärungen der XAI durchaus komplex sind und oft nur von Experten interpretiert werden können, sodass von einer echten Transparenz und Nachvollziehbarkeit der Entscheidung nicht gesprochen werden kann:

⁵² Bunte, Die chinesische KI-Revolution, 2020, S. 54, auch mit Nachweisen zur Intelligenz-Debatte; Sheikh/Prins/Schrijvers, Mission AI, 2023, S. 15 ff.; Russel/Norvig, Artificial Intelligence, A modern approach, 4. Aufl. 2021, S. 19 ff.; Heinrichs/Heinrichs/Rüther, Künstliche Intelligenz, 2022, S. 16 ff.; vgl. auch Ertel, Grundkurs Künstliche Intelligenz, 5. Aufl. 2021, S. 1 ff.; Independent High-level Expert Group on Artificial Intelligence, A Definition of AI: Main Capabilities and Disciplines, 8.4.2019, abrufbar unter https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341 (18.11.2024);

zu allgemeinen Begriffsbestimmungen exemplarisch Cope-land, Encyclopedia Britannica, abrufbar unter <https://www.britannica.com/technology/artificial-intelligence> (18.11.2024).

⁵³ Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz), ABl. 2024 EU Nr. L vom 12.7.2024.

⁵⁴ Ertel (Fn. 52), S. 13; zu einer kurzen Geschichte der KI dort S. 6 ff.

⁵⁵ Ertel (Fn. 52), S. 321 ff.; LeCun/Bengio/Hinton, nature 521 (2015), 436 ff.; Saker, SN Computer Science 2 (6/2021), 1 (5 ff.); Alzubaidi u.a., Journal of Big Data 8 (2021), 1 (4 ff.) jeweils auch zu den technischen Hintergründen.

⁵⁶ Statt vieler nur Dreyer/Schmees, CR 2019, 758 f.; Leopold, in: Altenburger/Schmidpeter (Hrsg.), CSR und Künstliche Intelligenz, 2021, S. 77 (83 ff.).

⁵⁷ Hüger, Künstliche Intelligenz und Diskriminierung, 2023, S. 76 ff.

⁵⁸ Beck u.a., Künstliche Intelligenz und Diskriminierung, 2019, S. 14 f., abrufbar unter https://www.plattform-lernende-systeme.de/publikationen-details/kuenstliche-intelligenz-und-diskriminierung-herausforderungen-und-loesungsansaeetze.html?file=files/Downloads/Publikationen/AG3_Whitepaper_250619.pdf (28.11.2024).

⁵⁹ v. Eschelbach, Philosophy & Technology 34 (2021), 1607 (1608); Castelvechi, nature 538 (2016), 20 (22 f.); Bathae, Harvard Journal of Law & Technology 31 (2018), 889 (901 ff.); Neu/Müller/Pothen/Zingel, Anwendungsfelder und Herausforderungen der Künstlichen Intelligenz, 2022, S. 79.

⁶⁰ Vgl. exemplarisch Rai, Journal of the Academy of Marketing Science 48 (2020), 137 (138 ff.); v. Eschelbach, Philosophy & Technology 34 (2021), 1607 (1615).

„One might come to some partial knowledge of the black box technology without understanding fully its inner workings.“⁶¹

XAI kann – so wird behauptet – auch selbst epistemische Unsicherheit in den Prozess einführen.⁶² Nicht ausgeschlossen ist freilich, dass die Weiterentwicklung der XAI-Technologie eine vollständige Nachvollziehbarkeit möglich macht – heute besteht sie jedenfalls aber nicht.

2. Die KI-Verordnung

Die EU hat mit der KI-Verordnung den ersten globalen Versuch einer Regulierung der KI-Technologie unternommen.⁶³ Für den vorliegenden Beitrag, der sich vor allem auf das Begründungsproblem der Strafzumessung fokussiert, ist eine umfassende Analyse der Verordnung und der rechtlichen Implikationen für die KI weder sinnvoll durchführbar noch notwendig. Daher soll es bei ein paar kurzen Ausführungen sein Bewenden haben: Die KI Verordnung geht von einer weiten Bestimmung des Begriffs der „Strafverfolgung“ aus. Hierunter sind gem. Art. 3 Nr. 46 KI-VO alle Tätigkeiten der Strafverfolgungsbehörden oder in deren Auftrag zur Verhütung, Ermittlung, Aufdeckung oder Verfolgung von Straftaten oder zur Strafvollstreckung, einschließlich des Schutzes vor und der Abwehr von Gefahren für die öffentliche Sicherheit, zu verstehen. Erfasst sind damit auch die Durchführung des Strafverfahrens und damit die Strafzumessung. Die KI-VO verfolgt einen risikobasierten Ansatz, den man in ähnlicher Weise beispielsweise aus dem Medizinprodukterecht kennt: KI-Systeme werden in unterschiedliche Risikoklassen eingeteilt und deren Einsatz je nach Risiko an unterschiedliche Voraussetzungen geknüpft: Hochrisiko-KI-Systeme müssen strenge Anforderungen erfüllen, damit ein Inverkehrbringen möglich ist (vgl. Abschnitte 2 und 3 der KI-VO). KI-Systeme mit begrenztem Risiko, insbesondere solche die mit Personen interagieren, unterliegen Transparenzpflichten gem. Art. 50 KI-VO und KI-Systeme mit minimalem Risiko unterliegen keiner spezifischen Regulierung durch die KI-VO. Daneben hat die KI-VO in Art. 5 solche Praktiken im KI-Bereich verboten, die ein unannehmbares Risiko bilden.

Im Rahmen der Strafverfolgung sind durch Art. 5 KI-VO sog. biometrische Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen grundsätzlich untersagt und nur in bestimmten Ausnahmesituationen zugelassen (z.B. zur Suche nach vermissten Personen). Andere Formen der KI zur Strafverfolgung sind jedenfalls nach Art. 5 KI-VO nicht verboten und unterliegen der Risikoklassifizierung. So führt

⁶¹ v. Eschelbach, *Philosophy & Technology* 34 (2021), 1607 (1615); Jiang/Kahai/Yang, *International Journal of Human-Computer Studies* 165 (2022), 1 (12).

⁶² Jiang/Kahai/Yang, *International Journal of Human-Computer Studies* 165 (2022), 1 (12).

⁶³ *Europäisches Parlament*, KI-Gesetz: erste Regulierung der künstlichen Intelligenz v. 8.6.2023, abrufbar unter <https://www.europarl.europa.eu/topics/de/article/20230601STO93804/ki-gesetz-erste-regulierung-der-kunstlichen-intelligenz> (18.11.2024).

beispielsweise Anhang II der KI-VO eine Reihe von Hochrisiko-KI-Systemen im Bereich der Strafverfolgung auf, beispielsweise KI-Systeme zur Lügendetektion, Beweismittelbewertung oder zur Bewertung persönlicher Merkmale und Eigenschaften oder vergangenen kriminellen Verhaltens einer natürlichen Person. Was bedeutet dies für einen Einsatz von KI-Systemen im Bereich der Strafzumessung? Es kommt auf die konkrete Ausgestaltung des Systems an – dieses ist dann je nach Risikoklasse einzuordnen und muss die spezifischen Anforderungen erfüllen. Eine allgemeingültige Aussage ist aufgrund des risikozentrierten Ansatzes daher nicht möglich. Festzuhalten bleibt, dass es jedenfalls nicht ausgeschlossen ist – und darauf kommt es hier an –, KI-Technologie im Rahmen der Strafzumessung zu verwenden.

3. KI bei der Strafzumessung: Auswirkungen der Begründungsprobleme

Damit ist nun der Weg bereitet, um darüber nachzudenken, ob KI die angesprochenen Probleme bei der Strafzumessung lösen kann bzw. welche Implikationen die Begründungsprobleme für den Einsatz von KI bei der Strafzumessung haben.

a) Trainingsdatenqualität

Wie schon angesprochen werden große Mengen an geeigneten Trainingsdaten für das Training einer KI benötigt. Dass diese Daten grundsätzlich zur Verfügung stehen, dürfte angesichts der Zahl von rechtskräftig verurteilten Personen, die in den letzten zehn Jahren immer über der Marke von 600.000 Personen lag,⁶⁴ nicht zu bestreiten sein. Diese Urteile sind zwar in der Welt und auch schriftlich niedergelegt, dürften allerdings nicht in einer entsprechend aufbereiteten Version vorliegen.⁶⁵ Für ein Training müsste eine umfassende Aufbereitung geschehen. Da der Algorithmus aufgrund des Feedbacks auf seine Berechnungen mit den Trainingsdaten lernt, müssen diese Daten menschlich verifiziert werden. Das führt schon zu einem ersten Problem, nämlich der Frage, wann eine Strafzumessungsentscheidung „richtig“ ist. Die Strafzumessung ist, das wurde bereits angesprochen, eine subjektiv geprägte Entscheidung innerhalb eines Ermessensspielraums: Es gibt – anders als früher vertreten – nicht die eine richtige Strafe (Punktstrafe),⁶⁶ sondern – nach der wohl herrschenden sog. Spielraumtheorie des BGH – vielmehr ein Kontinuum innerhalb dessen der Richter eine schuldangemessene Strafe

⁶⁴ *Statistisches Bundesamt*, Anzahl der rechtskräftig verurteilten Personen in Deutschland von 2011 bis 2022, abrufbar unter <https://de.statista.com/statistik/daten/studie/1060273/umfrage/rechtskraeftig-verurteilte-personen-in-deutschland/> (18.11.2024).

⁶⁵ Dreyer/Schmees, CR 2019, 758 (761) weisen darauf hin, dass besonders inhomogene Datensätze (z.B. unterschiedlicher Sprachgebrauch) und die große Bandbreite von Wortbedeutungen eine Herausforderung für das KI-Training darstellen.

⁶⁶ Frisch, *Revisionsrechtliche Probleme der Strafzumessung*, 1971, S. 167 ff.; A. Kaufmann, *Das Schuldprinzip, Eine strafrechtlich-rechtsphilosophische Untersuchung*, 1961, S. 261.

finden kann.⁶⁷ Dieses Kontinuum wird durch das Untermaßverbot nach unten und das Übermaßverbot nach oben begrenzt. Die gerade schon und gerade noch schuldangemessene Strafe bilden diese Grenzen und damit den Spielraum innerhalb dessen die konkrete Strafe auch unter Berücksichtigung der Strafzwecke und insbesondere präventiver Gedanken, zu finden ist.⁶⁸

In den aktuell vorliegenden Entscheidungen wird der Spielraum aber nicht angegeben (und muss es auch nicht werden)⁶⁹, sondern nur die konkrete Strafe, die das Gericht verhängt. Würde man nun die Spielraumtheorie nachbilden wollen, könnte der lernende Algorithmus gar nicht lernen, diesen Spielraum zu bilden, weil er in den Trainingsdaten nur konkrete Strafen vorfindet. Diese wiederum werden begleitet von – auch das wurde bereits erwähnt – qualitativ meist nicht ausreichenden Begründungen, womit die Gefahr mangelnder Trainingsdatenqualität – nämlich falsche oder diskriminierende Entscheidungen – besteht; denn der Algorithmus lernt aufgrund von Korrelationen zwischen Input- und Outputdaten. Hier können sich aber auch solche Korrelation aus den Trainingsdaten mathematisch ableiten lassen, die der menschliche Richter niemals als Abwägungsfaktor heranziehen würde. Auch die erwähnten regionalen Unterschiede in der Strafzumessung (Süd-Nord-Gefälle) aber auch andere tradierte Besonderheiten (z.B. besondere Härte gegenüber bestimmten Personengruppen oder bei bestimmten Deliktgruppen) würden sich durch die KI fortschreiben, weil diese gerade die (offensichtliche) Korrelation lernt, besonders, wenn wesentliche Unterscheidungskriterien die persönlichen Merkmale oder der Gerichtsort wären. Schon diese Punkte zeigen auf, dass die Trainingsdaten für eine Strafzumessungs-KI (noch) nicht in ausreichender Qualität vorliegen und jedenfalls aber Strategien zur Vermeidung von unerwünschten Ergebnissen gefunden werden müssen. Schließlich tut sich die KI – nach eigener Aussage – bei Einzelfallentscheidungen schwer, weil sie nur datenbasierte Vorschläge macht, aber die menschlichen Besonderheiten des Falles nicht berücksichtigen kann.⁷⁰ Zudem fehlen ihr die konzeptionellen Fähigkeiten (bislang in

den Daten), unbekannte Faktoren entsprechend verarbeiten zu können.⁷¹

Dreyer/Schmees weisen zudem auf ganz praktische Probleme hin: Jede Änderung der materiellen Rechtslage und -anwendung – sei es durch den Gesetzgeber oder auch die Rechtsprechung – würde ein komplett neues Training der KI notwendig machen und die dafür notwendigen Daten müssten zunächst generiert werden.⁷² Wenn aber jede höchstrichterliche Rechtsprechung die Rechtslage beeinflusst und ein neues Training notwendig macht, ist es im iterativen Prozess der Rechtsfindung kaum möglich, KI sinnvoll einzusetzen. Das dürfte auch für den Bereich der Strafzumessung gelten, ist diese doch nicht selten Gegenstand von Revisionen.

b) Nachvollziehbarkeit

Weiter dürfte auch die mangelnde Nachvollziehbarkeit von Strafzumessungsentscheidungen ein grundsätzliches Problem darstellen, das KI-Systeme an (rechtsstaatliche) Grenzen bringen dürfte. Denn wie das Gericht innerhalb des angesprochenen Spielraums zwischen Untermaß- und Übermaßverbot zu einer konkreten Strafe kommt, ist schlicht (oft) nicht nachvollzieh- oder erklärbar. Dieser Vorgang ist heute schon eine „black box“, ist er doch ein hochkomplexer Abwägungsvorgang, der von den Richtern auch nicht immer explizit gemacht oder begründet werden kann. Wie besprochen zeigt sich, dass oftmals die Abwägungsfaktoren nur aufgelistet, nicht aber zueinander in Bezug gesetzt werden. Welches Gewicht welcher Faktor hat, bleibt fast durchwegs im Unklaren. Wenn aber die menschlichen Richter den Vorgang der Abwägung nicht verbalisieren und umfassend darstellen können, wie soll die KI diesen Vorgang dann erlernen und – das ist die viel entscheidendere Frage – wie soll das Ergebnis kontrolliert werden? Wenn wir bereits heute Strafzumessungsentscheidungen nicht in einem engeren Sinne nachvollziehen, sondern lediglich auf ihre Plausibilität innerhalb des Spielraums prüfen können, dann potenziert sich diese mangelnde Nachvollziehbarkeit in der Kombination mit dem „black box-Problem“ der KI nochmals. Hinzuweisen ist auch darauf, dass die mangelnde Nachvollziehbarkeit der Entscheidungsweise der KI zu rechtsstaatlichen Problemen führen kann.⁷³

⁶⁷ So nur BGHSt 7, 28 (32); 20, 264 (266 f.); 24, 132 (133 f.); *Bußmann* (Fn. 12), § 46 Rn. 7; *Streng* (Fn. 12), § 46 Rn. 97 ff. jeweils m.w.N.

⁶⁸ Statt aller siehe nur BGHSt 29, 319 (321); *van Gemmeren* (Fn. 15), Rn. 882 ff., 1188; *Ziegler*, Das Strafurteil, 9. Aufl. 2021, Rn. 273 ff.; *Weber*, Einheitlichkeit und Transparenz in der Strafzumessung, 2024, S. 52 f.

⁶⁹ *van Gemmeren* (Fn. 15), Rn. 1188.

⁷⁰ GPT-4o (OpenAI), Prompt: „Kann KI Begründungen für Strafzumessungsentscheidungen generieren?“, Auszugsweise Antwort: „Strafzumessungsentscheidungen beinhalten oft ethische Überlegungen, die über eine rein mathematische oder datenbasierte Analyse hinausgehen. KI kann keine moralischen Urteile fällen, sondern nur datenbasierte Vorschläge machen. [...] Strafzumessungsentscheidungen erfordern eine individuelle Betrachtung jedes Falles. Obwohl die KI große Mengen an Informationen verarbeiten kann, könnte sie Schwierigkeiten haben, die einzigartigen menschlichen Aspekte jedes Falls vollständig zu berücksichtigen.“

⁷¹ GPT-4o (OpenAI), Prompt: „Wie geht KI damit um, wenn ihr unbekannte Variablen präsentiert werden?“, Auszugsweise Antwort: „Wenn eine KI auf völlig unbekannte Variablen stößt, die außerhalb des Trainingsbereichs liegen, könnte sie entweder gar keine Vorhersage treffen oder eine fehlerhafte bzw. nicht hilfreiche Antwort geben. Modelle sind in der Regel darauf optimiert, innerhalb eines bestimmten Datenraums zu arbeiten, und unbekannte Variablen können zu unsinnigen Ergebnissen führen. [...] In vielen Fällen fehlen der KI jedoch die konzeptionellen Fähigkeiten, die Komplexität unbekannter Faktoren voll zu erfassen.“

⁷² *Dreyer/Schmees*, CR 2019, 758 (760).

⁷³ Vgl. *Kaspar*, KriPoZ 2023, 1 (6).

Dieses Problem besteht auch dann, wenn die KI dem Gericht lediglich Entscheidungsvorschläge unterbreitet.⁷⁴ Denn dann würde – bei Übernahme dieser Vorschläge – letztlich ein „Abwägungsvorgang“ der KI nachträglich durch den Richter mit einer eigenen Abwägung unterfüttert. Damit würde lediglich ein vorgefundenes Ergebnis plausibilisiert, eine echte Abwägung hat aber niemals stattgefunden, sondern diese wird dem Ergebnis nur als Begründung beigelegt.⁷⁵

c) *Und nun?*

Der Einsatz von KI bei der Strafzumessung steht (noch) vor erheblichen Herausforderungen.⁷⁶ Die größte Herausforderung dürfte aktuell in der unzureichenden Trainingsdatenlage und -qualität liegen. Denn wenn der Mensch (noch) keine ausreichende Begründung für die Strafzumessung formuliert oder die „richtige“ Antwort nicht kennt, wie und wer soll dann die KI entsprechend trainieren? Sie könnte nur Korrelationen aus vorhandenen Urteilen herleiten. Die heute schon bestehenden Unzulänglichkeiten bei der Begründung der Strafzumessung übertragen sich dann auf die trainierte KI, aus unzureichenden Daten folgen dann unzureichende Algorithmen: „Garbage In, Garbage Out“⁷⁷. Künftige Entwicklungen können diese Probleme lösen, bis dahin bleibt der Einsatz von KI bei der Strafzumessung aber kritisch zu sehen.

V. Fazit

Das Begründungsproblem der Strafzumessung lässt sich durch KI – zum heutigen Zeitpunkt jedenfalls – nicht befriedigend auflösen. KI zur Begründung einer Entscheidung zu trainieren, würde voraussetzen, dass wir diese Entscheidung selbst begründen können, was in vielen Fällen an seine Grenzen stößt. Zudem bliebe dann die rechtsstaatliche Frage, ob die KI die Begründung für eine festgesetzte Strafe durch den Richter überhaupt übernehmen darf; denn der Richter ist es doch, der sich vor seiner Entscheidung die Begründung überlegen soll. Eine nachträgliche Erzeugung von Begründungen erscheint daher ausgesprochen problematisch. Vielmehr hat sich gezeigt, dass die Begründungsprobleme der Strafzumessung selbst dazu führen, dass dem Einsatz von KI-Systemen in der Strafzumessung (noch) erhebliche Hindernisse entgegenstehen. Ob sich diese jemals auflösen lassen, ist heute schwer zu prognostizieren; angesichts unserer eigenen, menschlichen Probleme bei der Begründung und Verbalisierung eines komplexen Abwägungsvorgangs ist dies aber grundsätzlich zweifelhaft. Ein spannender Einsatz von KI bei

der Strafzumessung könnte darin bestehen, die Spruchpraxis von Gerichten auf (unbewusst) diskriminierende Faktoren hin zu überprüfen; denn eine trainierte KI könnte hier Korrelationen aufzeigen, die es zu reflektieren gilt.⁷⁸

Bevor weiter über den Einsatz von KI bei der Strafzumessung nachzudenken ist, ist es zunächst notwendig, sich ausführlicher als bisher in Wissenschaft und Ausbildung mit der Strafzumessung zu beschäftigen: Es bedarf der Weiterentwicklung der Dogmatik der Strafzumessung und insbesondere der Reflexion der Begründungen von Strafzumessungsentscheidungen. Jedenfalls bei den Begründungen ließen sich deutliche Verbesserungen erreichen, beispielsweise durch die Angabe der Ober- und Untergrenze des Spielraums im Urteil, die Einordnung der Tat in qualitativ-relative Kategorien (z.B. leicht, mittel, schwer) oder durch die konsequente Darstellung der Bewertung und Inbezugsetzung der Abwägungsfaktoren. Welche weiteren Schritte man hin zu einer rationaleren und nachvollziehbaren Begründung der Abwägungsentscheidung gehen kann und soll, sollte Gegenstand weiterer Forschung sein.

Mit diesen ersten Denkanstößen lässt sich festhalten, dass der Einsatz von KI nur dann sinnvoll erscheint, wenn sich menschliche Defizite bei der Begründung von Strafzumessungsentscheidungen zuerst deutlich reduziert haben.

⁷⁴ Zum sog. „Automation Bias“ siehe *Greco*, RW 2020, 29; *Kaspar*, KriPoZ 2023, 1 (6); auch: *Kohn* (Fn. 28), S. 189 f.

⁷⁵ Zu einer anderen Bewertung, allerdings ohne spezifisch die Begründung in den Blick zu nehmen, gelangt *Kohn* (Fn. 28), S. 200 ff.: Verfassungsrechtliche Zulässigkeit.

⁷⁶ Zu weiteren Herausforderungen, insbesondere der technischen Infrastruktur sowie Vertrauensaspekten siehe *Kohn* (Fn. 28), S. 221 ff.

⁷⁷ Zu diesem geflügelten Wort vgl. exemplarisch *Kilkenny/Robinson*, Health Information Management Journal 47 (2018), 103; *de la Durantaye*, ZUM 2023, 645; *Kaspar*, KriPoZ 2023, 1 (6).

⁷⁸ Was nicht heißen muss, dass solche Korrelationen direkt Diskriminierung bedeuten. Die KI könnte aber bei der Reflexion der eigenen Spruchpraxis ein technisches Analyseinstrument sein.