

Automatisierte rechtliche Entscheidungsfindung – kommt der RoboJudge?¹

„Die Juristen könnte man erst durch Roboter und Computer ersetzen, wenn man wüsste, was Juristen tun. Und es ist eben schwierig, herauszufinden, was sie eigentlich tun.“²

Prof. Dr. Thomas Riehm

1 Einführung: Von guten und bösen Geistern

Ein Gespenst geht um: der „RoboJudge“³ – Algorithmen, gar „Künstliche Intelligenz“ sollen die Rechtsfindung oder gleich die Rechtsprechung übernehmen. Für die Apologeten der damals noch jungen Rechtsinformatik in den 1970er Jahren war das Gespenst der gute Geist, der Rechtsprechung endlich objektiv macht und allzu

¹ Schriftliche Fassung des Vortrags im Rahmen der Göttinger Kolloquien zur Digitalisierung des Zivilverfahrensrechts am 14.01.2022. Alle im Beitrag zitierten Internetadressen wurden zuletzt am 27.02.2023 abgerufen.

² *Tom Braegelmann* auf LinkedIn im Dezember 2021 (https://www.linkedin.com/posts/braegel_die-juristen-könnte-man-erst-durch-roboter-activity-6890016045705629696-FQH2).

³ S. etwa *Raabe/Wacker/Oberle u. a.*, *Recht ex machina*, 2012; *Kotsoglou*, JZ 2014, 451 ff.; *Engel*, JZ 2014, 1096 ff.; *Adrian*, RTh 48 (2017), 77 ff.; *Fries*, *Zeit für einen RoboJudge?*, 9.3.2018 (https://www.lto.de/persistent/a_id/27435/); *Erbguth*, DRiZ 2018, 130 ff.; aus früherer Zeit bereits *D'Amato*, *Georgia L.R.* 11 (1977), 1277 ff.; *Susskind*, *MLR* 49 (1986), 125 ff.

menschliche, ihren persönlichen Vorbehalten und Vorurteilen unterworfenen Richter durch den unbestechlichen, nüchtern und neutral urteilenden Computer ersetzt.⁴ Dieselbe Vorstellung findet sich in modernem Gewand auch heute noch in dem Hinweis, dass „das Demokratie- und Rechtsstaatsprinzip, wonach das Recht abstrakt generell für alle gleich gelten sollte, besser umgesetzt werden dürfte, wenn der juristische Umsatz, also das Recht, von der Maschine und nicht durch einen menschlichen Richter ‚konstruiert‘ wird.“⁵

Für andere ist es der böse Geist, der die Rückkehr zur Begriffsjurisprudenz einleitet,⁶ Vorurteile und Diskriminierungen versteinert,⁷ die Rechtsprechung entmenschlicht und sogar gegen das verfassungsrechtlich verbürgte Recht auf den gesetzlichen Richter verstößt.⁸ Und, wie das bei Gespenstern so ist: Die einen glauben an die Möglichkeit ihrer Existenz,⁹ die anderen halten sie für reine Hirngespinnste, die in der Realität nicht vorkommen können.¹⁰

Bei näherer Betrachtung haben die Befürworter und Kritiker von „RoboJudges“ allerdings ganz unterschiedliche Typen rechtlicher Entscheidungen vor Augen:¹¹ Die Befürworter behandeln überwiegend¹² einfache explizite Algorithmen im Sinne rechtlicher Expertensysteme, die einzelne Rechtsfragen in formalisierter Weise beantworten können, indem sie mit vorstrukturierten Sachverhalten „gefüttert“ werden und diese nach einem vorgegebenen Regelsystem verarbeiten, und deren Ergebnisse Richter (oder Anwälte¹³) bei der Arbeit lediglich unterstützen, aber nicht ersetzen sollen.¹⁴ Die Gegner denken dagegen vorwiegend an Systeme, die vollkommen selbständig Sachverhalte ermitteln, Gesetze auslegen und anwenden und damit Gerichte dereinst in Gänze ersetzen sollen.¹⁵

⁴ S. etwa *D'Amato*, Georgia L.R. 11 (1977), 1277, 1277.

⁵ *Adrian*, RTh 48 (2017), 77, 112.

⁶ *Zippelius*, Juristische Methodenlehre, 11. Aufl. 2012, 90 f.

⁷ *Erbguth*, DRiZ 2018, 130 ff.

⁸ *Enders*, JA 2018, 721, 723; Länderarbeitsgruppe Legal Tech: Herausforderungen für die Justiz, 2019 (https://www.schleswig-holstein.de/DE/landesregierung/ministerien-behoerden/II/Minister/Justizministerkonferenz/Downloads/190605_beschlusse/TOPI_11_Abschlussbericht.pdf?__blob=publicationFile&v=1), S. 60; *Winkelmann*, LTZ 2022, 163, 165.

⁹ Die Möglichkeit eines Richterautomaten will etwa *Adrian*, RTh 48 (2017), 77 ff. bewiesen haben.

¹⁰ Die Unmöglichkeit eines Richterautomaten will *Kotsoglou*, JZ 2014, 451 ff. bewiesen haben.

¹¹ Auf die Bandbreite denkbarer „RoboJudges“ hinweisend auch *Klug*, in: JbJZivRWiss 2019, 2020, S. 63, 63.

¹² Eine Ausnahme bildet insoweit *Adrian*, RTh 48 (2017), 77 ff.

¹³ Dazu s. etwa *Wagner*, BB 2017, 898 ff.; *Postinett*, Künstliche Intelligenz: Die Robo-Anwälte kommen, 17.5.2016 (<https://www.handelsblatt.com/13601888.html>).

¹⁴ Etwa *Burr*, BB 2018, 476 (Prüfung des § 6a GrEStG); *Fries*, in: Breidenbach/Glatz, Rechtshandbuch Legal Tech, 2018, Kap. 7.3 (Durchsetzung von Fluggastrechten); *Klug*, in: JbJZivRWiss 2019, S. 63 (Prüfung des Eigentumsübergangs nach §§ 929 ff. BGB).

¹⁵ Etwa *Kotsoglou*, JZ 2014, 451 ff.; *D'Amato*, Georgia L.R. 11 (1977), 1277 ff.

In diesem Spannungsfeld soll der nachfolgende Beitrag Möglichkeiten und Grenzen rechtlicher Entscheidungen durch Algorithmen ausloten. Dabei beschränkt sich die Untersuchung auf die Entscheidung von zivilrechtlichen Streitigkeiten, wie sie heute die Gerichte beschäftigen. Weder die Anwaltperspektive noch andere Rechtsgebiete werden näher betrachtet.

2 Deterministische Algorithmen und KI¹⁶

Im Hinblick auf den Einsatz von Algorithmen zur rechtlichen Entscheidungsfindung werden gemeinhin zwei verschiedene Arten von Algorithmen unterschieden: deterministische (explizite) Algorithmen auf der einen Seite und Machine-Learning-Algorithmen (auch: implizite Algorithmen) auf der anderen Seite, die häufig auch mit dem schillernden Begriff „Künstliche Intelligenz“ belegt werden.

Als deterministische oder explizite Algorithmen sollen im Folgenden solche Algorithmen bezeichnet werden, die zur Lösung von Aufgaben bestimmt sind, welche vorab klar definiert sind. Es handelt sich um endliche Wenn-dann-Strukturen, die zu einem bestimmten Input stets denselben vorher bestimmten Output generieren. Dabei sind sämtliche Ergebnisse, die der Algorithmus erzielen kann, und die hierfür erforderlichen Bedingungen im Voraus festgelegt. Auch wenn solche Algorithmen sehr komplex sein können, besteht ihr hier in den Vordergrund gestelltes Charakteristikum darin, dass ihr Verhalten vom Programmierer ex ante vollständig festgelegt wird und für Menschen zumindest in abstracto nachvollziehbar ist. Beispiele für solche Algorithmen sind einfache Taschenrechner, aber auch Unterhaltsberechnungen der gängigen Programme oder die automatische Steuerberechnung beim Finanzamt.

Demgegenüber können Algorithmen, die auf Machine Learning-Technologien beruhen, auch Aufgaben lösen, die ihnen nicht zuvor explizit einprogrammiert wurden. Derartige Algorithmen „lernen“ aus großen Zahlen von Beispielen und entwickeln sich gegenüber dem Stand zum Zeitpunkt ihrer Programmierung durch „Training“ weiter. Der Ablauf dieses „Trainings“ wird zwar von Menschen gesteuert. Am Ende dieses „Trainings“ steht jedoch ein Algorithmus, dessen Verhalten im Einzelnen für Menschen nicht mehr nachvollziehbar ist, auch nicht *in abstracto*. Die Komplexität des Algorithmus' lässt sich nicht mehr in Wenn-dann-Strukturen übertragen, die für Menschen verständlich wären, selbst wenn sie im eigentlichen Sinne noch deterministisch sind. Typische Beispiele derartiger Algorithmen sind solche zur Sprach- und Bilderkennung, zur Erkennung von Mustern in großen Datenmengen, aber auch zum automatischen Generieren von Texten oder Bildern.

¹⁶ S. zu diesen Abgrenzungsfragen eingehend *Stiemerling*, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning, 2020, Kap. 2.1.

3 Richterliche Entscheidungsfindung

Betrachtet man nun die richterliche Entscheidungsfindung vor dem Hintergrund dieser Typen von Algorithmen, so scheint diese auf den ersten Blick einem deterministischen Algorithmus zu entsprechen.¹⁷

3.1 Der Justizsyllogismus als Algorithmus

Klassisch wird die richterliche Entscheidungsfindung durch den sogenannten Justizsyllogismus beschrieben. Danach besteht die Rechtsanwendung aus den drei Elementen Obersatz, Untersatz und Subsumtion.¹⁸ Der Obersatz lautet in abstrakter Beschreibung: „Für Tatbestand T (bestimmt durch die Merkmale T₁ bis T_n) gilt die Rechtsfolge R.“ Im Untersatz wird festgestellt, dass ein Sachverhalt S die Merkmale T₁ bis T_n enthält. Unter dieser Voraussetzung ergibt die Subsumtion, dass für Sachverhalt S die Rechtsfolge R gilt.

Dieses Vorgehen lässt sich unmittelbar mit der Anwendung eines Algorithmus vergleichen: Der Obersatz beschreibt den Algorithmus selbst, also die anzuwendende Wenn-Dann-Struktur. Der Untersatz beschreibt den Input, und die Subsumtion ergibt den Output.

Am Beispiel einer Entschädigungsforderung nach der Fluggastrechteverordnung könnte das folgendermaßen aussehen: Der Obersatz/Algorithmus lautet „verfügt ein Fluggast über eine bestätigte Buchung für einen Linienflug, der in der EU startet oder von einer EU-Airline ausgeführt wird, und wird ihm die Beförderung auf diesem Flug verweigert, obwohl er sich rechtzeitig zur Abfertigung eingefunden hat, oder kommt der Flug mit einer Verspätung von mehr als 3 Stunden am Zielort an, so hat der Fluggast Anspruch auf Entschädigung in Höhe von 250/400/600 €, sofern dies nicht auf außergewöhnliche Umstände zurückgeht.“ Ein entsprechender Untersatz bzw. Input wäre „X hatte eine bestätigte Buchung für den Iberia-Flug IB045 von München nach Barcelona am 3.6.2020, planmäßige Ankunft 12:15 Uhr und kam um 18:00 Uhr in Barcelona an.“ Die Subsumtion bzw. der Output würde dann lauten „X hat einen Anspruch gegen Iberia auf Entschädigung in Höhe von 250 €.“

Das sieht zunächst so aus, als könnte diese Fallkonstellation leicht algorithmisch abgebildet und die rechtliche Entscheidungsfindung automatisiert werden. Doch ist das wirklich so einfach?

¹⁷ S. hierzu auch überblicksweise *Rühl*, JZ 2020, 809, 814 f.

¹⁸ S. dazu im vorliegenden Zusammenhang auch *Kotsoglou*, JZ 2014, 451 ff.; *Deichsel*, Digitalisierung der Streitbeilegung, 2021, 251 ff.

3.2 Die reale Welt des Justizsylogismus

Um diese Frage zu beantworten, ist für alle drei Elemente des Justizsylogismus' zu klären, wie sie tatsächlich entstehen.¹⁹

3.2.1 Der Obersatz

Woher kommt also der Obersatz? Idealerweise wird dieser gebildet aus dem Gesetz selbst bzw. durch dessen Auslegung und ggf. Fortbildung. Diese folgen den Regeln der Methodenlehre, die allerdings ihrerseits nicht deterministisch sind und sich algorithmisch bisher nicht abbilden lassen. Die Auslegung eines Gesetzes zur Anwendung in einem konkreten Einzelfall erfordert ein ständiges „Hin-und-her-Wandern des Blickes“ zwischen Rechtsnorm und Sachverhalt,²⁰ um all diejenigen Elemente der Rechtsnorm passend zu konkretisieren, die für den jeweiligen Sachverhalt relevant sind. Zugleich wird der Sachverhalt nach verallgemeinerungsfähigen Gesichtspunkten untersucht, die wiederum für eine Normkonkretisierung durch Auslegung und Fortbildung fruchtbar gemacht werden. Dieser Prozess erfolgt nicht rein mechanisch, sondern setzt eine Rückbindung an dahinterliegende Werte und Wertungen voraus, der notwendigerweise ein subjektives Element inneohnt. Gesetzesauslegung und -fortbildung sind auch bei Einhaltung der anerkannten Regeln der Methodenlehre und ihrer verfassungsrechtlichen Grenzen nie intersubjektiv zwingend, sondern folgen unvermeidlich auch persönlichen Präferenzen der jeweiligen Rechtsanwender.²¹ Diese Subjektivität ist auch institutionell im Grundgesetz verankert, nämlich in der richterlichen Unabhängigkeit nach Art. 97 Abs. 1 GG. Gäbe es nur eine objektiv richtige methodenkonforme Gesetzesauslegung, so wäre die sachliche richterliche Unabhängigkeit nicht erforderlich; die einzig richtige Auslegung könnte schlicht von oben verordnet werden.

Realiter erfolgt die „amtliche“ Auslegung und Fortbildung der Gesetze dagegen durch die hierzu kompetenziell berufenen Gerichte. Dies ist für das deutsche Recht der Bundesgerichtshof (§ 543 Abs. 2 S. 1 Nr. 2 ZPO), für das europäische Recht der Europäische Gerichtshof in Luxemburg (Art. 267 AEUV). Jedoch gibt es in beiden Fällen keine formelle Präjudizienbindung außerhalb des konkret entschiedenen Verfahrens. Die Auslegung und Fortbildung durch das zuständige Obergericht ist daher nur faktisch, nicht aber rechtlich für Untergegerichte bindend. Vielmehr bedeutet die richterliche Unabhängigkeit (Art. 97 Abs. 1 GG) gerade auch, dass Untergegerichte von der Rechtsauffassung der Obergerichte abweichen dürfen. Die Verfahrensordnungen sehen für diesen Fall vor, dass Rechtsmittel zuzulassen sind, um

¹⁹ S. zum Folgenden näher und m.w.N. *Riehm*, Abwägungsentscheidungen in der praktischen Rechtsanwendung, 2006, S. 16 ff.

²⁰ Dazu *Engisch*, Logische Studien zur Gesetzesanwendung, 3. Aufl. 1963, S. 15.

²¹ S. im vorliegenden Zusammenhang auch *Rollberg*, Algorithmen in der Justiz, 2020, S. 74; *Nink*, Justiz und Algorithmen, 2021, 28 ff.; *Dörr*, Rethinking: law 02.2021, 41, 42.

eine höchstrichterliche Klärung herbeizuführen (z. B. §§ 511 Abs. 4; 543 Abs. 2 S. 1 Nr. 2 ZPO). Auch im europäischen Recht ist eine solche Abweichung grundsätzlich zulässig; das abweichende Gericht hat allerdings zuvor eine Vorabentscheidung des europäischen Gerichtshofs einzuholen.²²

Die Bildung des Obersatzes ist mithin im tatsächlichen Rechtssystem nicht objektiv vorgegeben, sondern obliegt jeweils den im Einzelfall entscheidenden Gerichten. Die Einheitlichkeit der Rechtsanwendung ist nicht durch die Vorgabe zwingender Obersätze gesichert, sondern allein prozedural durch entsprechende Vorlage- bzw. Rechtsmittelzulassungspflichten.

3.2.2 *Der Untersatz*

Auch hinsichtlich des Untersatzes stellt sich die Sachlage in der Realität wesentlich komplexer dar als in der Idealwelt des Justizsyllogismus⁴. In dieser Idealwelt steht der Sachverhalt fest, weil er entweder zwischen den Parteien unstreitig ist oder aus unbestrittenen öffentlich zugänglichen Dokumenten nachgewiesen wurde. In der Realität ist der Sachverhalt dagegen zwischen den Parteien streitig, was eine Beweisaufnahme und Beweiswürdigung durch das Gericht nötig macht. Die Beweiswürdigung ist jedoch nach § 286 ZPO der subjektiven richterlichen Überzeugung zugewiesen, die ihrerseits wiederum nicht deterministisch ist. Ziel der Beweiswürdigung ist nicht die Ermittlung der objektiven Wahrheit, sondern lediglich die hinreichende (subjektive) Überzeugung des Gerichts von der Wahrheit einer Tatsachenbehauptung der beweispflichtigen Partei.²³ Das System der Beweiswürdigung akzeptiert daher die Möglichkeit, dass unterschiedliche Gerichte dieselben Beweismittel trotz Einhaltung des rechtlich vorgegebenen Beweisverfahrens unterschiedlich würdigen, ohne dass eine dieser Würdigungen im Rechtsmittelwege als falsch identifiziert würde.

Auch hier gibt es also nicht die eine, ex ante feststehende „objektiv richtige“ Beweiswürdigung, sondern eine subjektive, der konkret entscheidenden Person überantwortete Entscheidung über den zugrunde zu legenden Sachverhalt.

3.2.3 *Die Subsumtion*

Was schließlich die Subsumtion anbelangt, also die Zuordnung der einzelnen Sachverhaltselemente des Untersatzes zu den Tatbestandselementen des Obersatzes und die sich daraus ergebende rechtliche Schlussfolgerung, so stellt sich diese idealiter als simple logische Operation dar: Eine Buchung ist eine Buchung, und Verspätung ist Verspätung. Doch auch hier erweist sich die Realität als komplexer. Eine simple logische Operation liegt nur dann vor, wenn der Obersatz ausschließlich aus deskriptiven Tatbestandsmerkmalen besteht, denen die Sachverhaltselemente ohne jede rechtliche Wertungsoperation zugeordnet werden können.

²² Kotzur/Dienelt, in: Geiger/Khan/Kotzur u. a., EUV, AEUV, 7. Aufl. 2023, Art. 267 AEUV Rn. 38.

²³ Dazu ausf. Riehm, Abwägungsentscheidungen in der praktischen Rechtsanwendung, 2006, S. 77 ff.

Bei näherer Betrachtung zeigt sich indessen, dass es solche „rein deskriptiven“ Tatbestandsmerkmale kaum je gibt. Wann ist etwa eine Flugbuchung eine „bestätigte Buchung“ im Sinne der Fluggastrechte-Verordnung – was ist beispielsweise, wenn die Buchung mit Falschangaben erschlichen wurde (z. B. Buchung für „noch unbekannt“²⁴)? Worauf kommt es an, um die Verspätung eines Fluges festzustellen: auf das Aufsetzen der Maschine auf der Landebahn, das Andocken am Gate oder die Ankunft der Koffer auf dem Gepäckband?²⁵ Selbst vermeintlich klare und deskriptive Tatbestandsmerkmale haben Grenzfälle, deren Zuordnung nicht mehr selbstverständlich ist. Erst recht gilt das bei wertausfüllungsbedürftigen Tatbestandsmerkmalen, also solchen, bei denen die Zuordnung des Sachverhalts zwingend eine rechtliche Wertungsoperation erfordert. Im Beispiel sind das die „außergewöhnlichen Umstände“, die einen Entschädigungsanspruch ausschließen, wenn die Flugverspätung auf ihnen beruht. Zur Konkretisierung dieser „außergewöhnlichen Umstände“ enthält zwar Erwägungsgrund 14 der Fluggastrechte-Verordnung eine beispielhafte Aufzählung; diese ändert aber nichts an dem Erfordernis einer Abwägung aller Umstände in jedem Einzelfall.²⁶ Zudem führen auch außergewöhnliche Umstände nach der Rechtsprechung des EuGH nur dann zu einem Ausschluss des Ausgleichsanspruchs, wenn diese nicht durch „zumutbare Vorkehrungen“ hätten vermieden werden können. Auch die Prüfung, welche Aufwendungen „zumutbar“ in diesem Sinne sind, erfordert eine umfassende Einzelfallabwägung.²⁷

Auch diese Abwägungsentscheidungen im Rahmen der Subsumtion sind ein nicht-deterministischer Vorgang. Trotz geradezu ausufernder Kasuistik geschehen tagtäglich neue Sachverhalte, deren rechtliche Beurteilung im Hinblick auf ihre Einordnung als „außergewöhnliche Umstände“ solange ungewiss ist, bis sie höchststrichterlich entschieden sind – durch den BGH oder den EuGH, deren Entscheidung niemand zu 100% sicher vorhersagen kann, und die Untergerichte nicht über den entschiedenen Einzelfall hinaus rechtlich bindet, noch nicht einmal für Entschädigungsansprüche anderer Passagiere desselben Fluges.

3.3 Zwischenfazit

Nach alledem bleibt festzuhalten, dass deterministische Algorithmen nur dann Recht sprechen könnten, wenn der Obersatz strikt vorgegeben wäre, der zu subsumierende Sachverhalt feststünde, und für die Subsumtion keinerlei Wertungsoperationen erforderlich wären. Da jedoch unter den Bedingungen der realen Rechtsanwendung keine einzige dieser Voraussetzungen erfüllbar ist, scheidet eine „echte“ Rechtsanwendung durch deterministische Algorithmen bereits aus grundsätzlichen

²⁴ BGH NJW 2013, 598.

²⁵ BeckOGK BGB/*Steinrötter*, 1.11.2022, Fluggastrechte-VO Art. 6 Rn. 46 ff.

²⁶ BeckOGK BGB/*Steinrötter*, Art. 5 Rn. 39.

²⁷ S. zu deren Rechtsnatur *Riehm*, RW 2013, 1 ff.

rechtstheoretischen Erwägungen aus. Der Prozess der Rechtsanwendung ist auf jeder seiner Stufen durch nicht-deterministische Elemente geprägt, die eine Abbildung durch deterministische Algorithmen unmöglich machen.

Aber schließt das auch eine rechtliche Entscheidungsfindung durch Machine Learning-Algorithmen aus? Diese Algorithmen zeichnen sich ja gerade dadurch aus, dass sie keine *ex ante* feststehenden Wenn-dann-Strukturen benötigen, sondern aus (sehr) vielen Beispielen Muster erkennen und lernen. Studien haben gezeigt, dass solche Algorithmen auch Abwägungsmuster erkennen können – auch wenn die hierzu am häufigsten zitierten verschiedenen Studien zur Vorhersage von Entscheidungen des Europäischen Gerichtshofs für Menschenrechte (EGMR)²⁸ aus methodischen Gründen wertlos sind.²⁹ Bei einer entsprechend großen Menge an Trainingsdaten wäre es tatsächlich in nicht allzu ferner Zukunft denkbar, dass eine solche „Künstliche Intelligenz“ auch rechtliche Wertungsentscheidungen nach dem Vorbild vergangener Gerichtsentscheidungen treffen und rechtsmittelfest begründen kann. Wegen der grundsätzlichen Fähigkeit derartiger Systeme zu Ähnlichkeitschlüssen wäre dies auch nicht auf die Reproduktion vergangener Entscheidungen begrenzt, sondern könnte auch unbekannte Sachverhalte in Analogie zu bekannten Mustern entscheiden und dabei gewissermaßen Rechtsfortbildung betreiben.

4 Bedenken

Allerdings bestehen gegenüber dem Einsatz derartiger impliziter Algorithmen auf der Basis von Machine Learning-Technologien eine Reihe gravierender Bedenken, von denen nur die wichtigsten im Folgenden kurz ausgeführt werden sollen.³⁰ Diese Bedenken betreffen zum einen die technische Umsetzbarkeit einer derartigen „Justiz-KI“, zum anderen deren rechtsstaatliche Zulässigkeit.

²⁸ *O'Sullivan/Beel*, Predicting the Outcome of Judicial Decisions made by the European Court of Human Rights, 2019 (https://ceur-ws.org/Vol-2563/aics_26.pdf); *Şulea/Zampieri/Vela u. a.*, in: *Mitkov/Angelova*, Proceedings of the International Conference Recent Advances in Natural Language Processing, 2017, S. 716 ff.; *Aletras/Tsarapatsanis/PreoŃiu-Pietro u. a.*, PeerJ Computer Science 2 (2016), e93 ff.

²⁹ Die verwendete Software analysierte nicht die Verfahrensakten mit dem streitigen Parteivortrag, sondern lediglich den vom EGMR *ex post facto* erstellten Tatbestand, und orientierte sich nicht an dessen Inhalt, sondern an unspezifischen Konjunktionen, deren Verwendung im Tatbestand durch den EGMR das gefällte Urteil verriet, s. *Aletras/Tsarapatsanis/PreoŃiu-Pietro u. a.*, PeerJ Computer Science 2 (2016), e93 (13 ff.); krit. daher auch *Bull/Steffek*, ZKM 2018, 165, 167 f.

³⁰ S. auch die Überblicke bei *Rühl*, in: *Kaulartz/Braegelmann*, Rechtshandbuch Artificial Intelligence und Machine Learning, Kap. 14.1; *Huber/Giesecke*, in: *Ebers/Heinze/Krügel/Steinrötter*, Künstliche Intelligenz und Robotik, 2020, § 19 Rn. 23 ff.

4.1 Technische Umsetzbarkeit

4.1.1 Das Black box-Problem

Wie bereits eingangs ausgeführt, sind Algorithmen, die mithilfe von Machine Learning-Technologien entwickelt bzw. „trainiert“ wurden, für Menschen nicht mehr verständlich. Im Kern führen diese Algorithmen eine unvorstellbar große Zahl statistischer Operationen durch, die die Kapazität des menschlichen Verstandes um Größenordnungen überschreitet. Das menschliche juristische Denken, das auf Kausalitäten und logische Erklärungen ausgerichtet ist, ist mit diesen Verfahren schlicht nicht kompatibel. Anders als bei deterministischen Wenn-dann-Strukturen bilden solche Verfahren nicht mehr die Grundstruktur des Justizsyllogismus ab, sondern beruhen auf einem hochkomplexen Ähnlichkeitsvergleich unter Berücksichtigung unzähliger Parameter. Die Kausalität oder Bedeutung einzelner Parameter für das Endergebnis ist selbst für die Programmierer des Algorithmus oder für die „Trainer“ nicht mehr nachvollziehbar. Der eigentliche Entscheidungsweg kann daher nicht mehr in einer für Menschen verständlichen Weise dargestellt werden, weil er nicht über sprachlich darstellbare Kausalitätsbeziehungen läuft – das System ist für Menschen eine „Black box“.³¹

Zwar zeigen moderne generative KI-Algorithmen eindrucksvoll, dass eine „Künstliche Intelligenz“ in absehbarer Zeit ohne weiteres in der Lage sein wird, rechtmittelfeste Entscheidungsbegründungen zu verfassen.³² Es wäre daher technisch möglich, für praktisch jedes Entscheidungsergebnis eine rechtlich vertretbare Begründung automatisiert zu erstellen, die einer Überprüfung durch höhere Instanzen standhalten würde. Diese Begründung hätte allerdings mit der eigentlichen Entscheidungsfindung nichts zu tun; sie könnte sogar für ein rein zufällig ausgewürfeltes Ergebnis gegeben werden. Dies würde das Black box-Problem nicht lösen, sondern eher verschärfen, weil eine überzeugend formulierte, aber letztlich aus der Luft gegriffene Entscheidungsbegründung umso wirkungsvoller verschleiern würde, dass die eigentliche Entscheidung gerade nicht auf den in der Begründung dargestellten Erwägungen beruht.

Hiergegen lässt sich allenfalls anführen, dass auch ein menschlicher Richter für die Parteien letztlich eine Black box darstellt, und dass auch bei menschlichen Richtern die „Herstellung“ einer Entscheidung nicht zwingend mit der „Darstellung“ ihrer Begründung übereinstimmen muss. Ob die gegebene rechtliche Begründung tatsächlich den Gedankengang bei der Entscheidungsfindung widerspiegelt, oder *ex post* zu deren Rechtfertigung konstruiert wird, ist für die Parteien nicht ersichtlich

³¹ Dazu etwa *Rühl*, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning, Kap. 14.1 Rn. 19 ff. m.w.N.

³² S. etwa den GPT-Chatbot von open.ai, s. <https://openai.com/blog/chatgpt/>.

und rechtlich auch unerheblich.³³ Allerdings stellt sich in diesem Fall das menschliche Gericht immerhin mit seiner Unterschrift und damit mit der persönlichen Verantwortung hinter die Entscheidung und die hierfür gegebene Begründung. Diese Übernahme von Verantwortung ist ein zwischenmenschlicher kommunikativer Akt, der für die Akzeptanz einer gerichtlichen Entscheidung von wesentlicher Bedeutung sein dürfte und durch eine „Künstliche Intelligenz“ nicht abzubilden ist.

4.1.2 Trainingsdaten

Ein weiteres praktisches Problem ist das der Trainingsdaten: Ein Machine Learning-Algorithmus bedürfte einer extrem großen Menge von Trainingsdaten, die derzeit schlicht nicht verfügbar sind. Soll ein Algorithmus tatsächlich einen Fall auf der Grundlage streitiger Tatsachen- und Rechtsbehauptungen der Prozessparteien entscheiden, müsste er auf der Grundlage vollständiger Verfahrensakten trainiert werden. Es genügt also nicht, lediglich Urteile als Trainingsdaten zu verwenden, weil die darin enthaltenen Sachverhaltszusammenfassungen im Tatbestand bereits auf einer eigenständigen Würdigung durch das Gericht beruhen und in Kenntnis des Entscheidungsergebnisses erstellt worden sind. Es ist daher davon auszugehen, dass die Sachverhaltsdarstellung im Tatbestand bereits „zwischen den Zeilen“ das Ergebnis verrät, wie dies in der EGMR-Studie auch deutlich geworden ist.³⁴

Eine tatsächlich leistungsfähige juristische „Künstliche Intelligenz“ müsste vielmehr auf der Basis von Millionen vollständiger Verfahrensakten (von der Klageschrift bis zum Endurteil) trainiert werden, um die Arbeit eines Gerichtes zu übernehmen. Diese Daten sind erst seit kurzem und noch längst nicht flächendeckend digital verfügbar und müssten über viele Jahre aus den e-Akten-Systemen der Justiz extrahiert und gesammelt werden. Gegenüber diesem Vorgehen bestehen selbstverständlich auch noch gravierende datenschutzrechtliche Bedenken.

Die gegenwärtigen Versuche, Systeme allein auf der Basis der verfügbaren veröffentlichten Rechtsprechung zu trainieren, führen indessen zu nicht aussagefähigen Ergebnissen. In Deutschland werden derzeit – bedauerlicherweise – nur ca. ein bis zwei Prozent der tatsächlich ergehenden Gerichtsentscheidungen veröffentlicht.³⁵ Die Auswahl dieser Entscheidungen ist dabei keineswegs repräsentativ. Veröffentlicht werden vielmehr vorwiegend diejenigen Entscheidungen, die von den Parteien oder den Gerichten für „publikationswürdig“ gehalten werden, weil sie „besonders“ sind. Veröffentlicht werden mit anderen Worten nicht die alltäglichen „Brot-und-Butter-Entscheidungen“, in denen Standardfälle nach dem Vorbild der ständigen Rechtsprechung behandelt werden. Stattdessen wird gerade der verschwindend geringe Teil „Ausreißer“-Entscheidungen publiziert, in denen etwa ein

³³ Zu diesem Problem näher *Riehm*, Abwägungsentscheidungen in der praktischen Rechtsanwendung, 160 ff.

³⁴ Dazu oben bei und in Fn. 29.

³⁵ Dazu näher *Hamann*, JZ 2021, 656 ff.

unterinstanzliches Gericht von der höchstrichterlichen Rechtsprechung abweicht oder eine neue Rechtsfrage erstmalig entschieden wird.

Wollte man allerdings einer „künstlichen Intelligenz“ beibringen, wie deutsche Gerichte tatsächlich entscheiden, so müsste man ihr die 98-99 % unveröffentlichten Entscheidungen zeigen, die das Alltagsgeschäft der Justiz repräsentativ abbilden. Der geringe Prozentsatz veröffentlichter „besonderer“ Entscheidungen erweckt demgegenüber ein völlig falsches Bild von der tatsächlichen Rechtsprechung (sogeannter *publication bias*). Auf die wirklich relevanten 98-99 % besteht aber gegenwärtig kein digitaler Zugriff.

4.1.3 Sachverhaltsermittlung – Inputdaten

Problematisch bleibt ferner die Sachverhaltsermittlung. Zwar könnte eine „Künstliche Intelligenz“ die Tatsachenbehauptungen aus den Schriftsätzen der Parteien extrahieren und auswerten; eine umfassende Beweisaufnahme mit Zeugenvernehmungen (einschließlich Glaubwürdigkeitsbeurteilung) und Augenscheinnahmen im Ortstermin stellen jedoch noch einmal ganz andere Anforderungen an die Fähigkeiten entsprechender Algorithmen.³⁶ Jeder Algorithmus kann indessen nur so gute Ergebnisse erzielen, wie es die Qualität seiner Inputdaten zulässt – es gilt das Prinzip „*rubbish in, rubbish out*“. Hier dürfte die „Schnittstelle“ zwischen der realen Welt der Beweismittel und ihrer Auswertung durch Software noch für lange Zeit unüberwindbar sein.

4.1.4 Adverse Patterns

Nach dem gegenwärtigen Stand der Technik können Systeme, die auf Machine Learning-Technologien beruhen, verhältnismäßig leicht durch sogenannte „*adverse patterns*“ getäuscht werden. Darunter versteht man Muster, die für das menschliche Auge gänzlich unverdächtig und nicht erkennbar sind, die algorithmische Mustererkennung aber überlisten. Bekannt ist dies etwa von unscheinbaren Aufklebern, die eine automatische optische Verkehrszeichenerkennung dazu bringen, ein Tempolimitschild, das für einen Menschen ohne weiteres als ein solches erkennbar bleibt, als Stoppschild zu identifizieren.³⁷ Tierbilder können mit einem für Menschen unsichtbaren Bildrauschen überlagert werden, was eine Bilderkennungs-Software dazu bringt, eine andere Tierart in dem Bild zu erkennen. Für eine Justiz-KI wäre ein banales Beispiel für ein *adverse pattern* die Einfügung von weißem Text auf weißem Grund in einen PDF-Schriftsatz, der für Menschen nicht sichtbar wäre, aber entscheidende Schlüsselwörter für das KI-System enthalten könnte.

³⁶ Nink, ZdiW 2021, 1 (2); zweifelnd daher sogar *Quarch/Hähmle*, NJOZ 2020, 1281, 1285.

³⁷ S. etwa <https://www.heise.de/newsticker/meldung/Autonome-Autos-Forscher-fuehren-Bildererkennung-mit-manipulierten-Schildern-in-die-Irre-3974483.html>.

Der Einsatz solcher *patterns* könnte zwar wiederum technologisch erkannt werden, es würde allerdings potenziell zu einem ewigen Wettlauf zwischen Täuschenden und Getäuschten kommen. Gerade wenn bekannt ist, mit welchen Algorithmen die Justiz arbeitet – und etwas anderes dürfte schon aus rechtsstaatlichen Gründen nicht vorstellbar sein –, könnten Schriftsätze auf einer Parallelinstallation der Justizsoftware so lange getestet werden, bis das gewünschte Ergebnis erreicht wird.

4.1.5 Die Versteinerung der Rechtsprechung

Zuletzt wird gegen den Einsatz von Machine Learning-Algorithmen in der Justiz immer wieder zu Recht ins Feld geführt, dass diese zu einer „Versteinerung“ der Rechtsprechung führen würden.³⁸ Die Algorithmen würden anhand der Entscheidungen aus der Vergangenheit trainiert werden und könnten diese stets reproduzieren oder auf neue Sachverhalte entsprechend der gelernten Muster anwenden. Neue Interpretationen der bestehenden Gesetze aufgrund eines geänderten Regelungsumfelds oder geänderter gesellschaftlicher Wertvorstellungen wären auf dieser Grundlage nicht möglich.³⁹ Auch bestehende Vorurteile der Vergangenheit, die die bisherige Rechtsprechung und damit die Trainingsdaten geprägt haben mögen, würden auf diese Weise auf Dauer fortgeschrieben.

Zudem würde jede Gesetzesänderung die Systeme obsolet machen, weil die Trainingsdaten, die auf der Grundlage der alten Gesetzeslage ergangen sind, wertlos würden. Da zudem nicht genau identifiziert werden kann, welcher Faktor der Trainingsdaten die Ergebnisse des Algorithmus’ in welcher Form beeinflusst, können neue Gesetze auch nicht durch punktuelle Modifikation des Algorithmus’ eingearbeitet werden. Der Algorithmus ist vielmehr insgesamt bei einer (und sei es auch nur marginalen) Gesetzesänderung wertlos und muss vollständig neu trainiert werden.

4.2 Verfassungsrechtliche Grenzen

Unter rechtsstaatlichen bzw. verfassungsrechtlichen Gesichtspunkten bestehen eine ganze Reihe gravierender Bedenken gegen den Einsatz künstlicher Intelligenz zur automatisierten rechtlichen Entscheidungsfindung bei Gericht.⁴⁰

³⁸ Etwa *Rühl*, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning, Kap. 14.1 Rn. 15 ff.

³⁹ S. auch *Heil*, IT-Anwendung im Zivilprozess, 2020, S. 76 ff.

⁴⁰ S. auch *Rühl*, in: Kaulartz/Braegelmann, Rechtshandbuch Artificial Intelligence und Machine Learning, Kap. 14.1 Rn. 22 ff. mit weiteren rechtlichen Bedenken; *Bernzen*, RD 2023, 133 ff.

4.2.1 Recht auf den gesetzlichen Richter und Richtervorbehalt

Ein zentraler Einwand betrifft zunächst das grundrechtsgleiche Recht auf den gesetzlichen Richter (Art. 101 Abs. 1 S. 2 GG) und die Übertragung der rechtsprechenden Gewalt auf „die Richter“ (Art. 92 GG). Es entspricht ganz herrschender Auffassung im Verfassungsrecht, dass „Richter“ im Sinne dieser Vorschriften ausschließlich Menschen sein können,⁴¹ die richterliche Unabhängigkeit nach Art. 97 Abs. 1 GG genießen. Soweit also nach dem geltenden (Prozess-)Recht eine rechtsprechende Tätigkeit vorgesehen ist, kann diese Aufgabe nur durch einen Menschen wahrgenommen werden.

Zwar unterliegt nicht jede rechtliche Entscheidung einem derartigen Richtervorbehalt, wie bereits die zahlreichen Aufgaben belegen, die nach § 3 RPfG den Rechtspflegern übertragen sind, welche keine richterliche Unabhängigkeit genießen⁴² und daher auch nicht vom Recht auf den gesetzlichen Richter umfasst sind.⁴³ Das – den Rechtspflegern übertragene (§ 3 Nr. 3 lit. a) i.V.m. § 20 Abs. 1 Nr. 1 RPfG) – Mahnverfahren kann hier auch durchaus zu rechtskräftigen Sachentscheidungen in Gestalt von Vollstreckungsbescheiden gem. § 699 ZPO führen. Gleichwohl besteht ein zentraler Unterschied zur eigentlichen Spruchrichtertätigkeit, die verfassungsrechtlich gem. Art. 101 Abs. 1 S. 2 und Art. 97 Abs. 1 GG privilegiert ist, darin, dass Mahn- und Vollstreckungsbescheide von Rechtspflegern ohne inhaltliche Sachprüfung erlassen werden. Eine tatsächliche inhaltliche Richtertätigkeit lässt sich nicht auf vergleichbare Weise dem Richtervorbehalt (und damit den Art. 97 Abs. 1, 101 Abs. 1 S. 2 GG) entziehen und einem Softwaresystem übertragen.

Das Recht auf den gesetzlichen Richter gilt auch nicht nur für die letzte Instanz, sondern in jeder Instanz der gerichtlichen Entscheidung. Sobald also nach der ZPO eine rechtliche Streitentscheidung aufgrund inhaltlicher Sachprüfung gefordert ist, muss diese aus verfassungsrechtlichen Gründen von einem Menschen getroffen werden. Ob de lege ferenda in der ZPO eine Art automatisch generierter „Vorbescheid“ vorgesehen werden kann, der keine richterliche Entscheidung darstellt, erscheint angesichts des klaren Wortlauts von Art. 92 GG äußerst zweifelhaft: Zu der von dieser Vorschrift den (menschlichen) Richtern exklusiv anvertrauten rechtsprechenden Tätigkeit gehört nach der Rechtsprechung des BVerfG die Entscheidung

⁴¹ Enders, JA 2018, 721, 723; *Länderarbeitsgruppe*, Legal Tech: Herausforderungen für die Justiz, 2019 (https://www.schleswig-holstein.de/DE/landesregierung/ministerien-behoerden/II/Minister/Justizministerkonferenz/Downloads/190605_beschluesse/TOPI_11_Abschlussbericht.pdf?__blob=publicationFile&v=1), S. 60; *Winkelmann*, LTZ 2022, 163 (165); *Nink*, Justiz und Algorithmen, 261 ff.; offener *Quarcb/Hähnle*, NJOZ 2020, 1281, 1283.

⁴² Dürig/Herzog/Scholz/*Hilgruber*, Grundgesetz, März 2022, Art. 97 Rn. 20 sowie Dürig/Herzog/Scholz/*Hilgruber*, Art. 92 Rn. 70.

⁴³ Dürig/Herzog/Scholz/*Jachmann-Michel*, Art. 101 Rn. 35.

bürgerlicher Rechtsstreitigkeiten vermögensrechtlicher Art⁴⁴ (nach dem Vorstehenden zu ergänzen: aufgrund inhaltlicher Sachprüfung). Schon hier stößt der Einsatz von KI als Richterersatz an verfassungsrechtliche Grenzen, die allenfalls durch eine Verfassungsänderung überwunden werden können.⁴⁵

4.2.2 Rechtsstaatsprinzip

Auch aus dem Rechtsstaatsprinzip – das einer Verfassungsänderung nach Art. 79 Abs. 3 GG entzogen ist – ergeben sich Grenzen für den Einsatz von KI zur rechtlichen Streitentscheidung. Aus diesem folgt zum einen ein Begründungserfordernis für rechtliche Entscheidungen, das richtigerweise nur so verstanden werden kann, dass die tatsächlich tragenden Erwägungen für die Entscheidung wiedergegeben werden müssen. Beim gegenwärtigen Stand von Machine Learning-Algorithmen kann aber gerade keine für Menschen verständliche Begründung im Sinne einer kausalen Ableitung der Entscheidung aus Rechtsnormen und den einzelnen entscheidungserheblichen Sachverhaltselementen gegeben werden. Gerade diese Begründung ist jedoch von erheblicher Bedeutung für die prozedurale Gerechtigkeitsgewähr: Wo schon das Entscheidungsergebnis wegen der unvermeidlichen Auswirkungen subjektiver Wertungen nur eingeschränkt kontrollierbar ist, muss wenigstens die *Herleitung* der Entscheidung nachvollziehbar und durch höhere Instanzen kontrollierbar sein, um die Entscheidung an Recht und Gesetz rückzubinden. Eine nur *ex post* konstruierte, den tatsächlichen Entscheidungsvorgang nicht wiedergebende Begründung genügt diesem rechtsstaatlichen Erfordernis gerade nicht.⁴⁶ Hier wären neuartige, derzeit noch nicht existierende Formen von „*explainable AI*“ nötig, um tatsächlich in menschenverständlicher Weise aufzudecken, welche Erwägungen für die Entscheidung tragend waren.

4.2.3 Richterliche Unabhängigkeit (Art. 97 Abs. 1 GG)

Ein entscheidender Gesichtspunkt gegen die Zulässigkeit des Einsatzes von Algorithmen zur rechtlichen Streitentscheidung folgt schließlich aus dem Gebot der richterlichen Unabhängigkeit (Art. 97 Abs. 1 GG):⁴⁷ Soweit die Entscheidung bürgerlicher Rechtsstreitigkeiten Richtern übertragen ist (s. Art. 92 GG), müssen diese Richter sachliche Unabhängigkeit genießen. Jeder Richter muss also befugt sein, selbst und in eigener inhaltlicher Verantwortung seine Entscheidung zu treffen. Damit ist es nicht vereinbar, wenn Richter gezwungen sind, eine bestimmte Software einzusetzen und deren Ergebnisse als eigene Entscheidungen zu erlassen.

⁴⁴ BVerfG 14, 56, 66; Dürig/Herzog/Scholz/Hillgruber, Art. 92 Rn. 36.

⁴⁵ Bernzen, RD 2023, 133 Rn. 10 ff.

⁴⁶ BVerfG NJW 2001, 2161, 2162 sub B.I.2.c.cc.

⁴⁷ Dazu eingehend Kable, LTZ 2022, 170 ff.

Wie oben ausgeführt, ist der einzelne Richter gerade nicht an die Rechtsprechung übergeordneter Instanzen gebunden und kann auch nicht durch deren softwaremäßige Implementierung mittelbar dazu gezwungen werden, diese zu übernehmen.⁴⁸

4.2.4 Gewaltenteilung

Das führt zu einem weiteren Problem, das in der bisherigen Diskussion noch kaum erörtert wird: Welche staatliche Gewalt darf eigentlich legitimerweise über den Einsatz und vor allem die Konfiguration von rechtlicher Entscheidungssoftware bestimmen? In der Praxis dürften es derzeit vorrangig die Justizministerien der Länder sein, die entsprechende Software zentral beschaffen oder in Auftrag geben. Diese gehört allerdings der Exekutive an und ist nicht befugt, die inhaltliche Entscheidungstätigkeit von Gerichten zu bestimmen. Keinesfalls dürfen Ministerialbeamte darüber entscheiden, welche Algorithmen in welcher Konfiguration anstelle der unabhängigen Gerichte rechtliche Entscheidungen treffen. Aus dem gleichen Grund darf eine solche Entscheidung auch nicht vom Gesetzgeber getroffen werden.

Die Auswahl der entsprechenden Software müsste also der Judikative übertragen werden, die freilich keine zentralen passenden Organe für derartige Entscheidungen hat. Allenfalls die Gerichtspräsidien wären derartige Gremien – freilich könnte die Entscheidung dann nur je Gericht getroffen werden, weder bundes- noch länder einheitlich. Selbst diese wäre indessen noch unzulässig, weil sie, wie vorstehend gezeigt, in die richterliche Unabhängigkeit der einzelnen Spruchkörper eingreifen würde. Letztlich kann daher nur ein Spruchkörper selbst die Software auswählen bzw. zumindest deren konkrete inhaltliche Konfiguration bestimmen, also über das „Ob“ und das „Wie“ des Softwareeinsatzes entscheiden.

5 Fazit und Ausblick

Das Fazit nach den vorstehenden Ausführungen ist ernüchternd: Deterministische Algorithmen sind schon strukturell nicht in der Lage, rechtliche Entscheidungen streitiger Sachverhalte selbständig anstelle von menschlichen Richtern zu treffen. Auch Systeme auf der Basis von Machine Learning können rechtsstaatlichen Maßstäben an richterliche Entscheidungen nicht genügen; sie stoßen an technische wie an verfassungsrechtliche Grenzen.

Der Einsatz rechtlicher Entscheidungssysteme in der Justiz ist damit aber nicht von vornherein ausgeschlossen. Allerdings dürfen derartige Systeme – das dürfte inzwischen weitgehend Konsens sein⁴⁹ – nur in unterstützender Weise eingesetzt

⁴⁸ Bernzen, RDi 2023, 133 Rn. 24.

⁴⁹ S. etwa Engel, JZ 2014, 1096 ff.; Quarch/Hähnle, NJOZ 2020, 1281 ff.; Rollberg, Algorithmen in der Justiz, 125 ff.; Kahl, LTZ 2022, 170, 172 f.; Bernzen, RDi 2023, 133 Rn 24 ff.

werden. Die inhaltliche Verantwortung für die getroffenen Entscheidungen muss – bei aller Vorbereitung durch Algorithmen – beim menschlichen Gericht verbleiben.

Selbst ein bloß unterstützender Einsatz von Algorithmen zur Vorbereitung rechtlicher Entscheidungen ist indessen nicht grenzenlos zulässig. Auch insoweit kann die richterliche Unabhängigkeit tangiert sein. Selbst wenn Richter frei sind, die Ergebnisse der unterstützenden Software zu übernehmen oder davon abzuweichen, ist es eine Frage der konkreten Ausgestaltung, ob damit ein Eingriff in die richterliche Unabhängigkeit verbunden ist. Bei der Beurteilung der Ausgestaltung ist insbesondere darauf zu achten, wie sich eine Abweichung des Richters von dem „Angebot“ der Unterstützungssoftware konkret darstellt:⁵⁰ Erfordert die Abweichung einen erheblichen manuellen Mehraufwand, und wird dieser bei der richterlichen Kapazitätsberechnung nicht berücksichtigt, so kann das dazu führen, dass ein Richter mittelbar gezwungen wird, das automatische Ergebnis der Unterstützungssoftware in den meisten Fällen zu übernehmen, um sein Arbeitspensum überhaupt in vertretbarer Zeit bewältigen zu können. Der dadurch mittelbar entstehende Übernahmedruck kann gegen die sachliche richterliche Unabhängigkeit verstoßen.⁵¹ Wird jedoch umgekehrt die Kapazitätsberechnung darauf ausgerichtet, dass die Richter gleichwohl jeden Fall manuell und ohne inhaltliche Unterstützungssoftware bearbeiten können, verfliegen die mit der Softwareeinführung verbundenen Hoffnungen auf eine Entlastung der Gerichte. Eine Lösung kann hier wohl nur darin liegen, dass die Unterstützungssoftware für jeden konkreten Richter personalisierbar ist, um dessen persönliche Rechtsauffassung abzubilden. Im Extrembeispiel: Wenn ein Richter am Landgericht sich entschließt, aus persönlicher rechtlicher Überzeugung eine bestimmte Rechtsfrage in Einzelrichtersachen stets entgegen der Rechtsprechung des BGH zu entscheiden, dann müsste die Unterstützungssoftware auch in der Lage sein, ihm entsprechende Entscheidungsvorschläge zu unterbreiten, die der höchstrichterlichen Rechtsprechung widersprechen.

Mein Plädoyer lautet daher, dass rechtliche Entscheidungen durch Algorithmen besser außerhalb von Gerichten stattfinden sollten, wo die vorstehend geschilderten rechtsstaatlichen und sonstigen verfassungsrechtlichen Anforderungen nicht gelten. Teils erfolgt das bereits heute, nämlich im Rahmen der Streitschlichtungs-Algorithmen von PayPal, eBay, Amazon Marketplace & Co.⁵² Diese Algorithmen erledigen eine sehr große Zahl von Streitfällen – die Rede ist von mehreren Millionen pro

⁵⁰ Zum Problem des sog. „Automation Bias“ s. auch *Nink*, Justiz und Algorithmen, 295 ff.; *Bernzen*, RDJ 2023, 133 Rn. 26 ff.

⁵¹ *Enders*, JA 2018, 721, 723; *Starosta*, DÖV 2020, 216 ff.; *Rühl*, in: Kaulartz/Braegelmann, Rechts- handbuch Artificial Intelligence und Machine Learning, Kap. 14.1 Rn. 28 f.; zum Verhältnis zwischen Arbeitspensum und richterlicher Unabhängigkeit s. BGH (Dienstgericht des Bundes) NJW 2018, 158; NJW 2022, 3320; BVerfG NJW 2021, 3717.

⁵² S. hierzu auch *Fries*, NJW 2016, 2860 ff.; *Quarab/Hähne*, NJOZ 2020, 1281, 1282; *Rühl*, JZ 2020, 809, 811.

Jahr, also mehr als die gesamte deutsche Ziviljustiz – in hochgradig automatisierter Weise. Die entsprechenden Anbieter haben aufgrund ihres Plattformcharakters bereits Zugriff auf viele streitrelevante Daten vom Bestellverlauf über Zahlungsdaten bis zur Sendungsverfolgung und können daher ihre Entscheidung auf relativ gesicherter Sachverhaltsgrundlage treffen. Zudem sind die einschlägigen Fälle rechtlich typischerweise einfach gelagert. Dem Vernehmen nach genügt für einen erheblichen Teil der Fälle im Versandhandel die einfache Regel „Geld und Ware dürfen niemals in den gleichen Händen sein“. Die Entscheidungen dieser Algorithmen sind rechtlich indessen nicht bindend, sondern unterliegen gegebenenfalls einer Korrektur durch die staatlichen Gerichte.⁵³

Ein anderes mögliches Einsatzgebiet streitentscheidender Algorithmen ist die Verbraucherschlichtung nach dem VSBG.⁵⁴ Auch hier werden nach gegenwärtigem Stand keine rechtlich bindenden Entscheidungen getroffen, sondern gem. § 5 Abs. 2 VSBG lediglich (für Verbraucher) unverbindliche Schlichtungsvorschläge unterbreitet, deren Annahme dem freien Willen der Streitparteien unterliegt.⁵⁵ Diese Schlichtungsvorschläge unterliegen gemäß § 19 Abs. 1 S. 2 VSBG auch keiner strengen Rechtsbindung.⁵⁶ Ein Einsatz in diesem Feld wäre zudem mit dem Vorteil verbunden, dass diese Schlichtungsstellen typischerweise branchenspezifisch ausgerichtet sind und daher in der Regel nur wenige Typen von Fällen bearbeiten müssen. Zudem sind die Schlichtungsstellen nicht an eine gesetzliche Verfahrensordnung gebunden, sondern geben sich ihre eigene Verfahrensordnung jeweils privatautonom selbst (§ 5 Abs. 1 VSBG). Dies wirkt wie ein ideales Testfeld für Algorithmen zur rechtlichen Entscheidung, die dementsprechend auch zunächst an einzelnen Falltypen entwickelt werden können. Freilich steht auch hier gegenwärtig noch die Regelung des § 6 Abs. 1 und 2 VSBG entgegen, wonach Streitmittler nur eine natürliche Person sein kann, die die Befähigung zum Richteramt besitzt oder zertifizierter Mediator ist. Diese Regelung ist allerdings nicht verfassungsrechtlich vorgegeben, sondern könnte durch den einfachen Gesetzgeber geändert werden. Freilich wäre hierfür zunächst auf europäischer Ebene eine Änderung von Art. 6 Abs. 1 der ADR-Richtlinie 2013/11/EU erforderlich, die ebenfalls nur natürliche Personen als Streitmittler vorsieht.

Der ebenfalls gelegentlich zu hörende Vorschlag, ein automatisiertes Entscheidungsverfahren innerhalb der Justiz als „automatisiertes Bagatellverfahren“ im Rahmen einer „nullten Instanz“ vorzusehen,⁵⁷ dürfte allerdings de lege lata an Art. 92 GG scheitern, denn selbst die Tätigkeit einer „nullten Instanz“ wäre inhalt-

⁵³ BGH NJW 2018, 537; MMR 2018 156 (jeweils zu PayPal).

⁵⁴ Hierzu auch *Deichsel*, VuR 2020, 283 ff.; *Deichsel*, Digitalisierung der Streitbeilegung, 311 ff.

⁵⁵ Freilich können sich die Unternehmen vorab den Schlichtungsvorschlägen unterwerfen.

⁵⁶ *Riehm*, JZ 2016, 866 ff.

⁵⁷ S. etwa *Fries*, RW 2018, 414, 422 ff.: „Automatische Vor-Urteile“.

lich rechtsprechende Tätigkeit und daher (menschlichen) Richtern vorbehalten. Insofern wäre zumindest eine Änderung des Grundgesetzes erforderlich; diese dürfte allerdings wohl nicht an Art. 79 Abs. 3 GG scheitern.⁵⁸

⁵⁸ Ebenso *Bernzen*, RDt 2023, 133 Rn. 10 ff.