

Die Vorhersage von Gerichtsentscheidungen – Methodische und inhaltliche Ausgangs- überlegungen

Prof. Dr. Christoph Sorge/Dr. Jochen Krüger, Saarbrücken*

Der Beitrag befasst sich mit aktuellen Fragen von Urteilsprognosen unter Einsatz von KI-Methoden. In diesem Zusammenhang werden auch einige mit dem allgemeinen Thema „Legal Tech und Recht“ verbundene Grundsatzfragen erörtert.

A. Problemstellung

Beim Generalthema „Legal Tech“¹ handelt es sich um Informationstechniken, die im juristischen Bereich zum Einsatz gelangen oder auch künftig gelangen können.² Dazu gehören auch die sog. Predictive Justice Tools (PJT), mit denen Gerichtsentscheidungen vorhergesagt werden sollen.³

Im Folgenden sollen einige der mit diesem Thema verbundenen Ausgangsüberlegungen angesprochen werden. Dafür werden zunächst die technischen Grundlagen der Urteilsvorhersage einschließlich der daraus resultierenden inhaltlichen Anforderungen dargestellt (B.). In Abschnitt C. soll mit Blick auf das Strafrecht sodann der Gedanke verdeutlicht werden, dass Prognosen ein elementarer Bestandteil der Rechtsordnung insgesamt sind. Daraus ergeben sich eine Vielzahl von potentiellen Anwendungsfeldern für das allgemeine Konzept – Einsatz von Künstlicher Intelligenz (KI) im Recht. Skizziert werden in diesem Zusammenhang daher auch einige der damit verbundenen rechtlichen und gesellschaftlichen Grundsatzfragen, die zurzeit verstärkt diskutiert werden. Der Beitrag schließt mit einer Zusammenfassung und einem Ausblick (D.), bei dem auch die Bedeutung der Legal-Tech-Diskussion für die Zukunft der juristischen Ausbildung kurz angesprochen wird.

B. Anmerkungen zu den technischen Grundlagen der Urteilsvorhersage einschließlich der daraus resultierenden inhaltlichen Anforderungen

I. Technische Grundlagen

Aus technischer Sicht wird eine Vorhersage von Gerichtsentscheidungen in aller Regel durch maschinelle Lernverfahren erfolgen, konkreter durch sogenannte „überwachte“ Verfahren (supervised machine learning)⁴. Das bedeutet, dass in einem ersten Schritt Trainingsdaten verwendet werden, um mittels eines Algorithmus ein Modell zu erzeugen. Anschließend wird dieses Modell dann auf weitere Daten angewendet. Bei der Vorhersage handelt es sich in der Terminologie eines Großteils der technischen Literatur um eine Klassifikationsaufgabe.

Die Trainingsdaten entsprechen in ihrer Struktur den Daten, die später als Grundlage der Vorhersage verwendet werden. Sollen etwa Entscheidungen anhand der jeweiligen Klageschrift vorhergesagt werden, bestehen auch die Trainingsdaten aus Klageschriften. Zusätzlich beinhalten sie aber bereits das vorherzusagende Attribut (z. B. die Zulässigkeit oder Begründetheit einer Klage). Durch das Lernverfahren kann dann das Modell erzeugt werden, also eine Repräsentation von Mustern in den Trainingsdaten, die eine Vorhersage des gesuchten Attributs für weitere Datensätze ermöglichen. Anhand von Testdaten – für die das vorherzusagende Ergebnis bekannt ist, die aber im Training nicht verwendet wurden – wird geprüft, wie genau diese Vorhersage ist. Welche Metrik für die Beschreibung der Genauigkeit verwendet wird und wie aussagekräftig diese Metrik ist, sollte bei der Einordnung der Vorhersagegenauigkeit genau betrachtet werden. Eine Genauigkeit – als Anteil korrekter Vorhersagen an allen Vorhersagen – von insgesamt 97 % dürfte beispielsweise in sehr vielen Anwendungsfällen maschineller Lernverfahren als hervorragendes Ergebnis gelten. Soll aber der Ausgang eines Verfassungsbeschwerdeverfahrens beim BVerfG vorhergesagt werden, ließe sich eine noch höhere Genauigkeit schlicht dadurch erzielen, dass immer ein Misserfolg der Beschwerde vorhergesagt wird. Der Anteil der erfolgreichen unter allen entschiedenen Verfassungsbeschwerden

* Der Autor Prof. Dr. Christoph Sorge ist Inhaber des Lehrstuhls für Rechtsinformatik an der Universität des Saarlandes. Der Autor Dr. Jochen Krüger ist Vizepräsident des AG Saarbrücken a. D. sowie wissenschaftlicher Mitarbeiter am zuvor genannten Lehrstuhl.

¹ Als Abkürzung für Legal Technology.

² Vgl. Hartung, in: Hartung/Bues/Halbleib, Legal Tech – Die Digitalisierung des Rechtsmarkts, 2018, S. 7.

³ Dazu näher Hoch, MMR 2020, 295 ff.

⁴ Zu diesem Begriff und der Abgrenzung zum nicht überwachten maschinellen Lernen vgl. Sugiyama, Introduction to Statistical Machine Learning, 2015, 3 f.

liegt seit 1995 in jedem Jahr unter 3 %.⁵ Von einem sehr guten Erfolg dieses Ansatzes sollte man deshalb aber nicht sprechen.

Grundsätzlich sind nahezu beliebige Attribute als Eingabedaten in ein maschinelles Lernverfahren möglich, wenn auch nicht jedes Verfahren für alle Arten von Attributen geeignet ist. Im Beispiel der Klageschrift ist etwa denkbar, deren Länge, das zuständige Gericht bzw. den zuständigen Spruchkörper oder den Verfahrensgegenstand (etwa „Diesel-Verfahren“) heranzuziehen. Diese Attribute sind dann noch in eine für das maschinelle Lernverfahren geeignete Repräsentation zu überführen.

Die Auswahl der zu betrachtenden Attribute benötigt Sachverstand. Einerseits kann ein maschinelles Lernverfahren Muster erkennen, die auch Experten im Vorhinein nicht aufgefallen waren, und somit können scheinbar unbedeutende Attribute für die gewünschte Vorhersage nützlich sein. Andererseits können – insbesondere bei kleinen Trainingsdatensätzen – aber auch rein zufällige, auf die Trainingsdaten beschränkte Zusammenhänge mit einem Attribut auftauchen und deren Berücksichtigung die Vorhersagekraft auf anderen Daten sogar verschlechtern. Dieses Risiko ist umso höher, je mehr Attribute einbezogen werden, so dass ein wahlloses Hinzufügen weiterer Attribute problematisch ist. Schließlich können Attribute durchaus die Vorhersagequalität verbessern, ohne aber für den konkreten Anwendungsfall nützlich zu sein. Stelle sich etwa heraus, dass in vergleichbaren Sachverhalten längere Klageschriften zu höheren Erfolgsaussichten führten, könnte dies im Rahmen einer wissenschaftlichen Untersuchung interessant sein. Für die Praxis muss das aber nicht gelten. Ein denkbarer Grund für diesen (hypothetischen!) Zusammenhang wäre das Vorliegen umfangreicher, die Klageschriften verlängernden Textbausteinsammlungen bei einschlägig spezialisierten Kanzleien. Diese können aufgrund ihrer Erfahrungen möglicherweise auch die Erfolgsaussichten besser einschätzen und deshalb einen höheren Anteil erfolgreicher Klagen vorweisen. Die Länge der Klageschrift wäre in diesem Beispiel mit den Erfolgsaussichten korreliert – aber ohne direkten Kausalzusammenhang. Würde ein maschinelles Lernverfahren ohne das Verständnis solcher Hintergründe in einer Kanzlei eingesetzt, wäre die Verwendung des Attributs „Länge der Klageschrift“ also nicht sinnvoll.

Eine besondere Schwierigkeit stellt sich dabei, Texte so zu repräsentieren, dass sie durch ein maschinelles Lernverfahren sinnvoll verarbeitet werden können. Hier sind unterschiedliche (und unterschiedlich aufwendige) Möglichkeiten der Vorverarbeitung denkbar.⁶ Üblich ist unter anderem die Entfernung von sogenannten Stoppwörtern,

die keine (in den Folgeschritten verwertbare) Information tragen, etwa Artikel oder Konjunktionen. Auch werden meist Stammformreduktionen (Stemming) angewendet, also Wörter auf ihren jeweiligen Wortstamm zurückgeführt.

Die Repräsentation des so bereinigten Texts kann dann u. a. durch einen sogenannten „Bag of Words“-Ansatz erfolgen.⁷ Hier wird im einfachen Fall lediglich gezählt, wie häufig jedes Wort vorkommt. Möglich ist aber auch die Betrachtung von N-Grammen statt einzelner Wörter. Ein N-Gramm ist die Zusammenfassung von jeweils N (z. B. zwei oder drei) aufeinanderfolgenden Wörtern. Aus „Die Klage ist zulässig und begründet“ könnte nach der Vorverarbeitung etwa „Klage zulässig begründet“ und daraus die 2-Gramme (Bigramme) „Klage zulässig“ und „zulässig begründet“ werden.

Dass in diesen Schritten Information verloren geht, ist offensichtlich, denn auch die Position von Wörtern oder Sätzen trägt Bedeutung. Ob der Bag-of-Words-Ansatz trotzdem ausreicht bzw. die Schwächen durch eine geeignete Vorverarbeitung ausgeglichen werden können, ist vom konkreten Anwendungsfall abhängig. Es sind aber durchaus vielversprechende Ergebnisse bereits mit diesem im Grunde sehr einfachen Ansatz erzielt worden. Auch hier gilt jedoch, dass manche Muster sich gerade in kleinen Trainingsdatensätzen rein zufällig ergeben könnten. Werden 100 Klageschriften analysiert, von denen 10 das Bigramm „Zahlung Kaufpreis“ enthalten und diese 10 Klagen alle erfolgreich sind, muss darin also noch keine verallgemeinerbare Erkenntnis liegen.

II. Grundlage und Ziel der Entscheidungsvorhersage

Wie oben bereits dargestellt, können ganz unterschiedliche Daten für die Entscheidungsvorhersage herangezogen werden. Zu Recht weist Hoch⁸ darauf hin, dass das Training eines Systems zur Entscheidungsvorhersage anhand von Sachverhalten, die durch das Gericht selbst (und mithin in Kenntnis des Ergebnisses) verfasst worden sind, problematisch sein kann. Einem praktischen Anwendungsfall, etwa in Bezug auf die Prüfung, ob eine noch einzureichende Klage aussichtsreich ist, entspricht dieses Vorgehen sicherlich nicht. Die für das Training maschineller Lernverfahren notwendige Anzahl an Datensätzen, die praktische Anwendungsszenarien abbilden (etwa die vollständigen Akteninhalte einiger tausend Verfahren aus einem Rechtsgebiet), stehen der Wissenschaft jedoch schlicht nicht zur Verfügung, so dass eine Beschränkung auf die Entscheidungen selbst oft notwendig ist. Erkenntnisse für die Forschung sind dennoch möglich, denn maschinelle Lernver-

⁵ Bundesverfassungsgericht, Jahresstatistik 2020, Anteil der stattgegebenen an den entschiedenen Verfassungsbeschwerdeverfahren pro Jahr, https://www.bundesverfassungsgericht.de/DE/Verfahren/Jahresstatistiken/2020/gb2020/A-IV-2.pdf?__blob=publicationFile&v=2, Abruf v. 14.7.2021.

⁶ Ausführlich zu möglichen Vorverarbeitungsschritten *Dop*, Hands-On Natural Language Processing with PyTorch 1.x, 2020, S. 88 ff.

⁷ So etwa bei *Sulea/Zampieri/Vela/van Genabith*, Predicting the Law Area and Decisions of French Supreme Court Cases, Proceedings of the International Conference Recent Advances in Natural Language Processing, 2017, S. 716–722, https://doi.org/10.26615/978-954-452-049-6_092.

⁸ Hoch (Fn. 3), 296.

fahren können etwa statistische Zusammenhänge sichtbar machen und somit eine Grundlage für weitergehende, empirische Untersuchungen darstellen. Auch in anderen Wissenschaftsgebieten müssen Schlussfolgerungen aus unvollständigen oder systematisch verzerrten Informationen gezogen werden, ohne dass diese Forschung deshalb nutzlos wäre. Im Folgenden gehen wir daher auch von Systemen zur Entscheidungsvorhersage aus, die lediglich auf Basis der Sachverhalte aus den Entscheidungen selbst trainiert worden sind.

III. Zur Verfügbarkeit von einschlägigen Entscheidungen

Für eine valide Prognose wäre es zunächst hilfreich, wenn auch eine ausreichende Anzahl von einschlägigen Entscheidungen des zuständigen Spruchkörpers vorhanden und für eine Analyse zugänglich wäre. Damit wird bereits deutlich, dass die Themenbereiche „Vorhersage von Gerichtsentscheidungen“ und „möglichst umfassende Veröffentlichungspraxis des zuständigen Gerichts“ auf das Engste miteinander verbunden sind.

Dies erklärt auch, warum Fragen der Vorhersage gerichtlicher Entscheidungen auf nationaler Ebene⁹ typischerweise mit Blick auf Obergerichte thematisiert werden. Nur hier gibt es eine regelmäßige und leicht zugängliche Veröffentlichungspraxis. Hinzu kommt in diesem Zusammenhang ein weiterer Aspekt: Vorhersagen aufgrund der Analyse früherer Entscheidungen haben insbesondere dann eine Chance auf Verlässlichkeit, wenn der Spruchkörper personell weitestgehend identisch ist und die relevanten Ausgangsentscheidungen nicht von tatsächlichen Umständen wie z. B. der Glaubhaftigkeit einer Zeugenaussage abhängen.¹⁰ Daher eignen sich z. B. normale Schöffengerichte als Tatsacheninstanz bei einem Amtsgericht mit einem Vorsitzenden und zwei oftmals wechselnden Laienrichtern (Schöffen)¹¹ weniger für das Thema Urteilsprognose. Dies gilt umso mehr, als diese Schöffen den Vorsitzenden theoretisch überstimmen können.¹² Die zuvor skizzierten Gesichtspunkte könnten die Auffassung nahe legen, das Thema „Vorhersagen von Gerichtsentscheidungen“ schwerpunktmäßig auf revisionsrechtlicher Ebene zu erörtern. Jedenfalls sind die Eigenheiten des konkreten Anwendungsfalls bei Fragen der Prognosemöglichkeiten mit zu berücksichtigen.

IV. Schwierigkeiten der Prognose bei neuartigen Sachverhalten oder einschlägigen Gesetzesänderungen

Selbst wenn die zuvor skizzierten Rahmenbedingungen gegeben sind, zeigen sich bei näherer Analyse weitere Schwachstellen für eine zuverlässige Prognose. Dies gilt insbesondere mit Blick auf neuartige Sachverhalte oder einschlägige Gesetzesänderungen.¹³ In diesem Zusammenhang kann z. B. der Themenkomplex „Mordvorwurf bei illegalem Autorennen in der Stadt mit Todesfolge bei Unbeteiligten“ genannt werden. Hier gibt es erst seit 2017 eine intensivere wissenschaftliche Diskussion, auch im Zusammenhang mit einschlägigen obergerichtlichen Entscheidungen.¹⁴ Zudem ist zu berücksichtigen, dass durch Gesetz vom 30.9.2017 § 315d StGB und damit eine neue Vorschrift über verbotene Kraftfahrzeugrennen eingeführt wurde.¹⁵ Dazu liegen zwar bereits einige Entscheidungen vor.¹⁶ Es gibt jedoch noch keine umfassende gesicherte Rechtsprechung.

C. Anmerkungen zu weiteren Fragestellungen im Rahmen von rechtlich relevanten Prognosen

Der vorliegende Beitrag befasst sich im Ansatz mit Fragen von IT-gestützten Urteilsprognosen und den dazu erforderlichen Rahmenbedingungen. Das damit verbundene gesellschaftliche und rechtliche Problempotential dürfte jedenfalls auf den ersten Blick als eher gering eingestuft werden. Vorrangiger Ansprechpartner für Urteilsprognosen in der Praxis ist die Rechtsanwaltschaft. Diese ist darin geschult, im Rahmen ihrer Beratungspflicht auf etwaige Risiken der Rechtsverfolgung hinzuweisen. Entspricht die abschließende tatsächliche Entscheidung des Gerichts nicht der Anwaltsprognose, ist das rechtliche Gesamtsystem nicht erschüttert.

I. Problematische Aspekte im Konzept der Urteilsprognose

Bei weiterer Analyse werden jedoch auch problematische Aspekte im Konzept der Urteilsprognosen sichtbar. So ist das Recht oftmals nicht auf eindeutige Antworten angelegt. Offenkundig ist dies insbesondere bei Sachverhalten, in denen die sog. Generalklauseln wie z. B. § 138 Abs. 1 BGB¹⁷ eine zentrale Rolle spielen. Aber auch allgemein

⁹ Zur Vorhersage von Entscheidungen des EGMR die Nachweise in Fußnote 13 bei Hoch (Fn. 3), 296.

¹⁰ Zu diesem – bisher wenig diskutierten – Aspekt auch Hoch (Fn. 3), 298; Enders, JA 2018, 721 (727).

¹¹ § 29 Abs. 1 S. 1, 43 Abs. 1, 2 Gerichtsverfassungsgesetz (GVG).

¹² § 192 Abs. 1, 3 GVG, § 196 Abs. 1 GVG, § 197 GVG – danach stimmen (S. 2) die Schöffen vor dem Richter.

¹³ Dazu auch von Bünau, in: Breidenbach/Glatz, Rechtshandbuch Legal Tech, 1. Auflage 2018, S. 57 Rn. 38.

¹⁴ Vgl. dazu z. B. die Entscheidung BGH NJW 2020, 2900 ff. mit Anmerkung von Grünwald; zu früheren Entscheidungen näher Preuß, NZV 2018, 345 ff. m. w. N.

¹⁵ Vgl. BGBI I, 3532, dazu näher Fischer, Kommentar StGB, 67. Auflage 2020, § 315d Rn. 1 ff.

¹⁶ Z. B. LG Osnabrück BeckRS 2021, 5342 bzw. LG Koblenz NZV 2021, 222 mit weiterführenden Hinweisen von Fromm.

¹⁷ Gem. § 138 Abs. 1 BGB ist ein Rechtsgeschäft, das gegen die guten Sitten verstößt, nichtig.

gilt für das deutsche Recht,¹⁸ dass es im Bereich der Einzelfallentscheidung tendenzmäßig auf Entwicklung und Anpassung unter neuen Rahmenbedingungen angelegt ist. Zwar sollte im Idealfall der Richter die Eigenheiten seiner Person zurücknehmen und dem Regelungsgehalt des Gesetzes Rechnung tragen.¹⁹ Aber selbst bei ernsthaftem Bemühen lässt sich nicht vermeiden, dass dabei auch das individuell geprägte Vorverständnis die endgültige Entscheidung mit beeinflussen kann.

Dies führt im Ergebnis jedoch zu einem elementaren Interesse der Prognosewissenschaft an den personenbezogenen Daten des Entscheiders selbst. Daraus können erkennbar Gefahren für das Recht auf Privatheit auf der Richterebene entstehen. Denn in den zuvor skizzierten Überlegungen ist auch eine Analyse des allgemeinen Sozialverhaltens der Entscheider mit angelegt. Dieser Umstand hat dazu geführt, dass z. B. Frankreich ein in der wissenschaftlichen Diskussion viel beachtetes Gesetz erlassen hat. Darin wird geregelt, dass Identitätsdaten von Richtern und Angehörigen der Justiz nicht Gegenstand eines analytischen Verfahrens sein dürfen, um Entscheidungen zu vergleichen, zu bewerten oder zu Vorhersagezwecken zu nutzen.²⁰

II. Prognosen als elementarer Bestandteil des allgemeinen rechtlichen Regelwerks

Prognosen sind allgemein ein elementarer Bestandteil des rechtlichen Regelwerks. Dies lässt sich besonders deutlich am Beispiel des Strafrechts aufzeigen. So hängt die Strafaussetzung zur Bewährung gem. § 56 Abs. 1, 2 StGB von einer günstigen Sozialprognose ab. Eine noch gravierendere Bedeutung für die rechtliche Entscheidung hat die Prognose im Bereich der sog. Maßregeln der Besserung und Sicherung nach §§ 61 ff. StGB.²¹ Zentraler Maßstab des Rechts ist dabei nicht die Schuld des Betroffenen, sondern der Grad der zu erwartenden Gefährlichkeit.²² Für diese Fragen benötigt ein Gericht oftmals die Hilfe von erfahrenen Sachverständigen.²³

Die rechtlichen Einzelheiten stehen hier nicht im Vordergrund. Wichtig ist der allgemeine Gedanke, dass Prognosen in diesem Bereich eine ausschlaggebende Rolle spielen können. Daher kann nicht überraschen, dass in den USA insoweit auch PJT eingesetzt werden.²⁴ Bekannt geworden ist z. B. die Anwendung der Software „COMPAS“.²⁵ Das System errechnet aus Antworten auf 137 Fra-

gen, wie hoch die Wahrscheinlichkeit einer weiteren Straftat bei der analysierten Person ist. Auf dieser Ebene hat diese technikorientierte Methode eine ganz andere soziale und rechtliche Wirkung als bei der bloßen Vorhersage von Gerichtsentscheidungen. In diesem Bereich werden aber auch allgemeine Grundsatzprobleme beim Einsatz von KI im Recht sichtbar. Dies gilt etwa für das Stichwort „diskriminierende KI“. So ist die Software „COMPAS“ z. B. in der Kritik, weil sie bei Prognoseentscheidungen Afroamerikanern ein höheres Rückfallrisiko zuordnet als Weißen.²⁶ Die Abweichung zwischen der – negativen – Prognose und der tatsächlichen Rückfallhäufigkeit ist bei ihnen doppelt so hoch wie bei Weißen.²⁷

Derartige Grundsatzfragen sind nicht auf die USA beschränkt. In Deutschland ist der Einsatz von KI im Recht ebenfalls ein viel erörtertes²⁸ und aktuelles²⁹ Thema. Dies gilt z. B. für den Bereich des sog. Predictive Policing. Dort treffen Algorithmen Vorhersagen über Straftaten, Straftäter oder Tatorte.³⁰ Allgemein bietet sich dieses Konzept als Methode der vorausschauenden Kriminalintervention an.³¹ Denkbar ist auch, dass derartige Prognosesysteme nicht nur vom Staat, sondern z. B. von Arbeitgebern zur Vermeidung von Diebstählen im eigenen Betrieb eingesetzt werden.³² Daraus ergeben sich erkennbar zentrale gesellschaftliche und rechtliche Grundsatzfragen. Erörtert wird z. B. die Frage, ob eine großflächige Anwendung von automatisierten Entscheidungssystemen mit dem Demokratieprinzip³³, dem Rechtsstaatsprinzip³⁴ und dem Menschenbild des Grundgesetzes³⁵ zu vereinbaren ist.

D. Zusammenfassung und Ausblick

Der Beitrag hat sich zunächst mit Fragen der Urteilsprognosen unter Einsatz von KI-Methoden befasst. Verdeutlicht werden sollte dabei auch, dass Prognosen ein elementarer Bestandteil der Rechtsordnung insgesamt sind. Aus den Überlegungen resultiert ein klares Ergebnis: Die allgemeine Digitalisierung in der Gesellschaft macht auch vor der Justiz nicht Halt.³⁶ Diese Entwicklung muss von gesellschaftlichen, ethischen und rechtlichen Überlegun-

¹⁸ Zu den methodischen Unterschieden ggü. den USA – dort vorrangige Orientierung an Präzedenzfällen – *Hoch* (Fn. 3), 297.

¹⁹ Dazu näher *Sommer*, ZRP 2017, 60.

²⁰ Dazu Redaktion MMR-Aktuell 2019, 418742 – zu den Einzelheiten des Gesetzes *Hoch* (Fn. 3), 300.

²¹ Z. B. Unterbringung in einem psychiatrischen Krankenhaus oder Sicherungsverwahrung.

²² Dazu näher *Fischer* (Fn. 15), Vorbem. § 61 StGB, Rn. 3.

²³ Zu den rechtlichen Rahmenbedingungen bei Prognosen im Strafverfahren näher *Boetticher u. a.*, NSTZ 2019, 553 ff.

²⁴ Dazu näher *Steege*, MMR 2019, 715 ff.

²⁵ Correctional Offender Management Profiling for Alternative Sanctions.

²⁶ Zu Einzelheiten *Steege* (Fn. 24), 716; *Martini*, JZ 2017, 1017 (1018).

²⁷ *Martini* (Fn. 26), 1018.

²⁸ Vgl. z. B. *Rollberg*, Algorithmen in der Justiz, 1. Auflage 2020, m. w. N. im dortigen Literaturverzeichnis.

²⁹ In diesem Zusammenhang z. B. *Willcke*, MMR-Aktuell 2020, 425656.

³⁰ Dazu *Steege* (Fn. 24), 717.

³¹ In diesem Zusammenhang der Einsatz der Software „Pre Crime Observation System“ (PreCobs) in deutschen Bundesländern – dazu näher *Singelstein*, NSTZ 2018, 1 ff.

³² Zu diesem Beispiel und den daraus folgenden Grundsatzfragen *Steege* (Fn. 24), 718.

³³ *Klaas*, MMR 2019, 84 ff.

³⁴ *Völmann*, DÖV 2021, 474 (478), u. a. mit dem Hinweis, dass z. B. durch den Einsatz von Legal Tech auch die Rechtsweggarantie gestärkt werden kann.

³⁵ Dazu z. B. *Golla*, DÖV 2019, 673 ff.

³⁶ *Rollberg* (Fn. 28), S. 22.

gen begleitet werden. Insoweit besteht ein Grundkonsens sowohl auf europäischer als auch nationaler Ebene. So hat die EU-Kommission im April 2021³⁷ den nach eigener Darstellung weltweit ersten Rechtsrahmen für eine vertrauenswürdige KI vorgelegt, der u. a. den Weg für eine ethische³⁸ Technik ebnen soll. Auf bundesdeutscher Ebene ist stellvertretend das Gutachten der Datenethikkommission vom Oktober 2019 zu nennen.³⁹ Vor diesem Hintergrund erscheint auch eine weitere Prognose relativ gesichert. Angesichts der zunehmenden Digitalisierung im Recht wird das allgemeine Thema „KI und Recht“ auch Auswirkungen auf die juristische Ausbildung haben bzw. haben müssen.⁴⁰ Dies gilt sowohl mit Blick auf die neuen Möglichkeiten⁴¹, aber auch Risiken des Einsatzes von Algorithmen im Recht. Zu diesen Risiken gehört die typische Intransparenz und die damit oftmals verbundene latente Diskriminierung.⁴² Allgemein fehlt Algorithmen ein ethischer Kompass.⁴³ Ernst zu nehmen ist auch die Gefahr, dass ein großflächiger Algorithmeinsatz einer – im Ergebnis notwendigen – Fortentwicklung des Rechts im Wege steht und daher zu einem „Einfrieren des Rechts“ führen kann.⁴⁴ Nicht ohne Grund lautete das Motto des 30. EDV-Gerichtstags 2021 in Saarbrücken „Mensch oder Maschine – Wer prägt die Zukunft der Rechtsanwendung?“.

³⁷ Vgl. *EU-Kommission*, Mitteilung v. 21.4.2021, Für vertrauenswürdige Künstliche Intelligenz: EU-Kommission legt weltweit ersten Rechtsrahmen vor, https://ec.europa.eu/germany/news/20210421-kuenstliche-intelligenz-eu_de, Abruf v. 14.7.2021; vgl. zu den Vorarbeiten näher *Unger*, ZRP 2020, 234 ff.

³⁸ Vgl. in diesem Zusammenhang auch die Europäische Ethik-Charta über den Einsatz Künstlicher Intelligenz in der Justiz und in ihrem Umfeld, Dezember 2018.

³⁹ *Schlee*, ZD-Aktuell 2019, 06897; *Kelber*, ZD 2020, 73 ff.

⁴⁰ Zu derartigen Fragen näher *Zwickel*, JA 2018, 881 ff.; *Herberger*, NJW 2018, 2825 ff. mit der Forderung nach einem KI-bezogenen Rechtsverständnis; zur aktuellen Diskussion näher *Omlor/Meister*, ZRP 2021, 59 ff.

⁴¹ Dazu näher *Buchholtz*, JuS 2017, 955.

⁴² Dazu z. B. *Martini* (Fn. 26), 1018; speziell zu Fragen der Diskriminierungsgefahr beim Einsatz von KI *Leicht/Karst/Zimmer*, Diskriminierende Frauenfeindlichkeit – KI als Spiegel unserer Gesellschaft, Tagungsband IRIS 2020, 73 ff.

⁴³ So die Formulierung von *Martini* (Fn. 26), 1018.

⁴⁴ *Enders* (Fn. 10), 727.