



Künstliche Intelligenz, Bias und Versicherungen – Eine technische und rechtliche Analyse

Petra Pohlmann · Gottfried Vossen · Jan Everding · Johanna Scheiper

Angenommen: 4. Juli 2022 / Online publiziert: 2. August 2022
© Der/die Autor(en) 2022

Zusammenfassung Künstliche Intelligenz (KI) kann in der datengetriebenen Versicherungswirtschaft erhebliche Effizienzgewinne mit sich bringen. Das gilt auch für den Einsatz von KI im Verhältnis zu potenziellen Kunden, Versicherungsnehmern und Begünstigten von Versicherungsleistungen. KI trägt jedoch das Risiko eines Bias, eines systematischen Prognosefehlers, und damit der Diskriminierung, in sich. Dieser interdisziplinäre Beitrag untersucht aus Sicht der Informatik die Gefahr eines Bias und ihre Ursachen. Darauf aufbauend wird aus juristischer Sicht erörtert, inwieweit bei Bias die Diskriminierungsverbote eingreifen und welche Rechtsfolgen sie haben. Eine anschließende Analyse der technischen Vermeidbarkeit von Bias führt zu der juristischen Frage, wie das Versicherungsaufsichtsrecht sicherstellt, dass Bias soweit wie möglich verhindert wird. Abschließend wird die Bedeutung des EU-Entwurfs einer KI-Verordnung für die Vermeidung von Bias beleuchtet. Wie man mit der in diesem Beitrag angenommenen Unvermeidbarkeit von Bias informationstechnologisch und versicherungsrechtlich umzugehen hat, ist eine der wichtigen, in weiterer interdisziplinärer Forschung zu klärenden Fragen.

Petra Pohlmann (✉) · Johanna Scheiper
Forschungsstelle für Versicherungswesen, Universität Münster,
Universitätsstraße 14–16, 48143 Münster, Deutschland
E-Mail: petra.pohlmann@uni-muenster.de

Gottfried Vossen · Jan Everding
Institut für Wirtschaftsinformatik, Universität Münster, Leonardo-Campus 3, 48149 Münster,
Deutschland

Artificial intelligence, bias and insurance – A technical and legal analysis

Abstract Artificial intelligence (AI) can bring significant efficiencies to the data-driven insurance industry. This is also true for the use of AI in relation to potential customers, policyholders, and beneficiaries. However, AI carries the risk of bias, a systematic forecasting error, and thus of discrimination. This interdisciplinary article examines the risk of bias and its causes from the perspective of computer science. Based on this, it discusses from a legal point of view the extent to which non-discrimination law applies in the case of bias and what the legal consequences are. The subsequent analysis of whether and how data scientists can technically eliminate bias leads to the legal question of how insurance supervisory law ensures that bias is avoided as far as possible. Finally, we examine the implications of the EU draft of an AI regulation for the avoidance of bias. How to cope with the assumed inevitability of bias is one of the important questions that future interdisciplinary research of information technology and insurance law has to solve.

1 Einleitung

Künstliche Intelligenz (KI) ist ein Begriff, der aus der Diskussion um die Zukunft der Versicherungsbranche nicht mehr wegzudenken ist. Unter Künstlicher Intelligenz werden hier selbstlernende Algorithmen zusammengefasst, die mithilfe von Daten Lösungen für wiederkehrende Entscheidungsprobleme finden können.¹ KI wird schon jetzt in verschiedenen Bereichen eingesetzt und hat das Potenzial, die Arbeit der Versicherungsbranche grundlegend zu verändern. Die Anwendungsmöglichkeiten reichen dabei von Produktentwicklung über Underwriting, Vertrieb und Bestandsmanagement bis zur Schadenmeldung, -bewertung und -regulierung.

In diesem Beitrag stehen KI-Anwendungen im Fokus, die im Verhältnis zu potenziellen Kunden, zu Versicherungsnehmern und zu Begünstigten von Versicherungsleistungen eingesetzt werden.² Im Vertrieb können intelligente Systeme entweder den Vermittler bei seiner Arbeit unterstützen oder sogar selbst die Beratung übernehmen. Des Weiteren kann KI in Verbindung mit Big Data auch dazu genutzt werden, den Bedarf des Kunden besser zu analysieren und dem Kunden bessere oder zusätzliche Produkte anzubieten. KI ermöglicht insbesondere eine laufende Beobachtung bestimmter versicherter Risiken und eine damit verbundene flexible, an das Risiko angepasste Gestaltung des Versicherungsschutzes (pay as you live). Hinzu kommt der Einsatz in der Schadenbearbeitung. So ist es schon heute machbar, dass KI die Schadenmeldung des Kunden „liest“, bewertet und zum Beispiel weitere Informa-

¹ *Buxmann/Schmidt*, in *Buxmann/Schmidt* (Hrsg.), *Künstliche Intelligenz* (2019, S. 3, 6f.); *Kreutzer/Sirrenberg*, *Understanding Artificial Intelligence* (2020, S. 2ff.) – Das Manuskript wurde im August 2021 eingereicht. Neuere Literatur wurde vereinzelt berücksichtigt.

² Anwendungsbeispiele auch bei *Bartel/Khuat/Wehling-Mauntel*, in *Chibanguza/Kuß/Steege* (Hrsg.), *Künstliche Intelligenz, Recht und Praxis automatisierter und autonomer Systeme* (2022, S. 1049–1064, Rn. 3–16).

tionen vom Kunden anfordert oder den Schaden sofort reguliert. Insbesondere bei Großereignissen, sogenannten Kumulereignissen, wie den zunehmenden Unwetter-Ereignissen, könnten so die Schadenabteilungen der Versicherer entlastet und den Kunden Wartezeiten erspart werden.

KI ist dem Menschen in vielen Punkten überlegen. Sie kann mehr Informationen verarbeiten, und dies schneller, rund um die Uhr und immer präzise. Eine KI vergisst nichts, hat keinen schlechten Tag, lässt sich nicht ablenken, ist nie frustriert oder müde. Sie ist ferner modular ergänzbar, und die Quanteninformatik eröffnet aktuell ganz neue Perspektiven hinsichtlich der Leistungsfähigkeit von KI. Kurz gesagt – man verspricht sich, dass eine KI effizienter und effektiver arbeitet als der Mensch.³

KI ist aber auch nicht perfekt. So zeigen KI-Systeme durchaus Entscheidungs- bzw. Klassifikationsergebnisse, die diskriminieren.⁴ So bot bei einem in England durchgeführten Versuch im Jahr 2018 eine KFZ-Versicherungsvergleichsseite Männern mit vollkommen identischen Risikoprofilen höhere Prämien an, wenn sie keinen klassischen englischen Namen trugen, sondern etwa „Muhammad Khan“ hießen.⁵ Es wird zwar kein Unternehmen je ganz vermeiden können, dass einzelne Mitarbeiter ebenfalls voreingenommene Entscheidungen treffen. Im Unterschied zu diesen Einzelnen ist eine voreingenommene KI jedoch skalierbar. Wird eine solche KI zum Entscheider über z. B. die Höhe der Prämie, den Abschluss eines Versicherungsvertrags oder die Auszahlung der Leistung, so ist jede Entscheidung gegenüber den Kunden „belastet“; es können ganze Gruppen von notwendiger finanzieller Teilhabe ausgeschlossen werden.⁶ Und wird ein solches Verhalten einer KI aufgedeckt, leidet nicht nur der Ruf des Versicherers, sondern es drohen auch rechtliche Konsequenzen.

Der folgende Beitrag behandelt das Thema der Voreingenommenheit von KI, den sog. Bias, aus informationstechnologischer und juristischer Sicht. Wir klären zunächst die technischen Grundlagen und damit die Frage, wie es passieren kann, dass eine KI mit einem Bias belastete Entscheidungen trifft (Abschn. 2.). Im Anschluss untersuchen wir, wann die Diskriminierung von Versicherungsnehmern durch Versicherungsunternehmen verboten ist und welche zivilrechtlichen Folgen sich an einen Verstoß gegen das jeweilige Verbot für den Versicherer knüpfen (Abschn. 3.).⁷ Dann gehen wir auf die technischen Möglichkeiten ein, einem Bias und damit einer verbotenen Diskriminierung entgegenzuwirken (Abschn. 4.). Es folgen Überlegungen dazu, ob und inwieweit das Versicherungsaufsichtsrecht gewährleistet, dass Versicherer Versicherungsnehmer durch Ergreifen dieser Möglichkeiten vor Bias schützen (Abschn. 5.), sowie ein erster Blick auf den Kommissionsentwurf des Artificial Intelligence Act und seiner Bedeutung für die Verhinderung von Diskriminierung

³ Zum Vergleich künstlicher und menschlicher Intelligenz *Wilczek*, in Brockman (Hrsg.), *Possible Minds* (2019, S. 64, 70 ff.).

⁴ *Salimi/Howel/Suciu*, *ACM SIGMOD Record* 49 (2020), Heft 1, 34, 35.

⁵ *Kretchmer*, *Insurers 'risk breaking racism laws'*, 09.02.2018, abrufbar unter: <https://www.bbc.com/news/business-43011882> [Stand: 26.07.2021].

⁶ Zu diesem, hier nicht näher verfolgten Aspekt *EIOPA* (2021), *Artificial intelligence governance principles: towards ethical and trustworthy artificial intelligence in the European insurance sector*, 17.06.2021, S. 24 f.

⁷ Andere mögliche Anspruchsgegner (Hersteller, Verkäufer, Datenlieferanten u. a.) bleiben unberücksichtigt.

(Abschn. 6.). Im Fazit (Abschn. 7.) fassen wir die zentralen Ergebnisse zusammen und geben einen Ausblick auf zukünftig wichtige Forschungsfragen.

2 Technische Grundlagen

Das Feld der KI im oben genannten Sinne spannt sich in zwei Bereiche auf: Die *starke KI* befasst sich mit der künstlichen Entwicklung von menschlichem Verhalten und Emotionen. Der technologische Stand von dieser Art KI ist jedoch noch fern von einer Marktreife, sodass er vor allem im akademischen Bereich betrachtet wird.⁸

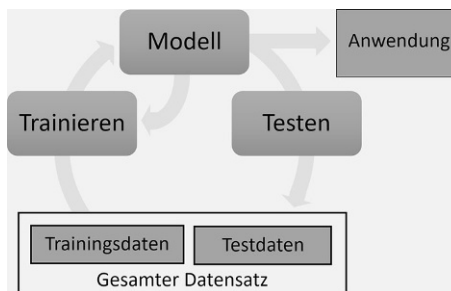
Anders sieht es hingegen mit dem Bereich der *schwachen KI* aus, welche sich mit der Lösung genau definierter Probleme beschäftigt und hierdurch immer häufiger im betrieblichen Kontext zur Unterstützung oder Übernahme von wiederkehrender Tätigkeiten eingesetzt wird.⁹ Vor diesem Hintergrund wird im Nachfolgenden lediglich diese Art von KI berücksichtigt.

2.1 Der Prozess des Lernens

Das Lernen einer Künstlichen Intelligenz basiert auf größeren Datenmengen, welche üblicherweise in Trainingsdaten und Testdaten, etwa im Verhältnis 80:20 oder 70:30, unterteilt werden.¹⁰ Während die Trainingsdaten zur Kalibrierung des Algorithmus bzw. zur Erstellung eines Modells genutzt werden, dienen die Testdaten zur Evaluierung der gelernten Muster. Zur Bewertung der algorithmischen Performance wird eine Zielfunktion eingesetzt, welche in der Regel einen möglichst maximalen Nutzen oder minimale Kosten erreichen soll. Eine Abstraktion dieses Lernprozesses wird in der Abb. 1 dargestellt.

Mit komplexen mathematischen und statistischen Verfahren, welche auch als Verfahren des maschinellen Lernens bekannt sind, wird im Rahmen des Trainings nach Datenmustern gesucht. Die Datenmuster, die in Kombination am besten zur Problemlösung geeignet sind, werden in ein Modell überführt und können mithilfe der

Abb. 1 Verwendung von Trainingsdaten und Testdaten zum Lernen einer KI; Quelle: Eigene Darstellung



⁸ Ashri, The AI-Powered Workplace: How Artificial Intelligence, Data, and Messaging Platforms Are Defining the Future of Work (2020, S. 15); Zimmer/Mitzner, BI-Spektrum (2019), Heft 5, 36, 37.

⁹ Gausling, in Graf Ballestrem/Bär/Gausling/Hack/von Oelffen (Hrsg.), Künstliche Intelligenz (2020, S. 11, 13).

¹⁰ Fader/Hardie, Journal of Interactive Marketing 23 (2009), 61 ff.

Testdaten evaluiert werden. Dabei können die Algorithmen auch sehr komplexe Probleme lösen und trainiertes (oder gelerntes) Wissen auf sehr große Datenbestände anwenden.¹¹ Liefert das Modell zufriedenstellende Ergebnisse, kann es in einer Produktivumgebung zur Behandlung unbekannter Datenbestände angewendet werden.

Bevor die Verfahren des maschinellen Lernens jedoch greifen können, müssen die Faktoren bestimmt werden, die im jeweiligen Anwendungsfall relevant sind. Die vorliegenden Merkmale werden untersucht und unter Annahme einer vertretbaren Irrtumswahrscheinlichkeit wird entschieden, ob die Merkmale im Rahmen des Lernens zur Verbesserung des Algorithmus beitragen. Während irrelevante Merkmale keine Anwendung im Rahmen des Lernprozesses finden, werden die relevanten Merkmale zum Trainieren der KI herangezogen, da davon ausgegangen wird, dass sie zur Beantwortung des Problems beitragen können.

Während dieses Vorgangs werden Datenmuster im Trainingsdatensatz gesucht, welche möglichst fehlerfrei auf den Testdatensatz übertragbar sein sollen. Werden zum Lernen einer KI zu wenig signifikante Einflussvariablen verwendet, führt dies dazu, dass nicht alle relevanten Faktoren abgebildet werden und die Fehlerhäufigkeit relativ hoch ist (sog. Underfitting). Die zusätzliche Aufnahme von weiteren bedeutsamen Merkmalen führt zu einer Erhöhung der Modellkomplexität, welche in der Regel zunächst auch eine allgemeine Verbesserung der Modellgüte mit sich zieht. Lernt eine KI hingegen zu viele Merkmale, hat dies zur Folge, dass das KI-Wissen nicht mehr auf andere Datenbestände übertragen werden kann, ohne zeitgleich eine Zunahme der Prognosefehler zu verursachen (sog. Overfitting).¹² Diese systematischen KI-Prognosefehler (Abweichungen zwischen dem wahren Wert und dem KI-Schätzwert) heißen im Kontext des Lernprozesses *Bias*. Der Begriff des Bias wird somit zur Beschreibung einer systematischen Verzerrung bzw. einer systematischen Abweichung zwischen der Modellierung und der Wirklichkeit genutzt.¹³

Neben der direkten Bias-Wirkung (Disparate Impact) zwischen zwei Variablen können auch indirekte Beziehungen (Disparate Treatment) auftreten. Im Kontext des maschinellen Lernens ist die Differenzierung dieser Begriffsverständnisse von fundamentaler Bedeutung, da der Einsatz von KI-Algorithmen nicht zwangsläufig zu einer Vermeidung von Bias-Effekten führt, da die Verwendung der Algorithmen genau auf die zugrundeliegende Situation abgestimmt werden muss. Algorithmen, welche mithilfe von Daten trainiert werden, können trotz des Ausschlusses von sensitiven Attributen (z. B. Alter, Religion oder Geschlecht) unbeabsichtigte Diskriminierung mit unterschiedlichen Auswirkung hervorrufen.¹⁴ Dabei können diese diskriminierenden Muster nicht nur über Variablen wie Alter oder Geschlecht, sondern auch über andere Variablen wie Wohnort, Einkommen und Bildungsniveau erfasst werden, sodass das reine Ausschließen sensitiver Merkmale (z. B. Alter, Religion

¹¹ *Martins*, Enabling Non-technical Users to Query and Purchase Data (2019, S. 6f).

¹² *Hastie/Tibshirani/Friedman*, The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2009, S. 219 ff.).

¹³ *Hastie/Tibshirani/Friedman*, The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2009, S. 219 ff.).

¹⁴ *Ghosh/Kleinberg*, in Papadimitriou (Hrsg.), 8th Innovations in Theoretical Computer Science Conference (ITCS) (2017, S. 9:1 ff.).

oder Geschlecht) nicht ausreicht. Zur Vermeidung von Bias-Effekten bei der datengetriebenen Beantwortung einer unbekanntem Zielgröße sind daher weitere Schritte notwendig,¹⁵ welche nachfolgend erneut aufgegriffen werden (siehe Abschn. 4.).

Die Suche nach der optimalen Datenrepräsentation und -auswahl gestaltet sich somit stets als Balanceakt, bei dem es gilt, zwischen Modellgenauigkeit (wenig Prognosefehler) und Modellkomplexität (viele Modellparameter) abzuwägen. Und eben diese Suche nach dem optimalen Grad des Lernens, um einerseits die vorliegende Datenbasis exakt zu beschreiben und andererseits zukünftige Ereignisse präzise vorhersagen zu können, ist von je her im Versicherungswesen verwurzelt, da Risiken innerhalb der Versicherungsgemeinschaft proportional auf die entsprechenden Gruppen umgelegt werden sollen.

Der Umgang mit Bias ist alles andere als eine Trivialität, da dem Lernerfolg häufig eine große Vielzahl von möglichen Einflussfaktoren zugrunde liegt und mit dem Lernen daher ein hohes Maß an Komplexität verbunden ist. In der Literatur werden zahlreiche Ursachen für Bias angeführt, welche auf unterschiedlichste Weisen strukturiert werden. Vor diesem Hintergrund folgt diese Arbeit dem in Pessach und Shmueli dargestellten Ansatz,¹⁶ um einen Einblick über potenzielle Bias-Quellen zu geben:

- Bias-Effekte, welche bereits in den Trainingsdatensatz enthalten sind und auf verzerrten Messungen, verzerrten menschlichen bzw. kognitiven Entscheidungen, fehlerhaften Berichten oder ähnlichem basieren.
- Bias-Effekte, welche durch das Fehlen von Daten hervorgerufen werden und dadurch für die Zielpopulation nicht repräsentativ sind. Beispiele hierfür sind Stichprobenverzerrungen, Auswahlverzerrungen oder fehlende Werte.
- Bias-Effekte, welche sich aus unpassenden algorithmischen Zielsetzungen ergeben, wie etwa eine stark aggregierte Vorhersage, da in der Regel die Mehrheitsgruppen die Minderheiten im Rahmen der meisten KI-Verfahren dominieren werden.
- Bias-Effekte, welche durch „Proxy“-Attribute hervorgerufen werden. Proxy-Attribute sind nicht-sensitive Attribute, die einen Rückschluss auf sensitive Attribute (z. B. Alter, Religion oder Geschlecht) erlauben.¹⁷

Es gibt folglich nicht „den einen Bias“, sondern eine Vielzahl, und durch die menschliche Interaktion mit digitalen Systemen werden Datenbestände erzeugt, die diese Verzerrungen reflektieren und durch die immanente Systematik dieser Verzerrungen auf Algorithmen übertragen. In diesem Beitrag beschreibt Bias eine signifikante systematische Abweichung des Outputs eines Algorithmus von der Realität.¹⁸

Insbesondere im Bereich der KI wird Bias oft als Problem angesehen, da Bias das algorithmische Verhalten stark beeinflussen kann und teils dazu führt, dass Individuen oder Personengruppen ungleiche, unfaire oder gar diskriminierende KI-

¹⁵ Kamishima/Akaho/Asoh/Sakuma, in Flach/Bie/Christianini (Hrsg.), *Machine Learning and Knowledge Discovery in Databases* (2012, S. 35 ff.).

¹⁶ Pessach/Shmueli, *Algorithmic fairness* (2020, S. 2).

¹⁷ Friedman/Nissenbaum, *ACM Transactions on Information Systems* (1996), Heft 3, 330, 332f.

¹⁸ Friedman/Nissenbaum, *ACM Transactions on Information Systems* (1996), Heft 3, 330, 332f.

Entscheidungsausgänge erfahren (Abschn. 2.2). Ob und in welchem Ausmaß die KI-Entscheidungsausgänge von einem Bias betroffen sind, beruht dabei auf der Beurteilung der KI-Fehlerhäufigkeiten in Bezug auf die jeweiligen Individuen oder Personengruppen.

2.2 Die Gefahr eines Bias

Schon heute zeigt KI in vielen Bereichen ihre Leistungsfähigkeit. Sie arbeitet rund um die Uhr, erreicht zumeist gewünschte Ergebnisse innerhalb kürzester Zeit und ist dabei dem Menschen häufig sogar hinsichtlich qualitativer Aspekte überlegen. Sie kann aber auch manchmal ziemlich danebenliegen. So forderte etwa Googles Bilderkennung asiatisch aussehende Menschen dazu auf, die Augen zu öffnen.¹⁹ Ein Algorithmus, der Amazon bei der Bewerberauswahl unterstützen sollte, schlug vor allem Männer vor,²⁰ die zum Beispiel „Jared“ hießen und Lacrosse spielten.²¹ Und YouTube assoziierte den Großbrand von Notre-Dame mit dem 9/11-Terroranschlag.²² Zwar ermöglichen es Verfahren des maschinellen Lernens, sämtliche eingehende Informationen in ihrer Diversität zu verarbeiten, aber offensichtlich war in diesen Fällen die KI nicht in der Lage, relevante Merkmale optimal zu erkennen. Die obigen Fauxpas der IT-Riesen unterstreichen die Komplexität, die mit diesem für den Lernerfolg so kritischen Schritt einhergeht und lassen vermuten, dass KI einem Bias unterliegen kann.

Besonders schwerwiegend sind diese systematischen Fehler, wenn sie in KI-Systemen auftreten, welche mit weitreichenden Lebensentscheidungen verbunden sind. Das COMPAS-System von ProPublica ist ein vorprozessuales Risikomodell, das von zahlreichen US-Gerichten verwendet wird, um die Wahrscheinlichkeit für den Rückfall einer (inhaftierten) Person zu bestimmen. Auf Grundlage dieser Wahrscheinlichkeit wird entschieden, ob diese Person auf Kautionsfreilassung wird oder nicht. Dieses intensiv diskutierte System basiert auf Trainingsdaten, die die Rückfallraten von schwarzen Angeklagten aufgrund unzähliger Unterdrückungsformen, wie etwa übermäßige Polizeikontrollen in Schwarzen-Vierteln, überschätzen.²³ Das COMPAS-System weist somit eine signifikant höhere Fehlerhäufigkeiten in Abhängigkeit der ethnischen Gruppenzugehörigkeit auf und unterliegt demnach einem Bias. Neben dieser algorithmischen Voreingenommenheit zeigen Dressel und Farid

¹⁹ Fieser Code, 24.06.2019, abrufbar unter: <https://www.internetworld.de/marketing-trends/algorithmenvorurteile-ausspucken-1720064.html> [Stand: 26.07.2021].

²⁰ *Dastin*, Amazon scraps secret AI recruiting tool that showed bias against women (2018), abrufbar unter: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> [Stand: 26.07.2021].

²¹ *Gershgorin*, Companies are on the hook if their hiring algorithms are biased, 22.10.2018, abrufbar unter: <https://qz.com/1427621/companies-are-on-the-hook-if-their-hiring-algorithms-are-biased/> [Stand: 26.07.2021].

²² *Barkhausen*, Youtube kombiniert Notre-Dame-Feuer mit 9/11-Terror, abrufbar unter: <https://www.businessinsider.de/gruenderszene/business/youtube-spacex-apple-qualcomm-microsoft/> [Stand: 26.07.2021].

²³ *Green*, in FAT* *20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (2020, S. 594, 600).

auf, dass das COMPAS-System bei der Entscheidungsfindung keine besseren Ergebnisse erzielt als ein einfaches logistisches Regressionsmodell.²⁴ Dies bedeutet, dass der Einsatz von komplexen KI-Verfahren nicht zwangsläufig mit einem Informationsnutzen behaftet ist und stets hinterfragt werden sollte, ob die zusätzliche Modellkomplexität gerechtfertigt ist.

Ein weiteres aktuelles Anwendungsbeispiel von Algorithmen, die kognitive Verzerrungen lernen, ist Predictive Policing, das darauf abzielt, Kriminalität zu stoppen, bevor sie passiert.²⁵ Was zunächst nach einem dystopischen Szenario klingt, wird in einigen US-Staaten bereits praktiziert und zeigt in der Praxis auf, dass die KI-Algorithmen anstelle der kognitiven Vorurteile von Richtern (siehe COMPAS-Beispiel) die unerschwingliche Voreingenommenheit von Polizisten erkennen, lernen und replizieren.²⁶

Das Problem von KI-Bias lässt sich nicht auf eine bestimmte Anwendungsdomäne beschränken. Vor diesem Hintergrund haben sich in den letzten Jahren zahlreiche Institutionen, wie die Datenethikkommission der Bundesregierung, die Initiative *European Network of Human-Centered AI* (humane-ai.eu) oder „Data, Responsibly“ (dataresponsibly.github.io), etabliert, die sich mit moralischen KI-Aspekten auseinandersetzen. An der TU München gibt es hierzu bereits ein eigenes Institut (ieai.mcts.tum.de).

Die Wahrscheinlichkeit für einen Irrtum, sprich die Wahrscheinlichkeit für die Abbildung eines Bias, steigt, je mehr Entscheidungen getroffen werden. Dieser Bias-Effekt kann zu einer ungleichen bzw. unfairen Behandlung von Individuen oder Gruppen führen und sich unterschiedlich stark auswirken (sowohl positiv als auch negativ). Eine systematische Bevorzugung von Auszubildenden, Senioren oder Menschen mit Handicap etwa durch eine Tarifvergünstigung stößt eher auf gesellschaftliche Akzeptanz, während eine systematische Benachteiligung einer ethnischen Gruppe abgelehnt wird.

Eine (schwache) KI lernt, ohne die von ihr identifizierten Muster selbstständig zu hinterfragen. Werden diese Muster nicht weiter überprüft, kann es passieren, dass, wenn man eine KI bei der Arbeit beobachtet und die Entscheidungen, die diese trifft, analysiert, zu dem Ergebnis kommt, dass die KI die Entscheidungen nicht in der gewünschten Weise trifft. Kritisch ist dies vor allem, wenn festgestellt wird, dass die KI nicht gewollte beziehungsweise nicht zulässige Differenzierungen zwischen Versicherungsnehmern, potenziellen Kunden oder Dritten trifft, die KI also zum Beispiel Ergebnisse zeigt, die gegen das Verbot der Diskriminierung aufgrund der Rasse oder des Geschlechts verstoßen (s. hierzu Abschn. 3.1).

An diesem Beispiel und den anderen obigen Beispielen wird deutlich, dass ein Bias massiven Einfluss auf die algorithmische (Un-)Gerechtigkeit bzw. algorithmische Fairness, haben kann. Der Prozess des Lernens ist jedoch stets mit einem Bias behaftet. Vor diesem Hintergrund wird das Verständnis für algorithmische Fairness im weiteren Verlauf zunächst genauer betrachtet.

²⁴ Dressell/Farid, *Science Advances* 4 (2018), Heft 1, eea05580.

²⁵ McGrory/Bedi, *Tampa Bay Times*, Targeted, Artikel vom 03.09.2020.

²⁶ Barocas/Hardt/Narayanan, *Fairness and Machine Learning* (2019, S. 19).

2.3 Algorithmische (Un-)Fairness (?)

Zur Quantifizierung algorithmischer Fairness wurden bereits umfangreiche Arbeiten durchgeführt. Es existiert eine Reihe konkurrierender algorithmischer Fairness-Ansätze, welche im Rahmen der KI-Entwicklung eingesetzt werden kann. Jeder dieser algorithmischen Fairness-Ansätze repräsentiert dabei eine andere konzeptionelle Vorstellung von Fairness und entscheidet darüber, wie diese umgesetzt werden muss. Dies führt dazu, dass dieselbe Situation auf verschiedene Weisen interpretierbar ist, ob ein fairer (unverzerrter) Zustand eines KI-Systems vorliegt oder nicht. Die zugrundeliegende Fairness-Definition entscheidet über das genaue Ausmaß des Bias, sollte eine Abweichung zum fairen Zustand vorliegen. Es konnte jedoch gezeigt werden, dass die Umsetzung der unterschiedlichen Fairness-Definitionen zu ähnliche Ergebnissen führt.²⁷ Vor diesem Hintergrund scheint es zunächst als wichtiger, dass eine geeignete Fairness-Metrik im Rahmen der KI-gestützten Entscheidungsfindung eingesetzt wird, als die Frage zu beantworten, welche spezifische Fairness-Metrik verwendet werden soll. Zudem richtet sich die für eine KI zu fordernde Fairness zu einem wesentlichen Teil nach den Vorgaben der Rechtsordnung. Eine umfassende Diskussion der jeweiligen Fairness-Vorstellungen wird daher zurückgestellt und an dieser Stelle lediglich anhand von zwei Arten algorithmischer Fairness durchgeführt.

Eine Art der Fairness bildet die *Gruppenfairness*, welche darauf zielt, das Gesamtergebnis eines Systems auf Chancengleichheit zu überprüfen. Hierzu werden verschiedene demografische, ethnische oder sonstige Gruppen verglichen und es wird entschieden, ob eine signifikante systematische Abweichung von dem Konzept der Gruppenfairness vorliegt.²⁸ Im Beispiel eines Bewerbungsverfahrens fordert die Gruppenfairness, dass z.B. Männer und Frauen dieselbe Chance auf ein Vorstellungsgespräch haben. Dabei steht jedoch die Gruppenzugehörigkeit im Vordergrund. Es wird nicht betrachtet, ob ein Individuum einem Bias-Effekt unterliegt. Diese Prüfung auf systematischen Bevorzugung bzw. Benachteiligung ist Gegenstand der *individuellen Fairness*, welche eine ähnliche Behandlung von ähnlichen Personen fordert.²⁹ Im Zuge des Bewerbungsverfahrens sorgt die individuelle Fairness dafür, dass gleich qualifizierte Bewerber eingeladen werden.

Eine weiterführende Herausforderung zeigt sich jedoch im Zuge der Anwendung dieser algorithmischen Fairness-Ansätze, da in der Finanzbranche viele sensitive Attribute nicht erhoben werden und in der Datengrundlage somit wenige oder gar keine Informationen vorliegen.³⁰ Im Beispiel der Versicherung wird vom Versicherungsmakler nicht erfasst, ob etwa ein Akzent oder eine andere Hautfarbe gegeben ist. Eine Überprüfung, ob eine Beeinflussung (Bias) vorliegt, gestaltet sich insbesondere vor dem Hintergrund der DSGVO als herausfordernd, weil sich hierdurch zusätzli-

²⁷ Friedler/Scheidegger/Venkatasubramanian/Choudhary/Hamilton/Roth, in FAT* '19: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (2019, S. 329 ff.).

²⁸ Chouldechova, Big data 5 (2017), Heft 2, 153 ff.

²⁹ Ivento, in Roth (Hrsg.), 1st Symposium on Foundations of Responsible Computing (FORC) (2020, S. 2:1 ff.).

³⁰ Chen/Kallus/Maol/Svachal/Udell, in FAT* '19: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (2019, S. 339, 340f.).

che Anforderungen an die Speicherung personenbezogener Daten ergeben und eben diese zur Überprüfung von KI-Bias benötigt werden. Das Fehlen der Informationen verhindert jedoch einen Rückschluss der KI auf diese Information nicht zwangsläufig, da die „Wissenslücke“ über andere Merkmale erschlossen werden kann.³¹ Chen, et al. schlossen etwa mithilfe Bayesischer Verfahren unter Nutzung der Wohnorte und Vornamen auf die Ethnie zurück und konnten über diesen Weg die Fairness von Algorithmen evaluieren.³²

Das Ausmaß von algorithmischen Bias-Effekten wird am Beispiel von Obermeyer et al. deutlich. Sie untersuchten einen Algorithmus zur Vergabe von Pflegeprogrammen, mit dem in den USA jährlich ca. 200 Mio. Menschen in Berührung kommen, und zeigten auf, dass ein starkes Ungleichgewicht bei Pflegeprogrammvergabe zwischen Weißen und Schwarzen bei gleichem Erkrankungsgrad besteht.³³ Anhand dieser Größenordnung wird nicht nur das verheerende Potenzial deutlich, sondern auch, dass menschliche Interventionen zur Problemlösung eher ungeeignet sind.

3 Diskriminierungsverbote und ihre zivilrechtlichen Folgen

Technisch gesehen besteht, wie oben ausgeführt, die Gefahr, dass eine KI Entscheidungen trifft, welche durch Bias belastet sind. Ein durch Bias belasteter Output kann unter verschiedenen Aspekten unerwünscht sein. Der vorliegende Beitrag behandelt Fälle, in denen der Bias zu einer Ungleichbehandlung von Versicherungsnehmern oder sonstigen Beteiligten führt. Einer solchen Ungleichbehandlung zieht das Gesetz gewisse Grenzen. Rechtlich gesehen ist zu fragen, ob KI-Entscheidungen, die diese Grenzen überschreiten, eine Rechtsverletzung des Versicherers gegenüber dem Versicherungsnehmer darstellen.

Wie oben (Abschn. 1.) dargestellt kann KI gegenüber dem potenziellen Kunden, dem Versicherungsnehmer und auch gegenüber sonstigen Begünstigten in verschiedenen Szenarien zum Einsatz kommen. Hier wird als Beispiel ein Fall zugrunde gelegt, in dem der Versicherer zur Erstellung von Angeboten auf Abschluss von Versicherungsverträgen eine KI einsetzt, die aufgrund der vom Versicherungsnehmer eingegebenen Daten ein Angebot errechnet. Diese KI ist im oben beschriebenen Sinne ein selbstlernender Algorithmus, der mithilfe von Daten konkrete Vertragsangebote nach sich immer wieder verändernden, für den Versicherer im Einzelnen nicht voraussehbaren, aber am Ziel der Gewinnmaximierung orientierten Regeln errechnet. Im Beispielsfall wird unterstellt, dass die KI nach dem Abschluss einiger tausend Verträge gelernt hat, dass Männer höhere Versicherungsprämien akzeptieren als Frauen. Sie bietet daraufhin die betreffenden Versicherungsprodukte systematisch Männern zu höheren Preisen an als Frauen, auch wenn die Risiken identisch sind.

³¹ Kamishima/Akaho/Asoh/Sakuma, in Flach/Bie/Christianini (Hrsg.), *Machine Learning and Knowledge Discovery in Databases* (2012), S. 35 ff.).

³² Chen/Kallus/Maol/Svachal/Udell, in FAT* '19: *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency* (2019), S. 339 ff.).

³³ Obermeyer/Powers/Vogel/Mullainathan, *Science* 366 (2019), Heft 6464, 447 ff.

3.1 Verstoß gegen des Allgemeine Gleichbehandlungsgesetz

3.1.1 Adressat des Diskriminierungsverbots und Anspruchsgegner

Bias kann zu Ungleichbehandlung führen. § 19 Abs. 1 Nr. 2 AGG verbietet eine Benachteiligung aus Gründen der Rasse oder wegen der ethnischen Herkunft, wegen des Geschlechts, der Religion, einer Behinderung, des Alters oder der sexuellen Identität bei der Begründung, Durchführung und Beendigung zivilrechtlicher Schuldverhältnisse, die eine privatrechtliche Versicherung zum Gegenstand haben. Gemäß § 20 Abs. 1 S. 1 AGG liegt eine Verletzung des Benachteiligungsverbots ausnahmsweise nicht vor, wenn für eine unterschiedliche Behandlung wegen der Religion, einer Behinderung, des Alters oder der sexuellen Identität ein sachlicher Grund vorliegt.³⁴

In Bezug auf den hier diskutierten Einsatz von KI im direkten Kundenkontakt stellt sich zuerst die Frage, wer Adressat des Diskriminierungsverbots und damit Anspruchsgegner ist. Ist der KI selbst die Diskriminierung verboten? Oder richtet sich das Verbot an den die KI einsetzenden Versicherer? Auch derjenige, welcher die KI ursprünglich entwickelt hat, kommt als Adressat des Verbots in Betracht.

Grundsätzlich ist Adressat, wer das Schuldverhältnis begründet, durchgeführt oder beendet hat (sog. Anbieter).³⁵ Nach bisheriger Rechtsauffassung kann dies nur ein Rechtssubjekt sein. Unter dem Stichwort der E-Person wird aktuell diskutiert, ob eine KI mit einer Rechtspersönlichkeit versehen werden soll.³⁶ Solange dies jedoch aussteht, kann eine KI mangels Rechtspersönlichkeit nicht Adressat des Diskriminierungsverbots sein. Der KI-Entwickler scheidet als Adressat ebenfalls aus, da er mit den Kunden des Versicherers in keinem Schuldverhältnis steht. Somit bleibt als Adressat nur der Versicherer, der die KI einsetzt.

Zu klären ist weiter, worin die Benachteiligung konkret besteht und wer benachteiligt. Benachteiligt werden Männer, die eine höhere Prämie angeboten bekommen als Frauen. Dieses Angebot ist als Willenserklärung dem Versicherer zurechenbar.³⁷ Die Erklärung stellt sich für den Rechtsverkehr als eine vom Handlungs-, Erklä-

³⁴ Eine Rechtfertigung der Benachteiligung wegen des Geschlechts ist nach dem Test-Achats-Urteil des EuGHs (EuGH, Urteil vom 01.03.2011, C-236/09) und Umsetzung im AGG nicht mehr möglich (MüKo BGB/Thüsing, 8. Aufl. (2018), § 20 AGG, Rn. 55, 57 f.; Däubler/Bertzbach/Klose/Braunroth, Allgemeines Gleichbehandlungsgesetz, 4. Aufl. (2018), § 20 Abs. 2 AGG, Rn. 36; BeckOGK AGG/Mörsdorf, Stand: 01.06.2021, § 20 AGG, Rn. 48).

³⁵ MüKo BGB/Thüsing, 8. Aufl. (2018), § 19 AGG, Rn. 125, 130; BeckOGK AGG/Mörsdorf, Stand: 01.06.2021, § 19 AGG, Rn. 67.

³⁶ Linardatos, Autonome und vernetzte Aktanten im Zivilrecht, Grundlinien zivilrechtlicher Zurechnung und Strukturmerkmale einer elektronischen Person (2021); Eidenmüller/Wagner, Law by Algorithm (2021, S. 103 ff.); Kainer/Förster, ZfPW (2020), 275, 295 ff.; Specht/Herold, MMR (2018), 40, 43 f.; Groß, CB (2020), 155, 158; Linke, MMR (2021), 200, 201 ff.; Denga, CR (2018), 69, 77; Martini, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz (2019, S. 290 ff.).

³⁷ Ob eine KI selbst eine Willenserklärung abgeben kann und ob diese dem KI-Einsetzenden zurechenbar ist, ist bis heute umstritten. Es werden verschiedenen Lösungswege diskutiert (z. B. bei Groß, CB (2020), 155, 158; Haseloff/Friehoff, VersR (2020), 1363, 1363 ff., 1367 f.; Weingart, CR (2020), 701, 702 ff., 704; Specht/Herold, MMR (2018), 40, 43). Andere Stimmen in der Literatur lehnen eine Willenserklärung einer KI bzw. deren Zurechnung mit den Mitteln des geltenden Rechts dagegen ab (so etwa Graf von Westphalen, BB (2020), 1859, 1862; Kainer/Förster, ZfPW (2020), 275, 281 ff., 295). Zentral scheint bei dieser Diskus-

rungs- und Geschäftswillen des Versicherers getragene Erklärung dar, weil der Versicherer und nicht die KI als bewusst Handelnder auftritt. Subjektiv muss die Erklärung von Handlungswillen und (mindestens potentiell) Erklärungswillen gedeckt sein.³⁸ Ein Geschäftswille hinsichtlich des konkreten Geschäfts ist für eine wirksame Erklärung nicht erforderlich.³⁹ Auch diese subjektiven Voraussetzungen liegen vor. Zwar weiß der Versicherer bei Einsatz der sich stetig fortentwickelnden, selbstlernenden KI nicht konkret, nach welchen Kriterien die KI welche Angebote errechnet, mit welcher Prämie und für welche Kunden. Sein Wille umfasst aber jedenfalls das Handeln im rechtsgeschäftlichen Bereich, was für eine wirksame Willenserklärung genügt. Auch wird sein Wille i. d. R. sämtliche von der KI hervorgebrachten Angebote decken, wenn er die Funktionsweise der KI einschließlich ihrer unvorhersehbaren Elemente kennt. Im Ergebnis ist daher der Versicherer derjenige, der i. S. d. AGG benachteiligt.

Eine eigene Willenserklärung der KI, welche dem Versicherer etwa nach Stellvertretungsrecht oder analog § 278 BGB⁴⁰ zugerechnet werden könnte, ist hingegen schon mangels Rechtspersönlichkeit der KI abzulehnen.⁴¹

3.1.2 Kausalität und Erklärbarkeit

Dieses benachteiligende Verhalten des Versicherers muss zudem kausal an eines der in § 1 oder § 19 Abs. 1 AGG genannten Merkmale anknüpfen.⁴² Hierbei wird das Problem der Erklärbarkeit von KI relevant. Kausalität liegt vor, wenn die KI das benachteiligende Angebot aufgrund der Anknüpfung an eines der verbotenen Merkmale berechnet hat. Um dies feststellen zu können, muss nachvollzogen werden, warum die KI das Ergebnis, also das Vertragsangebot, errechnet hat. Dies ist nach heutigem Stand der Technik, insbesondere bei sog. tiefen neuronalen Netzen,

sion jedoch zu sein, stets eindeutig zu unterscheiden, von welcher Form von KI, in welchem Einsatzgebiet, mit welchem Autonomie-Grad gesprochen wird.

³⁸ BGH, Urteil vom 07.06.1984, IX ZR 66/83, in BGHZ 91, 324, Rn. 21 ff.; Bork, Allgemeiner Teil des Bürgerlichen Gesetzbuchs (2016), Rn. 589, 593, 596; BeckOK BGB/Wendtland (2021), Stand: 01.05.2021, § 133 BGB, Rn. 5 f., 8; MüKo BGB/Armbrüster, 8. Aufl. (2018), Vor § 116 BGB, Rn. 3; Schulze/Dörner, Bürgerliches Gesetzbuch, 10. Aufl. (2019), Vorbemerkung zu §§ 116–144 BGB, Rn. 4f.

³⁹ Bork, Allgemeiner Teil des Bürgerlichen Gesetzbuchs (2016), Rn. 600; BeckOK BGB/Wendtland (2021), Stand: 01.05.2021, § 133 BGB, Rn. 7; Schulze/Dörner, Bürgerliches Gesetzbuch, 10. Aufl. (2019), Vorbemerkung zu §§ 116–144 BGB, Rn. 6; MüKo BGB/Armbrüster, 8. Aufl. (2018), Vor § 116 BGB, Rn. 28.

⁴⁰ Im AGG-Recht ist das Verhalten Dritter grds. nach § 278 BGB analog zurechenbar (MüKo BGB/Thüsing, 8. Aufl. (2018), § 19 AGG, Rn. 128; Däubler/Bertzbach/Deinert, Allgemeines Gleichbehandlungsgesetz, 4. Aufl. (2018), § 21 AGG, Rn. 21). Über diese Figur des Erfüllungsgehilfen wird demjenigen, der einen anderen in seinem Rechts- und Pflichtenkreis für sich einsetzt, dessen Verschulden und analog dann auch dessen Pflichtverletzung, bzw. hier dann dessen benachteiligendes Verhalten zugerechnet.

⁴¹ So hinsichtlich § 278 BGB Weingart, CR (2020), 701, 703; a.A. Keßler, MMR (2017), 589, 592; Zech, ZfPW (2019), 198, 211 f.; differenziert Hacker, RW (2018), 243, 249 ff. Zum Stellvertretungsrecht siehe z. B. Weingart, CR (2020), 701, 702; Specht/Herold, MMR (2018), 40, 43; Keßler, MMR (2017), 589, 592.

⁴² BeckOGK AGG/Baumgärtner, Stand: 01.06.2021, § 3 AGG, Rn. 47 ff.; Däubler/Bertzbach/Schraderl Schubert, Allgemeines Gleichbehandlungsgesetz, 4. Aufl. (2018), § 3 AGG, Rn. 45.

nicht mehr möglich.⁴³ Problematisch ist dies im AGG-Kontext für den Versicherer deshalb, weil er gemäß § 22 AGG die Beweislast für die Kausalität trägt, wenn der Betroffene Indizien beweisen kann;⁴⁴ dass auch für den Betroffenen der Beweis von Indizien schwierig sein kann, liegt auf der Hand.⁴⁵ Ohne eine erklärbare KI läuft der Versicherer Gefahr, sich mangels möglichen Beweises fehlender Kausalität nicht von dem Vorwurf eines Verstoßes gegen § 19 Abs. 1 Nr. 2 AGG befreien zu können.⁴⁶ Ähnlich gelagert taucht das Problem auch bei der Rechtfertigung nach § 20 AGG auf bzw. im Tatbestand einer mittelbaren Benachteiligung i. S. d. § 3 Abs. 2 AGG.⁴⁷ Hier muss der Versicherer belegen, dass die KI die Benachteiligungen, die gerechtfertigt werden können (d. h. wegen der Religion, einer Behinderung, des Alters oder der sexuellen Identität), aus einem sachlichen Grund vorgenommen hat (§ 20 Abs. 1 S. 1 AGG) bzw. ein rechtmäßiges Ziel mit angemessenen und erforderlichen Mitteln verfolgt hat (§ 3 Abs. 2 AGG).

Eine diskriminierende Absicht oder subjektive Motivation des Versicherers ist für das Vorliegen eines AGG-Verstoßes hingegen nicht erforderlich.⁴⁸ Daher kann der Versicherer auch nicht argumentieren, er habe die Funktionsweise der KI nicht verstanden und kein diskriminierendes Verhalten durch den Einsatz der KI beabsichtigt.⁴⁹

Insgesamt lässt sich festhalten, dass ein Versicherer durch einen algorithmischen Bias der von ihm eingesetzten KI gegen das AGG verstoßen kann oder mangels Erklärbarkeit der KI nicht beweisen kann, dass kein Verstoß vorliegt.

Die Folgen für den Versicherer als Anspruchsgegner sind Beseitigungs- und Unterlassungsansprüche (§ 21 Abs. 1 AGG) sowie bei Vorliegen von Verschulden (hierzu ausführlicher unter Abschn. 3.5) Schadensersatzansprüche gerichtet auf den Ersatz materieller und immaterieller Schäden (§ 21 Abs. 2 AGG). Hinzu kommen kann (mit unterschiedlicher Begründung in der Literatur) ein Kontrahierungszwang.⁵⁰ Zu-

⁴³ Zur Problematik insgesamt siehe etwa *Zech, ZfPW (2019)*, 198, 202, 217 f.; *Kädel/Maltzan, CR (2020)*, 66 ff.; *Nguyen, VW (2019)*, Heft 9, 52 ff.

⁴⁴ Näher *Grünberger, ZRP (2021)*, 232, 233.

⁴⁵ *Grünberger, ZRP (2021)*, 232, 234 ff., schlägt deshalb eine weitergehende sekundäre Darlegungslast der KI-Anwender mit einem safe harbour vor.

⁴⁶ *Bilski/Schmid, NJOZ (2019)*, 657, 660; i.E. ebenso *Nguyen, VW (2019)*, Heft 9, 52, 54; zur Beweisproblematik auch *Zech, ZfPW (2019)*, 198, 217 f.; sowie *Kolleck/Orwat (Oktober 2020)*, Mögliche Diskriminierung durch algorithmische Entscheidungssysteme und maschinelles Lernen – ein Überblick, S. 40, welche jedoch trotz § 22 AGG ein Beweisproblem für den Betroffenen mangels Nachvollziehbarkeit der Wirkungsweisen von Algorithmen sehen.

⁴⁷ *Bilski/Schmid, NJOZ (2019)*, 657, 659 nimmt an, dass die „typisierende Vorgehensweise der Algorithmen“ eine mittelbare Diskriminierung darstellen kann. Ähnlich *Kolleck/Orwat (Oktober 2020)*, Mögliche Diskriminierung durch algorithmische Entscheidungssysteme und maschinelles Lernen – ein Überblick, S. 40.

⁴⁸ *Däubler/Bertzbach/Schrader/Schubert, Allgemeines Gleichbehandlungsgesetz, 4. Aufl. (2018)*, § 3 AGG, Rn. 46; *BeckOGK AGG/Baumgärtner, Stand: 01.06.2021*, § 3 AGG, Rn. 49; *Bilski/Schmid, NJOZ (2019)*, 657, 659.

⁴⁹ *Bilski/Schmid, NJOZ (2019)*, 657, 659.

⁵⁰ *Däubler/Bertzbach/Deinert, Allgemeines Gleichbehandlungsgesetz, 4. Aufl. (2018)*, § 21 AGG, Rn. 90; *MüKo BGB/Thüsing, 8. Aufl. (2018)*, § 21 AGG, Rn. 17; a.A. *Prölss/Martin/Armbrüster, Versicherungsvertragsgesetz, 31. Aufl. (2021)*, Einleitung, Rn. 320.

dem können das benachteiligende Rechtsgeschäft oder einzelne Klauseln (§ 21 Abs. 4 AGG) gemäß § 134 BGB unwirksam sein.⁵¹

3.2 Verstoß gegen das Gendiagnostikgesetz

§ 18 Gendiagnostikgesetz (GenDG) soll eine Benachteiligung auf Grund genetischer Eigenschaften verhindern und die Menschenwürde sowie das Recht auf informationelle Selbstbestimmung schützen (§ 1 GenDG). Das kann v. a. bei Personenversicherungen, wie etwa Berufsunfähigkeits- oder Lebensversicherungen relevant werden. Dem Versicherer wird in § 18 Abs. 1 S. 1 GenDG verboten, vor oder nach dem Abschluss des Versicherungsvertrages die Vornahme genetischer Untersuchungen oder Analysen oder die Mitteilung von Ergebnissen oder Daten aus bereits vorgenommenen genetischen Untersuchungen oder Analysen zu verlangen oder solche Ergebnisse oder Daten entgegenzunehmen oder zu verwenden. Es handelt sich somit weniger um ein Diskriminierungsverbot als vielmehr um ein früher ansetzendes Verbot, bestimmte Informationen zu erheben, zu verlangen, entgegenzunehmen oder zu verwenden.⁵²

Für die Nutzung von KI hat dieses Verbot vor allem zur Folge, dass sichergestellt werden muss, dass die KI weder direkt noch indirekt, also über andere Informationen, an die nach § 18 Abs. 1 S. 1 Nr. 2 GenDG mit einem Verwendungsverbot belegten Informationen gelangt und diese somit auch nicht verwenden kann.⁵³ In der Beratung durch eine KI ist zudem darauf zu achten, dass die KI keine nach § 18 Abs. 1 S. 1 GenDG unzulässigen Daten erhebt und den (potenziellen) Versicherungsnehmer nicht zu einer genetischen Untersuchung auffordert. Werden diese Sicherheitsvorkehrungen im Vorfeld getroffen, scheint es unwahrscheinlich, dass durch KI ein Verstoß gegen diese Vorgaben begangen wird. Hier ist somit weniger die Erklärbarkeit von KI relevant als vielmehr die im Vorfeld stattfindende Implementierung spezifischer rechtlicher Vorgaben in die KI, bevor diese vom Versicherer eingesetzt wird. Zudem ist zu überwachen, dass die KI nicht in Folge der stetigen Weiterentwicklung irgendwann von diesen Vorgaben abweicht.

Bei einem Verstoß gegen § 18 Abs. 1 S. 1 Nr. 2 GenDG drohen nach § 25 Abs. 1 Nr. 5 GenDG strafrechtliche Konsequenzen. Zivilrechtlich ist § 18 GenDG zudem als Verbotsgesetz i.S.d. § 134 BGB anzusehen, mit der Folge, dass Klauseln, die auf Verstößen gegen § 18 GenDG beruhen, nichtig sind.⁵⁴

⁵¹ MüKo BGB/*Thüsing*, 8. Aufl. (2018), § 21 AGG, Rn. 76 ff.; Däubler/Bertzbach/*Deinert*, Allgemeines Gleichbehandlungsgesetz, 4. Aufl. (2018), § 21 AGG, Rn. 104 ff.

⁵² MüKo BGB/*Thüsing*, 8. Aufl. (2018), § 20 AGG, Rn. 74; Ernst/Rogler/*Mertens*, Berufsunfähigkeitsversicherung (2018), § 6 BUV, Rn. 224.

⁵³ In Bezug auf sämtliche Fälle von Diskriminierungen so auch *Beck*, ZRP (2019), 185.

⁵⁴ *Präve*, VersR (2009), 857, 862.

3.3 Verstoß gegen Gleichbehandlungsgebote des Versicherungsrechts

Im Versicherungsrecht gilt zwar nach h.M. kein allgemeines Gebot der Gleichbehandlung der Versicherungsnehmer,⁵⁵ es finden sich aber sparten- oder rechtsformspezifische Gleichbehandlungsgebote.⁵⁶ Sie sind beim Einsatz von KI ebenso im Vorfeld zu implementieren, und ihre dauerhafte Einhaltung ist zu überwachen. Zudem ist zu fragen, wie der Versicherer gegenüber der Aufsichtsbehörde nachweisen kann, dass die eingesetzte KI diese versicherungsrechtlichen Vorgaben beachtet. Hierbei handelt es sich auch wieder um ein Erklärbarkeitsproblem.

Verstöße gegen diese Gebote haben vor allem aufsichtsrechtliche Konsequenzen (s. Abschn. 5.2). Zivilrechtliche Konsequenzen werden dagegen überwiegend abgelehnt.⁵⁷

3.4 Verstoß gegen Art. 3 GG

In den Grundrechten findet sich in Art. 3 Abs. 1 GG ein allgemeiner Gleichheitssatz beziehungsweise ein allgemeines Gleichbehandlungsgebot. Des Weiteren enthält Art. 3 Abs. 3 GG ein Diskriminierungsverbot in Bezug auf die Merkmale „Geschlecht“, „Abstammung“, „Rasse“, „Sprache“, „Heimat und Herkunft“, „Glauben“, „religiöse oder politische Anschauungen“ und „Behinderung“. Grundsätzlich gilt dieses Grundrecht auf Gleichbehandlung nur im Verhältnis des Staates zum Bürger. Eine Ausnahme ist unter dem Stichwort der sogenannten „mittelbaren Drittwirkung“ anerkannt, wonach unter anderem bei struktureller Überlegenheit des einen Vertragsteils das Diskriminierungsverbot des Art. 3 GG auch im Verhältnis Bürger-Bürger relevant werden kann und eine Ungleichbehandlung eines sachlichen Grundes bedarf.⁵⁸

Im Hinblick auf den Einsatz von KI durch einen Versicherer ist die Frage zu klären, ob und ab wann durch diesen Einsatz eine derartige strukturelle Überlegenheit des Versicherers gegenüber dem (potenziellen) Versicherungsnehmer angenommen werden kann, sodass Art. 3 GG zwischen Versicherer und Versicherungsnehmer anwendbar ist.⁵⁹ Angemerkt sei hierzu, dass diese mittelbare Drittwirkung des Gleichheitssatzes, welche im Ergebnis einer unmittelbaren Anwendung des Gleichheitssat-

⁵⁵ MüKo VVG/Looschelders, 2. Aufl. (2016), § 1 VVG, Rn. 59 ff., 64; Jung, VersR (2003), 282, 284; Wandt, Versicherungsrecht (2016), Rn. 140; Brand/Baroch Castellvi/Baroch Castellvi, Versicherungsaufsichtsgesetz (2018), § 138 VAG, Rn. 13; Prölss/Dreher/Präve, Versicherungsaufsichtsgesetz, 13. Aufl. (2018), § 138 VAG, Rn. 12.

⁵⁶ § 138 Abs. 2, §§ 146 Abs. 2, 147 i. V. m. § 138 Abs. 2, § 161 Abs. 1 i. V. m. § 138 Abs. 2, § 177 Abs. 1 VAG; dazu Bartel/Khuat/Wehling-Mauntel, in Chibanguza/Kuß/Steegen (Hrsg.), Künstliche Intelligenz, Recht und Praxis automatisierter und autonomer Systeme (2022, S. 1049–1064, Rn. 27 f.).

⁵⁷ Brand/Baroch Castellvi/Baroch Castellvi, Versicherungsaufsichtsgesetz (2018), § 138 VAG, Rn. 17; Prölss/Dreher/Präve, Versicherungsaufsichtsgesetz, 13. Aufl. (2018), § 138 VAG, Rn. 25; Prölss/Dreher/Weigel, Versicherungsaufsichtsgesetz, 13. Aufl. (2018), § 177 VAG, Rn. 17.

⁵⁸ BVerfG, Beschluss vom 11.04.2018, 1 BvR 3080/09, Rn. 41; Steege, MMR (2019), 715, 718; BeckOK GG/Kischel (2021), Stand: 15.05.2021, Art. 3 GG, Rn. 93 ff.

⁵⁹ Dazu Härtel, LKV (2019), 49, 57; Steege, MMR (2019), 715, 718.

zes zwischen Privaten gleichkommt,⁶⁰ aufgrund der darin liegenden Einschränkung der Privatautonomie restriktiv anzuwenden ist.⁶¹ Kann dennoch die Anwendbarkeit des Gleichbehandlungsgebots bejaht werden, stellen sich ähnliche Fragen wie bei der AGG-Prüfung. Wieder ist zu klären, ob KI „wegen“ der genannten Merkmale eine Entscheidung trifft, auch auf sachliche Gründe zur Rechtfertigung kommt es an. Zuletzt ist zu fragen, ob aus der Anwendung des Art. 3 GG ein Kontrahierungszwang und/oder ein Unterlassungs-/Beseitigungsanspruch folgt.⁶²

3.5 Verletzung vertraglicher Rücksichtnahmepflichten

Im Rahmen des KI-Einsatzes können zudem vertragliche Rücksichtnahmepflichten i. S. d. § 241 Abs. 2 BGB verletzt werden⁶³ und gegen den Versicherer Schadensersatzansprüche gemäß § 280 Abs. 1 BGB, im vorvertraglichen Bereich i. V. m. § 311 Abs. 2 BGB, entstehen. Betrachtet man das vor Abschn. 3.1 beschriebene Beispiel, so ist die jeweilige Rücksichtnahmepflichtverletzung in dem AGG-widrigen, Männer aufgrund ihres Geschlechts benachteiligenden, Angebot zu sehen, welches wie dargestellt eine Willenserklärung des Versicherers selbst darstellt (Abschn. 3.1.1). Auch hier sind somit Zurechnungsfragen, wie etwa über die Figur des Erfüllungsgehilfen (§ 278 BGB analog) nicht relevant.

Eine Schadensersatzpflicht des Versicherers entsteht jedoch nur dann, wenn ihn bezüglich der Pflichtverletzung nach § 280 Abs. 1 S. 2 BGB auch Verschulden trifft, er also vorsätzlich oder fahrlässig i. S. d. § 276 BGB gehandelt hat.⁶⁴ Dass ein Versicherer vorsätzlich AGG-widrige, diskriminierende Angebote abgibt, erscheint realitätsfern. Es ist somit zu fragen, wann dem Versicherer Fahrlässigkeit vorzuwerfen ist, das heißt, wann er die im Verkehr erforderliche Sorgfalt außer Acht gelassen hat (§ 276 Abs. 2 BGB). Klärungsbedürftig ist dabei, ob die Pflichtverletzung, also die Diskriminierung, vorhersehbar und vermeidbar war.⁶⁵

Wie dargestellt gibt es heutzutage keine KI ohne Bias. Dies muss auch dem die KI einsetzenden Versicherer bewusst sein. Damit sind jedoch auch jedes diskriminierende Angebot oder sonstige diskriminierende Verhalten und dessen Folgen in

⁶⁰ BeckOK GG/Kischel (2021), Stand: 15.05.2021, Art. 3 GG, Rn. 93a; Neuner, NJW (2020), 1851, 1854.

⁶¹ BeckOK GG/Kischel (2021), Stand: 15.05.2021, Art. 3 GG, Rn. 93a.

⁶² Hierzu etwa Jobst, NJW (2020), 11, 13 ff.

⁶³ Weingart, CR (2020), 701, 705 nimmt an, dass eine Verletzung von Sorgfaltspflichten „ohne weiteres“ auch von einem „Verhandlungsagenten“ begangen werden, welcher sich insbesondere auch über eine gewisse Autonomie auszeichnen soll (701 f.).

⁶⁴ Auch hieran würde eine Haftung für ein Verhalten der KI, welches über § 278 BGB zugerechnet wird scheitern. So müsste dem Versicherer auch das Verschulden der KI zugerechnet werden. Die Verschuldensfähigkeit einer KI wird jedoch bis heute überwiegend abgelehnt (Intveen, ITRB (2020), 70, 72 f.; Graf von Westphalen, BB (2020), 1859, 1863; Günther/Böglmüller, BB (2017), 53, 55; Zech, ZfPW (2019), 198, 211; Keßler, MMR (2017), 589, 592. Anders über Ansatz einer „funktionalen Verschuldensäquivalenz“ und analoge Anwendung des § 278 S. 1 BGB Hacker, RW (2018), 243, 259; dies für „denkbar“ haltend Zech, ZfPW (2019), 198, 211; ebenso über eine analoge Anwendung des § 278 S. 1 BGB oder direktes Abstellen auf das Verschulden des Schuldners ohne zwischen den Handelnden zu differenzieren zudem Keßler, MMR (2017), 589, 592.

⁶⁵ MüKo BGB/Grundmann, 8. Aufl. (2019), § 276 BGB, Rn. 68, 77; BeckOGK BGB/Schaub (2021), Stand: 01.03.2021, § 276 BGB, Rn. 60, 65.

ausreichendem Maße vorhersehbar – der Versicherer muss immer damit rechnen, dass eine KI, welche sich auch im Einsatz stetig weiterentwickelt, irgendwann an verbotene Merkmale anknüpft und entsprechend zu diskriminierenden Ergebnissen kommt.⁶⁶

Fraglich ist, ob eine Diskriminierung und deren Folgen für den Versicherer auch vermeidbar sind.⁶⁷ Mit Maßnahmen des Bias-Managements kann Bias einer selbstlernenden KI bis zu einem gewissen Grad, jedoch niemals vollständig unterbunden werden (näher unten Abschn. 4.). Soweit nach dem Stand der Technik bestimmte Maßnahmen des Bias-Managements geboten sind, handelt ein Versicherer, der sie nicht einsetzt, mindestens fahrlässig und damit schuldhaft. Da aber keine Maßnahme geeignet ist, Bias selbstlernender KI ganz zu verhindern, sind das diskriminierende Verhalten und seine Folgen nie vollständig vermeidbar.

Das bedeutet aber nicht, dass der Versicherer nicht fahrlässig handelt. Vermeiden kann er Diskriminierung durch selbstlernende KI, indem er ihren Einsatz im Verhältnis zum Kunden ganz unterlässt. Das bedeutet, dass ein Versicherer, der eine KI mit unvermeidbaren Diskriminierungsfolgen für Kunden einsetzt, stets fahrlässig handelt. Er ist dann verpflichtet, voraussehbare Schäden zu ersetzen. Dieses Ergebnis erscheint auf den ersten Blick innovationsfeindlich. Allerdings lässt sich die Nutzung von selbstlernender KI mit dem ihr imminenten Risiko der Diskriminierung keinesfalls als erlaubtes, sozialadäquates Risiko einordnen, das hinzunehmen ist,⁶⁸ und dessen Folgen der diskriminierte Versicherungsnehmer tragen muss. Inwieweit Gefahren durch KI zu akzeptieren sind und wer die Schäden trägt, ist eine der großen offenen Fragen unserer Zeit, die die öffentliche Diskussion und Gesetzgeber international beschäftigen. Von einem erlaubten Risiko kann nicht die Rede sein. Zudem drohen Versicherungsnehmern durch Diskriminierung durchaus existentielle Schäden, etwa durch Ablehnung von Versicherungsschutz. Dem Einwand der Innovationsfeindlichkeit ist zudem entgegenzuhalten, dass der Versicherer nicht daran gehindert wird, KI einzusetzen (zu den Vorgaben nach dem Kommissionsentwurf einer KI-VO s. unten Abschn. 6.). Er muss nur die Haftungsrisiken in seine geschäftliche Entscheidung einbeziehen und sie den Vorteilen, welche er durch den Einsatz von KI hat (Effizienzgewinne, Kostenvorteile, etc.) gegenüberstellen.

Hier wird somit in Zukunft sicherlich noch Diskussions- und ggf. auch gesetzgeberische Handlungsbedarf bestehen. Hingewiesen sei im diesem Kontext auf die aktuell diskutierte Gefährdungshaftung⁶⁹ für KI sowie eine mögliche Versicherungslösung.⁷⁰

⁶⁶ Hierzu auch *Spindler*, CR (2015), 766, 768; anders im Ansatz für das Verschulden des Herstellers so *Graf von Westphalen*, BB (2020), 1859, 1864.

⁶⁷ Zur Frage der Vermeidbarkeit bei der Haftung für KI generell siehe auch *Graf von Westphalen*, BB (2020), 1859, 1864; *Weingart*, CR (2020), 701, 705 f.

⁶⁸ Dazu etwa BeckOGK BGB/*Schaub* (2021), Stand: 01.03.2021, § 276 BGB, Rn. 69.

⁶⁹ Hierzu etwa *Günther/Böglmüller*, BB (2017), 53, 55; *Intveen*, ITRB (2020), 70, 73; *Denga*, CR (2018), 69, 76 f.; *Bilski/Schmid*, NJOZ (2019), 657, 660; *Groß*, CB (2020), 155, 158 f.; *Spindler*, CR (2015), 766, 775 f.; *Wagner*, VersR (2020), 717, 734 f.; *Zech*, ZfPW (2019), 198, 214 ff.; *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz (2019), S. 287 ff.

⁷⁰ Hierzu etwa *Basedow*, EuZW (2021), 1 f.; *Denga*, CR (2018), 69, 76 f.; *Graf von Westphalen*, BB (2020), 1859, 1864 f.; *Graf von Westphalen*, VuR (2020), 248, 254; *Günther/Böglmüller*, BB (2017), 53,

3.6 Verstöße gegen Deliktsrecht

Wurde der Versicherungsnehmer durch die KI diskriminiert, kommt auch ein Verstoß gegen das allgemeine Deliktsrecht, insbesondere eine Verletzung des Allgemeinen Persönlichkeitsrechts in Betracht.⁷¹ Für die Frage nach dem Verschulden kann weitgehend auf die Ausführungen unter Abschn. 3.5. verwiesen werden. Ein wichtiger Unterschied liegt in der Beweislast – während bei § 280 BGB der Versicherer beweisen muss, nicht schuldhaft gehandelt zu haben (§ 280 Abs. 1 S. 2 BGB), ist es hier am Versicherungsnehmer zu beweisen, dass der Versicherer schuldhaft gehandelt hat. Im Übrigen wird diskutiert, ob der Versicherer auch für ein Verhalten der KI nach § 831 BGB haften muss. Hierfür müsste die KI jedoch Verrichtungsgehilfe sein, was wiederum vor dem Hintergrund fehlender Rechtspersönlichkeit zweifelhaft ist.⁷² Des Weiteren kann eine Haftung aus § 823 Abs. 2 BGB i. V. m. § 19 AGG⁷³ oder § 18 GenDG⁷⁴ als Schutzgesetz relevant werden. Auch kommen Beseitigungs- oder Unterlassungsansprüche aus § 1004 Abs. 1 BGB analog in Betracht.

3.7 Verstöße gegen Datenschutzrecht

Das Datenschutzrecht schützt neben dem Datenverkehr natürliche Personen bei der Verarbeitung personenbezogener Daten⁷⁵ und soll sie in bestimmten Konstellationen vor Diskriminierung bewahren. Art. 22 Abs. 4 DSGVO, der es grundsätzlich verbietet, eine automatisierte Entscheidung nach bestimmten Kategorien personenbezogener Daten zu treffen, ist auch als Diskriminierungsverbot zu verstehen.⁷⁶ Soweit ausnahmsweise aufgrund einer Einwilligung automatisierte Entscheidungen auf Grundlage der genannten Datenkategorien getroffen werden dürfen (Art. 9 Abs. 2 lit. a DSGVO), müssen „angemessene Maßnahmen zum Schutz der Rechte und Freiheiten sowie der berechtigten Interessen der betroffenen Person“ (Art. 22 Abs. 4 DSGVO i. V. m. Erwägungsgrund 71) ergriffen werden. So muss der Betroffene die Möglichkeit haben, eine Entscheidung einer Person zu erwirken und dabei individuelle Aspekte einzubringen.⁷⁷ Auch muss er über die Tatsache und die Logik der automatisierten Entscheidung informiert werden. Diese Pflichten erweitern den

55; *Intveen*, ITRB (2020), 70, 73; *Reusch*, K&R Beilage (2019), zu Heft 7/8, 20, 21; *Spindler*, CR (2015), 766, 775 f.; *Zech*, ZfPW (2019), 198, 216 f.

⁷¹ Zur Verletzung von Persönlichkeitsrechten durch Diskriminierung ausführlicher etwa *Kolleck/Orwat* (Oktober 2020), Mögliche Diskriminierung durch algorithmische Entscheidungssysteme und maschinelles Lernen – ein Überblick, S. 41 f.

⁷² Dazu *Wagner*, VersR (2020), 717, 730; *Denga*, CR (2018), 69, 74 f.; *Keßler*, MMR (2017), 589, 593; *Hacker*, RW (2018), 243, 265 ff.

⁷³ Umstritten, wie hier *Däubler/Bertzbach/Deinert*, Allgemeines Gleichbehandlungsgesetz, 4. Aufl. (2018), § 21 AGG, Rn. 100; a.A. *MüKo BGB/Thüsing*, 8. Aufl. (2018), § 21 AGG, Rn. 72; jeweils mwN.

⁷⁴ *Präve*, VersR (2009), 857, 862.

⁷⁵ *Specht/Mantz/Mantz/Marosi*, Handbuch europäisches und deutsches Datenschutzrecht (2019), § 3 Vorgaben der Datenschutz-Grundverordnung, Rn. 1.

⁷⁶ In diesem Sinne auch *Bilski/Schmid*, NJOZ (2019), 657, 659.

⁷⁷ *Kühling/Buchner/Buchner*, Bundesdatenschutzgesetz, 3. Aufl. (2020), Art. 22 DSGVO, Rn. 31.

Pflichtenkreis aus §§ 6, 61 VVG beim Vertrieb von KI-basierten Versicherungsprodukten.

Auch außerhalb von Art. 22 Abs. 4 DSGVO sind automatisierte Entscheidungen mit rechtlichen (oder ähnlich beeinträchtigenden) Wirkungen nach Art. 22 Abs. 1 DSGVO grundsätzlich unzulässig. Hiervon gibt es Ausnahmen, die für die Versicherer im Umgang mit ihren Versicherungsnehmern relevant werden können⁷⁸ oder sogar speziell für diesen Fall – konkret für die Leistungserbringung durch Versicherer – formuliert wurden (Art. 22 Abs. 2, 3 DSGVO, § 37 BDSG⁷⁹). Das grundsätzliche Verbot des Art. 22 Abs. 1 DSGVO zielt auch darauf ab, diskriminierende Wirkungen der Datenverarbeitung zu verhindern (s. Erwägungsgrund 71 S. 6 DSGVO).⁸⁰ Zudem stellen die Transparenzpflichten nach Art. 13 Abs. 2 lit. f, Art. 14 Abs. 2 lit. g DSGVO sowie der Auskunftsanspruch nach Art. 15 Abs. 1 lit. h DSGVO sicher, dass die Kriterien, aufgrund derer automatisiert entschieden wird, sowie die Tatsache der automatisierten Entscheidung für den Vertragspartner erkennbar sind. Eine Offenlegung des Algorithmus kann allerdings nicht verlangt werden (s. auch Erwägungsgrund 63).⁸¹

Sofern für die Datenverarbeitung gem. Art. 6 Abs. 1 S. 1 lit. a DSGVO eine Einwilligung notwendig ist, muss der betroffene Versicherungsnehmer über die Zwecke der Verarbeitung der Daten so informiert werden, dass er wirksam einwilligen kann (Art. 4 Nr. 11, Art. 5 Abs. 1 lit. b). Bei selbstlernender KI kann es schwierig oder sogar unmöglich sein, den Betroffenen über die Zwecke der Verarbeitung der Daten durch die KI so zu informieren, dass er wirksam in die Datenverarbeitung einwilligen kann. Die KI muss über die Zwecke der Verarbeitung womöglich selbständige Entscheidungen treffen, sodass sich die Zwecke erst im Laufe der Datenverarbeitung herauskristallisieren.⁸² KI in diesen Möglichkeiten zu beschränken, eben um eine wirksame Einwilligung zu ermöglichen, kann auch einen Bias verursachen, weil eine nicht repräsentative Datenbasis verwendet wird.⁸³

Versicherer haben zudem gem. Art. 35 Abs. 1 DSGVO vorab, also bevor sie die KI als neue Technologie einsetzen, eine Datenschutzfolgeabschätzung vorzunehmen, wenn die Datenverarbeitung „ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen zur Folge“ hat. Ein solches Risiko ist nach Erwägungsgrund 75 (s. auch 85) die Diskriminierung. Art. 35 Abs. 3 lit. a DSGVO erklärt eine solche Folgeabschätzung für erforderlich, wenn persönliche Aspekte natürlicher Personen systematisch und umfassend bewertet werden und Grundlage einer Entscheidung mit Rechtswirkung oder ähnlich beeinträchtigenden Folgen sind; hier ist insbeson-

⁷⁸ So nach Art. 22 Abs. 1 lit. a DSGVO, wenn sie für den Vertragsabschluss oder die Vertragserfüllung erforderlich sind; s. dazu Art. 13 Abs. 2 Verhaltensregeln für den Umgang mit personenbezogenen Daten durch die deutsche Versicherungswirtschaft; kritisch im Hinblick auf die Erforderlichkeit des KI-Einsatzes beim Vertragsabschluss und bei der Leistungsbearbeitung *Lüttringhaus*, VW (2020), Heft 6, 88 f.

⁷⁹ Näher dazu *Lüttringhaus*, VW (2020), Heft 6, 88 f.

⁸⁰ *Allen/Masters*, ERA Forum 20 (2020), 585, 597.

⁸¹ Artikel-29-Datenschutzgruppe, 03.10.2017, WP251rev.01, Leitlinien zu automatisierten Entscheidungen im Einzelfall einschließlich Profiling für die Zwecke der Verordnung 2016/679, S. 28; *Allen/Masters*, ERA Forum 20 (2020), 585, 596 f.; *Gausling*, ZD (2019), 335, 340 f.

⁸² *Walter*, K&R Beilage (2019), zu Heft 7/8, 21, 23.

⁸³ *Walter*, K&R Beilage (2019), zu Heft 7/8, 21, 23.

dere die automatisierte Entscheidung nach Art. 22 DSGVO erfasst.⁸⁴ Die in diesem Beitrag angesprochenen Einsätze von KI durch den Versicherer im Verhältnis zum Versicherungsnehmer werden nach dem Gesagten häufig eine Folgenabschätzung erfordern.⁸⁵

Verstößt ein Versicherer gegen diese Vorgaben, drohen neben Geldbußen nach Art. 83 Abs. 4 lit. a, Abs. 5 lit. b DSGVO und verwaltungsrechtlichem behördlichem Einschreiten auch zivilrechtliche Schadensersatzansprüche geschädigter Versicherungsnehmer nach Art. 82 DSGVO. Auch hier stellen sich Zurechnungsfragen, die im Zivilrecht parallel zu dem oben unter Abschn. 3.5 Gesagten zu beantworten sind.

4 Technische Möglichkeiten, einem Bias entgegenzuwirken (Bias-Management)

Algorithmische Bias-Effekte können nicht vollständig vermieden, aber durch den Einsatz technischer Möglichkeiten minimiert werden. Die stetig wachsende Ausbreitung der Digitalisierung, der vermehrte Einsatz von Künstlicher Intelligenz und die damit drastisch ansteigende Relevanz von Bias beflügeln die Entwicklung fairer KI-Verfahren, welche sich in drei Kategorien unterteilen lassen und nachfolgend vorgestellt werden.⁸⁶ Diese Dreiteilung bezieht sich auf den Zeitpunkt der Korrektur:

- Wie muss ich die Daten ändern, damit kein Bias gelernt wird?
- Wie muss ich meine Zielfunktion oder Nebenbedingungen anpassen, damit kein Bias gelernt wird?
- Wie muss ich meine Ergebnisse anpassen, damit der bereits gelernte Bias entfernt wird?

Als erste Kategorie können vorbereitende Algorithmen genutzt werden. Grundidee dieser Vorgehensweise ist es, dass in der Datengrundlage zwischen einer Reihe von sensitiven Attributen (z. B. Religion, Alter oder Staatsangehörigkeit) und einem vorherzusagenden Wert ein Muster beobachtet werden kann (z. B. niedrigere Annahmequote bei einer Minderheit). Damit dieses unerwünschte Muster nicht von einem Algorithmus erkannt und repliziert wird, zielen die vorbereitenden Algorithmen darauf, ein Trainingsdatenset ohne dieses unerwünschte Muster zu erstellen.⁸⁷ Hierdurch soll eine Chancengleichheit gewährleistet und Bias-Effekten präventiv entgegengewirkt werden, da bereits vor dem Lernprozess eine unverzerrte bzw. faire Datenrepräsentation erzeugt wird.

⁸⁴ BeckOK Datenschutzrecht/Hansen (2020), Stand: 01.08.2020, Art. 35 DS-GVO, Rn. 26; s. auch die Positivliste gem. Art. 35 Abs. 4 DSGVO.

⁸⁵ Vgl. Gausting, ZD (2019), 335, 337 für Chatbots und digitale Assistenten, die mit dem Kunden kommunizieren.

⁸⁶ Friedler/Scheidegger/Venkatasubramanian/Choudhary/Hamilton/Roth, in FAT* '19: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (2019, S. 329f.)

⁸⁷ Friedler/Scheidegger/Venkatasubramanian/Choudhary/Hamilton/Roth, in FAT* '19: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (2019, S. 329, 330).

Feldman, et al. greifen dieses Konzept auf und messen mithilfe einer balancierten Fehlerrate, ob innerhalb eines Datensets $D = (X, Y, C)$ die Ausprägungen der sensitiven Attribute X unterschiedliche Auswirkungen auf das jeweilige KI-Ergebnis C haben. Sofern solch ein verzerrter mathematischer Zusammenhang festgestellt wird, kann ein vorbereitender Algorithmus als Reparaturmechanismus zum Neutralisieren dieser Bias-Effekte eingesetzt werden. Diese Korrektur erfolgt über Adjustieren der nicht-sensitiven Attribute Y . Die Stärke der Korrektur kann dabei über einen Faktor λ gesteuert werden, sodass neben dem vollständigen Entfernen ($\lambda = 1$) und dem Beibehalten der Daten ($\lambda = 0$) auch das Abschwächen von Bias-Effekten ($0 < \lambda < 1$) möglich ist. Durch diese Maßnahme wird ein neues Trainingsset $\overline{D}_\lambda = (X, \overline{Y}, C)$ generiert, welches die verzerrten Muster nicht mehr oder in einer abgeschwächten Form beinhaltet und somit für eine KI im nachgelagerten Lernprozess aufbereitet ist.⁸⁸

Die Wirkungsweise des Reparaturmechanismus wird anhand eines Beispiels in der Abb. 2 verdeutlicht. In dem linken Teil wird die ursprüngliche Datenverteilung zweier Gruppen gezeigt. Verglichen werden eine privilegierte (orange) und eine benachteiligte Gruppe (blau), die unterschiedlich hohe Ausprägungen hinsichtlich einer algorithmischen Entscheidung ausweisen. Der rechte Teil der Abb. 2 zeigt hingegen einen reparierten Datensatz, in dem keine Gruppendifferenzierung anhand des Entscheidungsausgangs mehr beobachtet werden kann, da die beiden Verteilungen sich überschneiden. Der dargestellte Bias-Effekt wurde rechnerisch vollständig neutralisiert. In der Mitte wird ein Trade-off-Ansatz zwischen diesen Extremen dargestellt. Es ist zu erkennen, dass sich die Verteilungen der beiden Gruppen nicht vollständig überschneiden, sondern lediglich angenähert haben. Die Auswirkung des Bias-Effekts wurde somit abgeschwächt bzw. partiell entfernt.⁸⁹

Als zweite Kategorie zur Vermeidung von Bias-Effekten können im Rahmen von algorithmischen Modifikationen geeignete Regularisierungsterme in das Lernverfahren eingearbeitet werden. Durch die Nutzung von zusätzlichen Nebenbedingungen oder durch die Einbindung von additiven Kostenfunktionen kann das Lernverfahren in eine faire Form gesteuert werden.⁹⁰ Eine solche Vorgehensweise wurde etwa von Kamishima et al. eingesetzt.⁹¹ Die von ihnen eingesetzte algorithmische Regulierung erzwingt eine Unabhängigkeit von sensitiven Informationen. Anhand von verschiedenen Datensets zeigen sie, dass ihre Vorgehensweise zur einer Beseitigung von Vorurteilen in Daten beiträgt.⁹²

⁸⁸ Feldman/Friedler/Moeller/Scheidegger/Venkatasubramanian, in KDD '15: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2015, S. 259 ff.).

⁸⁹ Ronaghan, AI Fairness—Explanation of Disparate Impact Remover, 22.04.2019, abrufbar unter: <https://towardsdatascience.com/ai-fairness-explanation-of-disparate-impact-remover-ce0da59451f1> [Stand: 26.07.2021].

⁹⁰ Friedler/Scheidegger/Venkatasubramanian/Choudhary/Hamilton/Roth, in FAT* '19: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (2019, S. 329, 330).

⁹¹ Kamishima/Akaho/Asoh/Sakuma, in Flach/Bie/Christianini (Hrsg.), Machine Learning and Knowledge Discovery in Databases (2012, S. 35 ff.).

⁹² Kamishima/Akaho/Asoh/Sakuma, in Flach/Bie/Christianini (Hrsg.), Machine Learning and Knowledge Discovery in Databases (2012, S. 35 ff.).

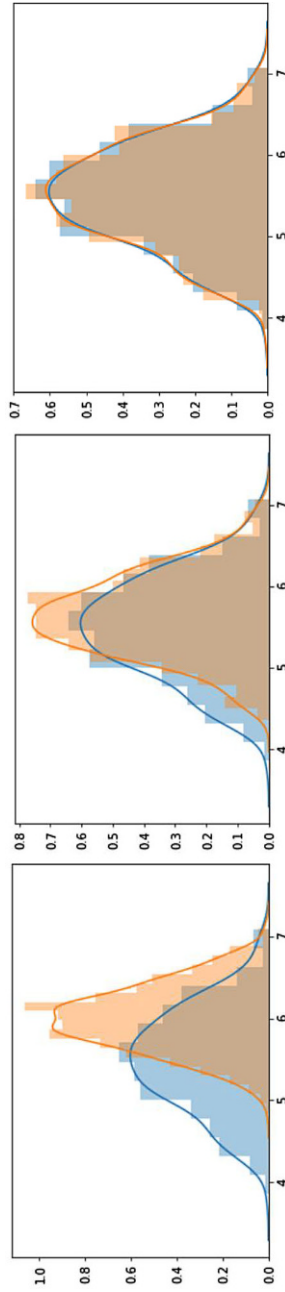


Abb. 2 Wirkungsweise des Reparaturmechanismus nach Feldman et al.; Quelle: In Anlehnung an Ronaghan (Fn. 89)

Als dritte Kategorie können nachbereitende Techniken genutzt werden, welche das Pendant zur ersten Variante darstellen, da diese Techniken nach der KI-Anwendung zum nachträglichen Adjustieren von ε -unfairen Outputs eingesetzt werden. Folglich werden hiermit Ergebnisse einer bestehenden KI-Lösung auf Bias-Effekte untersucht und nachträglich eine faire Fehlerverteilung sichergestellt.⁹³ Lohia et al. realisieren eine solche nachbereitende Technik, indem sie zuerst die individuelle Bias-Erkennung für ein bereits gelerntes KI-Modell durchführen und nachfolgend einen Nachbearbeitungsalgorithmus zur Bias-Minderung einsetzen. Hierzu nutzen sie das Konzept der Gruppenfairness und individuellen Fairness zur Ableitung eines Bias-Scores, welcher zur Bewertung der Klassifikationsergebnisse eingesetzt wird. Die anschließende Bias-Korrektur setzt eine einfache Abgleichlogik ein, welche die ursprüngliche Klassifikation der KI durch einen plausibleren Wert ersetzt, wenn der Score eine hohe Bias-Wahrscheinlichkeit suggeriert.⁹⁴

Jede dieser drei Kategorien weist Vor- und Nachteile auf. Vorbereitende Algorithmen können mit jeder Art von Klassifikationsalgorithmus verwendet werden. Gleichzeitig können sie jedoch die Erklärbarkeit der Ergebnisse beeinträchtigen, da sie die algorithmischen Eigenschaften in nachfolgenden Instanzen (etwa dem Lernen) nicht berücksichtigen und hierdurch die Auswirkungen auf die KI-Genauigkeit nicht abschätzbar ist. Auch nachbereitende Techniken sind mit jedem Klassifikationsalgorithmus verwendbar und kommen naturgemäß erst nach dem Lernen der KI zum Einsatz, wodurch typischerweise schlechtere Korrektur-Resultate als in den anderen Kategorien beobachtet werden. Allerdings bieten sie eine vergleichsweise einfache Möglichkeit, um spezielle Arten von Bias vollständig zu entfernen. Die Kategorie der algorithmischen Modifikationen bietet einen Trade-Off zwischen KI-Genauigkeit und der Fairness der Fehlerverteilung an, da sie direkt mit der Zielfunktion interagieren. Hierdurch entstehen im Lernprozess auch zusätzliche Abhängigkeiten, die einen Anstieg der algorithmischen Komplexität mit sich ziehen.⁹⁵

Die ausgewählten Beispiele des Bias-Managements bzw. der algorithmischen Fairness decken das breite Feld nicht in Gänze ab, da die zunehmende Relevanz der Thematik und der damit verbundene Anstieg des Forschungsinteresses zu vielen alternativen Umsetzungsmöglichkeiten in den jeweiligen Kategorien führt. Für eine tiefergehende Auseinandersetzung sei daher auf Forschungsplattformen und -communities verwiesen, die den Stand von Fairness-Algorithmen fortlaufend evaluieren und aggregieren. Solche Einrichtungen bieten häufig nicht nur zusätzliche Informationen, sondern auch qualitätsgesicherte Implementierungen dieser Algorithmen an und mindern hierdurch die Eintrittsbarrieren zur Nutzung. Das Alan Turing Institute⁹⁶, das interdisziplinäre Leverhulme Centre for the Future of Intelligence⁹⁷

⁹³ Friedler/Scheidegger/Venkatasubramanian/Choudhary/Hamilton/Roth, in FAT* '19: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency (2019, S. 329 ff.).

⁹⁴ Lohia/Natesan Ramamurthy/Bhide/Sahal/Varshney/Puri, in ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2019, S. 2847 ff.).

⁹⁵ Pessach/Shmueli, Algorithmic fairness (2020, S. 10).

⁹⁶ Siehe <https://www.turing.ac.uk/research/interest-groups/fairness-transparency-privacy> [Stand: 26.07.2021].

⁹⁷ Siehe <http://lcfi.ac.uk/projects/> [Stand: 26.07.2021].

der Cambridge Universität oder das fachbereichsübergreifende Berkman Klein Center for Internet & Society⁹⁸ an der Harvard Universität sind renommierte wissenschaftliche Beispiele. Es gibt jedoch auch eine Reihe an privatwirtschaftlichen Forschungsstellen, die ihre Ergebnisse frei im Internet anbieten. IBM Research betreibt etwa das Open Source-Toolkit *Fairness 360* zur Erkennung und Vermeidung von Bias-Effekten wie Diskriminierung entlang des gesamten KI-Lebenszyklusses.⁹⁹ Pymetrics hat eine Open-Source-Applikation zur Überprüfung von Bias im Rahmen von maschinellem Lernen entwickelt und auf GitHub öffentlich zugänglich gemacht.¹⁰⁰ Microsoft bietet das Open Source-Toolkit *Fairlearn* an.¹⁰¹ Google finanziert eine Forschungsgruppe¹⁰², deren Arbeiten regelmäßig auf führenden Konferenzen, wie der ACM FAccT Conference¹⁰³, erscheinen und ebenfalls freizugänglich anbietet.

Als Resümee kann festgehalten werden, dass der Stand des Bias-Managements soweit ausgereift ist, dass er die Nutzung der attraktiven KI-Vorteile unter zeitgleicher Bias-Minimierung ermöglicht. Bias im Rahmen von KI kann jedoch niemals gänzlich ausgeschlossen werden, da die einzelnen Verfahren des Bias-Managements Korrekturentscheidungen unter der Berücksichtigung von Wahrscheinlichkeiten treffen und zum Erreichen der bestmöglichen Resultate an das jeweilige Analyseszenario adaptiert werden. Es gibt nicht „den einen Bias“, welcher durch die Anwendung eines festen Schemas vermieden werden kann, sondern eine Vielzahl unterschiedlicher Bias-Formen, die durch den gezielten Einsatz von geeigneten Verfahren reduziert werden sollten. In diesem Kontext lassen sich Analogien zum Risiko-Management ziehen, da ebenso hier das unternehmerische Risiko nie vollkommen eliminiert werden kann, sondern durch geeignete Maßnahmen erkannt, gesteuert und überwacht werden soll. Folglich gilt es ein geeignetes Konzept im Umgang mit KI-Bias zu finden und hierdurch einen fahrlässigen Einsatz von KI auszuschließen.

5 Das Versicherungsaufsichtsrecht als Schutz vor Bias

Nach der Betrachtung der technischen Möglichkeiten der Versicherungsunternehmen, einem Bias entgegenzuwirken, ist zuletzt die Frage zu stellen, welche Vorgaben das Versicherungsaufsichtsrecht für die Verhinderung von Diskriminierung und damit, bezogen auf KI, für das Bias-Management macht.

Das Versicherungsaufsichtsrecht unterwirft Versicherungsunternehmen einer strengen Regulierung, um die Versicherungsnehmer und die Begünstigten von Versicherungsleistungen zu schützen (§ 294 Abs. 1 VAG). Es verlangt von den Versicherungsunternehmen eine wirksame, angemessene und ordnungsgemäße Geschäftsorganisation (Governance, § 23 Abs. 1 VAG). Sie ist eine der zentralen Vor-

⁹⁸ Siehe <https://cyber.harvard.edu/topics> [Stand: 26.07.2021].

⁹⁹ Siehe <https://aif360.mybluemix.net/> [Stand: 26.07.2021].

¹⁰⁰ Siehe <https://github.com/pymetrics/audit-ai> [Stand: 26.07.2021].

¹⁰¹ Siehe https://www.microsoft.com/en-us/research/uploads/prod/2020/05/Fairlearn_whitepaper.pdf [Stand: 26.07.2021].

¹⁰² Siehe <https://ai.google/responsibilities/responsible-ai-practices/?category=fairness> [Stand: 26.07.2021].

¹⁰³ Siehe <https://facctconference.org/> [Stand: 26.07.2021].

gaben, die das Versicherungsaufsichtsrecht den Versicherungsunternehmen macht. Soll eine Technologie wie die KI genutzt werden, ist die in der Generalklausel des § 23 Abs. 1 VAG verlangte Governance für diesen speziellen Fall zu konkretisieren. So betont auch die BaFin, dass die Nutzung von KI in die Governance eingebettet sein muss.¹⁰⁴ Was das im Einzelnen bedeutet, wie also die Governance bezüglich KI gestaltet sein muss, wird z.T. durch weitere gesetzliche Vorgaben, z.T. durch aufsichtsbehördliche Leitlinien sowie (selten) durch die Rechtsprechung konkretisiert. So muss das Geschäft z.B. so organisiert sein, dass die technische Robustheit und Sicherheit von KI-Systemen gewährleistet ist, und es muss insbesondere auch die Rechtmäßigkeit der KI sichergestellt sein.¹⁰⁵ Letzteres umfasst die Aufgabe, zu verhindern, dass ein Bias der KI Verstöße gegen die oben genannten Diskriminierungsverbote herbeiführt.¹⁰⁶ Welche Anforderungen das Versicherungsaufsichtsrecht unter diesem Aspekt stellt, wird zunächst untersucht (Abschn. 5.1). Im Anschluss wird erörtert, wie die Versicherungsaufsichtsbehörden Verstöße des Versicherungsunternehmens gegen die oben (Abschn. 3.) genannten Normen sowie gegen die KI-spezifischen Governancevorgaben verhindern und sanktionieren können (Abschn. 5.2).

5.1 Vorgaben für die Geschäftsorganisation

5.1.1 Allgemeine Governance-Vorgaben

§ 23 Abs. 1 S. 1 VAG bildet als „große“ Generalklausel den normativen Ausgangspunkt der Anforderungen an die Geschäftsorganisation. Letztere muss wirksam, ordnungsgemäß und für die Art, den Umfang und die Komplexität der Unternehmens-tätigkeit angemessen sein. Ziel der Geschäftsorganisation ist es nach § 23 Abs. 1 S. 2 VAG unter anderem, sicherzustellen, dass die Unternehmen die einschlägigen Gesetze, Verordnungen mit aufsichtsbehördlichen Anforderungen einhalten. Dazu gehören nach zutreffender Ansicht nicht nur die für den Betrieb des Versicherungsgeschäfts geltenden Normen (anders aber bei § 298 Abs. 1 S. 1 und 2 i. V.m. § 294 Abs. 2 S. 2 Fall 1 VAG, s. unten Abschn. 5.2.1), sondern all diejenigen, deren Nichteinhaltung die Belange der Versicherten gefährden kann.¹⁰⁷ Die oben genannten Diskriminierungsverbote und Datenschutzvorschriften zu respektieren, fällt also unter das Ziel der Geschäftsorganisation.

Innerhalb der Geschäftsorganisation ist eine Compliance-Funktion einzurichten, die die Einhaltung der einschlägigen Normen sicherstellen muss (§ 29 Abs. 2 VAG). Konkrete Vorgaben für den Umgang mit dem Diskriminierungsrisiko, das einer KI innewohnt, finden sich hier nicht; auch in Art. 258, 270 VO (EU) 2015/35

¹⁰⁴ BaFin, Big Data trifft auf künstliche Intelligenz (2018, S. 175).

¹⁰⁵ Zu weiteren Anforderungen s. Bartuschka, CB (2019), 340, 344.

¹⁰⁶ BaFin, Big Data trifft auf künstliche Intelligenz (2018, S. 181) (allgemein zu der Anforderung, beim Einsatz von Algorithmen Diskriminierung zu vermeiden).

¹⁰⁷ Näher Kaulbach/Bähr/Pohlmann/Pohlmann, Versicherungsaufsichtsgesetz, 6. Aufl. (2019), § 23 VAG, Rn. 36; a. A. Bürkle, in Bähr (Hrsg.), Handbuch des Versicherungsaufsichtsrechts (2011), S. 181, § 9 Rn. 14.

(Solvabilität II-VO) wird dieser Punkt nicht näher ausgeführt, ebenso wenig in den entsprechenden Leitlinien der EIOPA¹⁰⁸ oder in den Mindestanforderungen für die Geschäftsorganisation (MaGo) der BaFin.¹⁰⁹

Auch die Versicherungsaufsichtlichen Anforderungen an die IT (VAIT)¹¹⁰ der BaFin verhalten sich zu diesem Punkt nicht direkt. Sie verlangen aber eine IT-Strategie. In ihren jüngeren „Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen“ hat die BaFin die aufsichtsrechtlichen Mindestanforderungen für den Einsatz von KI vorläufig konkretisiert und auch die Verhinderung von Diskriminierung vorgegeben.¹¹¹ Da KI ein Fall von IT im Sinne der VAIT ist,¹¹² ist auch für diese eine „konsistente“ Strategie festzulegen (VAIT II. 1. 1), die sich nicht nur auf die eigene IT, sondern auch auf Ausgliederungen, Dienstleistungsbeziehungen und den Bezug von IT (Hard- und Software) erstreckt. Diese Strategie ist der Aufsichtsbehörde zur Kenntnis zu geben (VAIT II. 1. 4). Auch eine IT-Governance und ein IT-Risikomanagement sind zu implementieren (VAIT II. 2., 3.). Davon ist insbesondere die technische Seite der Verhinderung von Bias erfasst.¹¹³ Bei der Konzeption der Strategie, der Governance und des Risikomanagements müssen für den kundenbezogenen KI-Einsatz die oben beispielhaft dargestellten Bias-Gefahren (durch fehlerhafte oder fehlende Trainingsdaten, unpassende algorithmische Zielsetzungen oder durch Proxy-Attribute) jeweils gezielt adressiert werden. Speziell in der Phase der Entwicklung einer KI ist auf die Reproduzierbarkeit der Ergebnisse, Dokumentation und Validierung zu achten.¹¹⁴

Aus den Governancevorgaben kann auch die Pflicht der Unternehmensleiter und der Inhaber von Schlüsselfunktionen folgen, ihrerseits KI einzusetzen. Zu der letztgenannten Gruppe gehört auch der für die Beachtung der Diskriminierungsverbote zuständige Compliance-Beauftragte. Ist etwa eine vom Unternehmen verwendete KI im Hinblick auf mögliche Diskriminierung zu kontrollieren, wird sich im Massengeschäft der Versicherung häufig nur mithilfe von KI prüfen lassen, ob ein Bias vorliegt, der zu einer verbotenen Diskriminierung führt. Die Funktionsweise der KI

¹⁰⁸ EIOPA, Leitlinien zum Governance-System (EIOPA-BoS 14/253), abrufbar unter: https://www.eiopa.europa.eu/content/guidelines-system-governance_en [Stand: 26.07.2021]; Explanatory Texts zu diesen Leitlinien (abgedruckt in: EIOPA, Final Report on Public Consultation No. 14/017 on Guidelines on system of governance, abrufbar unter: https://register.eiopa.europa.eu/Publications/Consultations/EIOPA_EIOPA-BoS-14-253-Final%20report_Governance.pdf [Stand: 26.07.2021]).

¹⁰⁹ BaFin, Rundschreiben 2/2017 (VA) – Mindestanforderungen an die Geschäftsorganisation von Versicherungsunternehmen (MaGo), abrufbar unter: https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Rundschreiben/2017/rs_1702_mago_va.html [Stand: 26.07.2021].

¹¹⁰ BaFin, Versicherungsaufsichtliche Anforderungen an die IT (VAIT), abrufbar unter: https://www.bafin.de/SharedDocs/Downloads/DE/Rundschreiben/dl_rs_1810_vait_va.pdf?__blob=publicationFile&v=5 [Stand: 26.07.2021].

¹¹¹ BaFin (2021), Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen, Prinzipienpapier BDAI, 15.06.2021, S. 8, 11.

¹¹² Vgl. BaFin (2021), Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen, Prinzipienpapier BDAI, 15.06.2021, S. 6.

¹¹³ S. auch BaFin (2021), Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen, Prinzipienpapier BDAI, 15.06.2021, S. 8.

¹¹⁴ BaFin (2021), Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen, Prinzipienpapier BDAI, 15.06.2021, S. 9.

kann allein mit „natürlicher“ Intelligenz u. U. nicht überprüft werden.¹¹⁵ Eine angemessene Governance setzt dann voraus, dass sie sich der KI zur Unterstützung bedient. Hiervon ist nicht nur der Einsatz von nicht-diskriminierender KI betroffen. Auch andere Vorgänge im Unternehmen lassen sich nur mithilfe bestimmter IT-Anwendungen, darunter auch KI, überprüfen, sodass dann eine Pflicht zum Einsatz dieser IT besteht;¹¹⁶ hierfür ist der Begriff der „Corporate Technology (CorpTech)“ geprägt worden.¹¹⁷

Adressat der aufsichtsrechtlichen Governancepflichten ist das Versicherungsunternehmen. Entwickelt eine selbstlernende KI einen Bias, der im Einzelnen nicht, generell aber voraussehbar ist, und verstößt das Ergebnis gegen aufsichtsrechtliche Pflichten des Versicherungsunternehmens, stellen sich ebenfalls Zurechnungsfragen. Unmittelbar leuchtet ein, wenn die BaFin betont, dass die Verantwortung für automatisierte, auf Big Data und KI gestützte Prozesse bei der Geschäftsleitung bleibe;¹¹⁸ die KI selbst kann de lege lata nicht selbst verantwortlich sein. Diese Verantwortung bringt es mit sich, dass ein Unternehmen sich nicht darauf berufen kann, die Entscheidungsfindung durch eine KI sei nicht nachvollziehbar. In den Worten der BaFin: „Blackbox-Ausreden sind unzulässig“.¹¹⁹ Wie im Zivilrecht (Abschn. 3.5) wird man einem Versicherer, der KI nutzt, das Bias-Risiko auch aufsichtsrechtlich zurechnen können. Es ist für ihn voraussehbar und – durch Verzicht auf KI – vermeidbar.

Art. 32 DSGVO macht im Übrigen ebenfalls Vorgaben für das Risikomanagement im Hinblick auf die datenschutzrechtlichen Gefahren, zu denen, wie dargelegt, auch die Diskriminierung bei automatisierten Entscheidungen gehört. Hier überschneiden sich die beiden Schutzregime des VAG und der DSGVO und damit auch die Behördenzuständigkeiten.

5.1.2 Vorgaben für die Produktgovernance

§ 23 Abs. 1a bis 1c VAG normieren spezifische Governancepflichten des Versicherers für die Produktfreigabe, -überprüfung und die Information über Produkte im Vertrieb (zusammenfassend hier als Produktgovernance bezeichnet). Diese Pflichten sind keine zivilrechtlichen Pflichten gegenüber den Versicherungsnehmern, sondern aufsichtsrechtliche Vorgaben.¹²⁰

Das Produktfreigabeverfahren nach Abs. 1a muss alle Risiken des Produktes bewerten. Dazu gehören bei Produkten, die mithilfe von KI vertrieben werden oder in der Nutzung KI einsetzen, wie möglicherweise Pay-as-you-live-Produkte, die spezifischen KI-Risiken, also auch das Risiko der Diskriminierung. Sind Produkte zum Vertrieb freigegeben, müssen sie weiterhin auf diese Risiken regelmäßig überprüft

¹¹⁵ Dazu Noack, ZHR 183 (2019), 105, 118 f.

¹¹⁶ Lücke, BB (2019), 1986, 1989.

¹¹⁷ Z. B. Graewe, BB (2020), Heft 11, I.

¹¹⁸ BaFin, Big Data trifft auf künstliche Intelligenz (2018, S. 175); BaFin (2021), Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen, Prinzipienpapier BDAI, 15.06.2021, S. 6.

¹¹⁹ BaFin, Big Data trifft auf künstliche Intelligenz (2018, S. 175).

¹²⁰ Begr. RegE BT-Drs. 18/11627, S. 39.

werden, Abs. 1b. Schließlich verlangt Abs. 1c, dass alle Vertreter des Produkts über sämtliche sachgerechten Informationen über das Produkt verfügen. Dazu gehören z. B. die Informationen, die dem Versicherungsnehmer nach der DSGVO über die Funktionsweise einer automatisierten Entscheidung weiterzugeben sind (s. oben Abschn. 3.7).

5.1.3 Anforderungen an die Geschäftsleiter und Inhaber von Schlüsselfunktionen

§ 24 Abs. 1, 2 VAG verlangt bei Geschäftsleitern und sonstigen tatsächlichen Leitern u. a. die auf die konkrete Aufgabe bezogene (s. Art. 273 Abs. 3 Solvabilität II-VO) fachliche Eignung. Sie umfasst nach dem Merkblatt der BaFin zwar nicht ausdrücklich Qualifikationen, Kenntnisse und Erfahrungen im Bereich der IT, speziell der KI.¹²¹ IT-Kenntnisse theoretischer und praktischer Art werden aber von demjenigen erwartet, der für einen Versicherer im Vorstand das IT-Ressort leitet.¹²² Zudem müssen alle Mitarbeiter über die ihren Aufgaben entsprechenden IT-Kenntnisse verfügen.¹²³ Wenn ein Unternehmen Produkte entwickelt und vertreibt, die im Verhältnis zum Kunden KI einsetzen, setzt die von der BaFin ausdrücklich verlangte Kenntnis des Geschäftsmodells voraus, dass man mit den Diskriminierungsrisiken KI-basierter Produkte vertraut ist. Entsprechende IT-Kenntnisse sind allerdings nicht unbedingt zu verlangen. Immerhin aber erleichtert es die BaFin solchen Personen, die über IT-Kenntnisse verfügen, für Versicherungsunternehmen als Geschäftsleiter tätig zu werden, indem sie die Anforderungen an die Erfahrungen in Versicherungsgeschäften gesenkt hat.¹²⁴

5.1.4 Ausgliederung

Versicherungsunternehmen, die KI einsetzen, gliedern bestimmte Aufgaben in diesem Zusammenhang nicht selten auf andere Unternehmen, oftmals Tochtergesellschaften, aus. Sie müssen dann sicherstellen, dass die aufsichtsrechtlichen Vorgaben für das Bias-Management auch bei dem Unternehmen, auf das die Aufgabe übertragen wurde, gewahrt werden.

Eine solche Ausgliederung kann praktische, also z. B. betriebswirtschaftliche oder organisatorische, Gründe haben. Aber auch rechtliche Vorgaben können dazu führen, dass KI-Anwendungen ausgegliedert werden. Ein Beispiel dafür sind die Pay-as-you-Drive-Tarife in der KFZ-Versicherung. Würde der Versicherer selbst die Fahrdaten verarbeiten, könnte er sie mit den Personendaten des Versicherungsnehmers verknüpfen und so ein Bewegungsprofil des Versicherungsnehmers erstellen. Die

¹²¹ BaFin, VA Geschäftsleiter-Merkblatt, S. 14, abrufbar unter: https://www.bafin.de/SharedDocs/Downloads/DE/Konsultation/2016/dl_kon_0416_merkblatt_geschaeftsleiter_va.pdf?__blob=publicationFile&v=1 [Stand: 26.07.2021].

¹²² BaFin, BaFin Journal (2017), Heft 12, 15.

¹²³ BaFin, Versicherungsaufsichtliche Anforderungen an die IT (VAIT), II. 2. 10, abrufbar unter: https://www.bafin.de/SharedDocs/Downloads/DE/Rundschreiben/dl_rs_1810_vait_va.pdf?__blob=publicationFile&v=5 [Stand: 26.07.2021].

¹²⁴ BaFin, BaFin Journal (2017), Heft 12, 15.

Empfehlung des Datenschutzbeauftragten NRW lautet daher, die Datenkreise nicht zusammenzuführen.¹²⁵ Datenschutzrechtlich ist nach alldem erforderlich, die Verarbeitung der Fahrdaten auf eine unabhängige Einheit, i. d. R. eine Tochtergesellschaft, auszugliedern. So werden zwei getrennte Datenkreise geschaffen.

Nach § 47 Nr. 8 VAG muss der Versicherer solche Ausgliederungen anzeigen. Weitere Details der Ausgliederung regeln Art. 274 VO (EU) 2015/35 und die einschlägigen Leitlinien der BaFin¹²⁶ und der EIOPA.¹²⁷ Nach § 32 Abs. 1 VAG bleibt der Versicherer auch bei der Ausgliederung von Funktionen für die Erfüllung aller aufsichtsrechtlicher Vorschriften und Anforderungen, wie sie oben beschrieben wurden, verantwortlich. Das bedeutet zweierlei: Das Versicherungsunternehmen muss dafür sorgen, dass die Anforderungen eingehalten werden, und es kann bei Nichtbeachtung durch die Aufsicht in Anspruch genommen werden. Alle Anforderungen, die es beim Bias-Management zu erfüllen hat, muss es also, um seine Pflichten zu erfüllen, an den KI-Anbieter „weiterreichen“.¹²⁸ Ferner muss es den aufsichtsbehördlichen Zugriff auf das eingeschaltete Unternehmen vertraglich sicherstellen.

Das Datenschutzrecht stellt zudem spezifische Anforderungen an das Outsourcing, die an dieser Stelle nicht vertieft werden können.¹²⁹ Auch hier kommt es also zu sich überschneidenden Behördenzuständigkeiten.

5.2 Aufsichtsbehördliche Maßnahmen

5.2.1 Gesetze für den Betrieb des Versicherungsgeschäfts

Nach § 298 Abs. 1 S. 1 und 2 i. V. m. § 294 Abs. 2 S. 2 Fall 1 VAG kann die Aufsichtsbehörde alle Maßnahmen ergreifen, die geeignet und erforderlich sind,

¹²⁵ Empfehlung im 22. Datenschutz- und Informationsfreiheitsbericht des Landesbeauftragten für Datenschutz und Informationsfreiheit Nordrhein-Westfalen (2013/2014), 2015, S. 39, abrufbar unter: https://www.idi.nrw.de/mainmenu_Service/submenu_Berichte/Inhalt/22_DIB/DIB_22.pdf [Stand: 26.07.2021]; eine normative Begründung wird hierfür nicht gegeben. Eine solche Zusammenführung könnte gegen den Grundsatz der Datenminimierung in Art. 5 Abs. 1 lit. c DSGVO verstoßen. Der Versicherer selbst braucht für seine Zwecke nur die Auswertung der Fahrdaten, nicht diese selbst. Minimiert werden so zwar nicht die erhobenen Daten als solche, aber ihre Zusammenführung in einer Hand. Wird die Zulässigkeit der Datenverarbeitung aus Art. 6 Abs. 1 Unterabs. 1 lit. DSGVO hergeleitet, ergibt sich auch daraus, dass der Versicherer selbst nicht mehr Daten als zur Durchführung des Pay-as-you-drive-Tarifes verarbeiten darf (in diesem Sinne Specht/Mantz/Spitka, Handbuch europäisches und deutsches Datenschutzrecht (2019), § 12 Datenschutz in der Privatversicherung, Rn. 59). Ob die Trennung der Datenkreise als eine erforderliche angemessene Maßnahme i. S. v. Art. 22 Abs. 1, 2 lit. c i. V. m. Abs. 3 DSGVO anzusehen ist, um die berechtigten Interessen des Versicherungsnehmers in einem solchen Falle des Profilings zu wahren, ist bisher nicht diskutiert worden.

¹²⁶ BaFin, Rundschreiben 2/2017 (VA) – Mindestanforderungen an die Geschäftsorganisation von Versicherungsunternehmen (MaGo), Kap. 13, abrufbar unter: https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Rundschreiben/2017/rs_1702_mago_va.html [Stand: 26.07.2021].

¹²⁷ EIOPA, Leitlinien zum Governance-System (EIOPA-BoS-14/253), Leitlinien 14 und 60–64, abrufbar unter: https://www.eiopa.europa.eu/content/guidelines-system-governance_en [Stand: 26.07.2021].

¹²⁸ Ammann, CR (2020), 633, 635.

¹²⁹ Specht/Mantz/Spitka, Handbuch europäisches und deutsches Datenschutzrecht (2019), § 12 Datenschutz in der Privatversicherung, Rn. 66 ff.

um Missstände in Gestalt von Gesetzesverstößen zu vermeiden.¹³⁰ Dabei überwacht die Aufsicht allerdings nicht die Einhaltung der gesamten Rechtsordnung durch die Versicherungsunternehmen, sondern nur die Gesetze, die für den Betrieb des Versicherungsgeschäfts gelten. Dazu gehören eindeutig die besonderen Diskriminierungsverbote aus dem AGG (Abschn. 3.1), dem GenDG (Abschn. 3.2) und dem VAG (Abschn. 3.3), denn sie beziehen sich ausdrücklich auf Versicherungsnehmer. Auch die vertraglichen Rücksichtnahmepflichten (Abschn. 3.5) sind spezifisch im Versicherungsvertrag begründet, auch wenn sie aus allgemeinem bürgerlichem Recht hergeleitet werden. § 37 BDSG i. V. m. Art. 22 Abs. 1 DSGVO (Abschn. 3.7) spricht ebenfalls speziell die Leistungserbringung nach dem Versicherungsvertrag an, sodass die Norm der Aufsicht der BaFin unterfällt.

Hinzu kommen aber auch solche Vorschriften, die objektiv dazu dienen, Versicherungsnehmer und sonstige Begünstigte zu schützen.¹³¹ Die allgemeinen Diskriminierungsverbote aus Art. 3 GG (Abschn. 3.4) i. V. m. anderen Normen des Zivilrechts und aus dem Allgemeinen Persönlichkeitsrecht (Abschn. 3.6) betreffen in ihrer Anwendung auf Versicherungsnehmer deren Verhältnis zum Versicherungsunternehmen. Damit ist ein hinreichender Bezug zum Betrieb von Versicherungsgeschäften gegeben.¹³²

Ob und inwieweit beim Datenschutzrecht eine parallele Zuständigkeit der Versicherungsaufsicht neben derjenigen der Datenschutzbehörden gegeben ist, ist im Einzelnen nicht geklärt. Tatsächlich wird überwiegend angenommen, dass die Versicherungsaufsicht keine Normen überwacht, deren Durchsetzung anderen Behörden obliegt, es sei denn, das sei ausdrücklich vorgesehen.¹³³ Dann stünden die im Folgenden zu behandelnden versicherungsaufsichtsrechtlichen Maßnahmen zur Durchsetzung des Datenschutzrechts nicht zur Verfügung. Allerdings nennt die BaFin das BDSG – nicht aber die DSGVO – als ein Gesetz, dessen Einhaltung sie im Allgemeininteresse überwacht.¹³⁴ Für die Auffassung der BaFin lässt sich anführen, dass das Datenschutzrecht mit dem VVG und dem VAG eng verflochten ist, also seinerseits diesen Rechtsrahmen berücksichtigen muss.¹³⁵ Insbesondere kennt es spezielle Regelungen für Versicherer (s. oben Abschn. 3.7), und auch das VVG enthält datenschutzrechtliche Normen (§ 213 VVG). Hinzu kommt, dass der Versicherungs-

¹³⁰ Zu diesem Verständnis der Missstandsaufsicht s. VGH Kassel, Urteil vom 30.04.2020, 6 A 2158/18, in VersR (2020), 819–829, 823 ff.

¹³¹ BeckOK VAG/Kaulbach/Schäfers (2021), Stand: 01.06.2021, § 294 VAG, Rn. 7; Prölss/Dreher/Dreher, Versicherungsaufsichtsgesetz, 13. Aufl. (2018), § 298 VAG, Rn. 63.

¹³² So für das AGG auch BeckOK VAG/Kaulbach/Schäfers (2021), Stand: 01.06.2021, § 294 VAG, Rn. 7.

¹³³ Begr. RegE, BT-Drucks. 12/6959, S. 82; für das Beispiel der Kartellaufsicht in diesem Sinne Prölss/Dreher/Dreher, Versicherungsaufsichtsgesetz, 13. Aufl. (2018), § 298 VAG, Rn. 67; für das Beispiel des Kartell- und des Datenschutzes Brand/Baroch Castellvi/Brand, Versicherungsaufsichtsgesetz (2018), § 294 VAG, Rn. 31.

¹³⁴ BaFin., 16.03.2011, geändert am 16.05.2019, Vorschriften des Allgemeininteresses in Deutschland, abrufbar unter https://www.bafin.de/SharedDocs/Veroeffentlichungen/DE/Merkblatt/VA/mb_110308_zulassung_eu_liste_vorschriftenallgemeininteresse_va.html [Stand 19.04.2022].

¹³⁵ Specht/Mantz/Spitka, Handbuch europäisches und deutsches Datenschutzrecht (2019), § 12 Datenschutz in der Privatversicherung, Rn. 4.

sektor ein besonders datenintensiver Sektor ist,¹³⁶ da Daten über Risiken Grundlage des Geschäftsmodells der Versicherer sind. Datenschutzverstöße gehen im Übrigen in aller Regel zulasten der Versicherungsnehmer, deren Schutz das Hauptziel der Versicherungsaufsicht ist. Der BaFin hier Eingriffsmöglichkeiten abzusprechen und sich auf die unspezifischen und teils auch schwächeren Kompetenzen der Datenschutzbeauftragten zu verlassen, würde die Rechtsdurchsetzung schwächen. Zudem hat die BaFin die Fachkenntnisse, um versicherungsspezifische Diskriminierung, die zuweilen recht versteckt stattfinden kann, zu identifizieren. Daher ist die Wahrung der datenschutzrechtlichen Grenzen für Diskriminierung auch von der BaFin zu überwachen, soweit das Datenschutzrecht auf den Schutz der Versicherungsnehmer zielt. Das führt zu parallelen Zuständigkeiten, nicht zur Verdrängung der Versicherungsaufsicht.

5.2.2 Eingriffsschwelle

Des Weiteren ist zu klären, unter welchen Voraussetzungen die Aufsichtsbehörde bei Verstößen der oben genannten Art eingreifen kann. Im Grundsatz kann die Aufsichtsbehörde eingreifen, wenn das Versicherungsunternehmen gegen Gesetze verstößt. Sie ist beim „Ob“ und „Wie“ des Eingreifens an den Verhältnismäßigkeitsgrundsatz gebunden. Bei einmaligen, folgenlosen Gesetzesverstößen kann das etwa bedeuten, dass sich die Aufsicht mit einem Hinweis begnügt.¹³⁷ Bei den hier in Rede stehenden, KI-basierten Diskriminierungen wird ein Eingreifen jedoch schon deshalb geboten sein, weil ein Bias als systematische Abweichung zwischen der Modellierung und der Wirklichkeit eine Vielzahl von Fällen betreffen kann.

Auch eine innovationsspezifische Ausnahme ist nicht zu machen. Für den Bereich der neuen Finanztechnologien wird zwar seit langem, u. a. nach britischem Modell, diskutiert, ob Aufsichtsanforderungen abgesenkt werden müssen, um Innovation zu fördern (sog. regulatory sandbox).¹³⁸ Das europäisch harmonisierte Versicherungsaufsichtsrecht kennt zwar keine solche Innovationsklausel, lässt aber über den Verhältnismäßigkeitsgrundsatz Differenzierungen in der Aufsichtsintensität zu. Diskutiert wird insbesondere, ob man Versicherungsunternehmen, die ihr Geschäftsmodell auf innovative Technologien stützen (sog. InsurTechs), im Rahmen der Erlaubnisverfahrens entgegenkommt, indem man ihnen bei der Finanzausstattung mehr Spielraum lässt.¹³⁹ Die BaFin hat das bisher abgelehnt („Same business, same risks, same rules“)¹⁴⁰ und jüngst betont, eine volle Ausfinanzierung von InsurTechs schon

¹³⁶ Specht/Mantz/Spitka, Handbuch europäisches und deutsches Datenschutzrecht (2019), § 12 Datenschutz in der Privatversicherung, Rn. 2.

¹³⁷ Prölss/Dreher/Dreher, Versicherungsaufsichtsgesetz, 13. Aufl. (2018), § 298 VAG, Rn. 92 legt eine besondere Prüfung der Verhältnismäßigkeit in solchen Fällen nahe.

¹³⁸ Dazu etwa Möslein/Omlor, BKR (2018), 236, 238.

¹³⁹ Surminski, ZfV (2021), 98.

¹⁴⁰ Oster, Impulsreferat 3: Zulassung und laufende Aufsicht von Insurtechs in Zeiten der digitalen Transformation, Vortrag vom 29.10.2019, Präsentation S. 6, abrufbar unter: https://www.bafin.de/SharedDocs/Downloads/DE/Veranstaltung/dl_191029_va_jahreskonferenz_Vortragsunterlagen_3.pdf [Stand: 26.07.2021]; Surminski, ZfV (2021), 98.

bei Zulassung zu fordern.¹⁴¹ aber ein Bündel interner Maßnahmen ergriffen, um den neuen Fragestellungen gerecht zu werden.¹⁴² Ein großzügigerer Umgang mit Verstößen gegen Gesetze, die das Verhältnis zu den Versicherungsnehmern regeln, zählt allerdings zu Recht ohnehin nicht zu den für eine Sandbox diskutierten Erleichterungen. Der noch zu besprechende Entwurf der EU-Kommission für einen Artificial Intelligence Act sieht gewisse aufsichtsbehördliche Erleichterungen für die KI-Aufsicht vor; zu diesen und ihrer Bedeutung für die Versicherungsaufsicht s. unten Abschn. 6.

5.2.3 Eingriffsmittel

Klassische Eingriffsmittel der Aufsichtsbehörden sind zunächst, zur Aufklärung des Sachverhalts, das Auskunftsverlangen (§ 43 VAG i. V. m. § 305 VAG) und das Betreten und Durchsuchen von Räumen (§ 306 VAG). Nach Aufklärung und Würdigung des Sachverhalts sowie Anhörung der Betroffenen kann die BaFin alle verhältnismäßigen Maßnahmen ergreifen, um Verstöße abzustellen. Die Einführung von KI verändert die Prämissen der Versicherungsaufsicht schon bei der Erfassung des Sachverhalts grundlegend.¹⁴³ Die Aufsicht muss digital „aufrüsten“, um ihrerseits den Einsatz von KI durch das Versicherungsunternehmen überprüfen zu können.¹⁴⁴ Das entsprechende Stichwort lautet „Regulatory Technology“ (RegTech). Insbesondere im Vereinigten Königreich scheinen entsprechende Überlegungen, die technologische Kompetenz zu erhöhen, vorangetrieben zu werden. Befreit von den Fesseln der EU diskutiert man gar die – konsequente – Perspektive, aufsichtsrechtliche Normen in Computersprache zu fassen.¹⁴⁵ Fest steht, dass nicht nur die Versicherungsaufsicht, sondern jegliche staatliche Aufsicht über Unternehmen darauf hinwirken muss, den Unternehmen technologisch gewachsen zu sein. Wie schwierig das angesichts der unterschiedlichen finanziellen Basis ist, zeigt die globale Diskussion um die Regulierung der BigTechs wie Google, Amazon und Facebook. Ob und inwieweit angesichts der technischen Anforderungen bei der Aufsicht über KI eine – auch technische – Aufgabenteilung mit der nach EU-Recht zu schaffenden KI-Aufsicht (dazu sogleich) geboten ist, hängt von den Aufgaben dieser spezifischen KI-Aufsicht ab.

¹⁴¹ Näher dazu *Bürkle*, VW (2021), Heft 5, 84 ff.

¹⁴² *BaFin*, Digitalisierungsstrategie der BaFin, August 2018, S. 5 ff., abrufbar unter: https://www.bafin.de/DE/DieBaFin/ZieleStrategie/Digitalisierungsstrategie/digitalisierungsstrategie_node.html [Stand: 26.07.2021].

¹⁴³ *Swedloff* (2019), Regulating algorithmic insurance, 04.03.2019, S. 52, abrufbar unter: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3346753 [Stand 26.07.2021].

¹⁴⁴ *Kröner/Osman*, Handelsblatt, KI: Einsatz von stellt Regulierer vor Probleme, Artikel vom 18.11.2018.

¹⁴⁵ *Micheler/Whaley*, EBOR 21 (2020), 349 ff.

6 Kommissions-Entwurf eines Artificial Intelligence Act

Die EU-Kommission hat jüngst eine Verordnung vorgeschlagen, die das Inverkehrbringen, die Inbetriebnahme und die Nutzung von KI-Systemen regelt.¹⁴⁶ Auch bei diesen Vorschriften geht es zum Teil darum, Diskriminierung zu verhindern¹⁴⁷ und Transparenz zu schaffen, die ihrerseits sachlich nicht gerechtfertigte Ungleichbehandlungen aufzudecken hilft. Nach der Legaldefinition der Verordnung ist ein KI-System eine Software, die mit einer oder mehreren bestimmten Techniken entwickelt wurde und für eine gegebene Menge von durch den Menschen definierten Zielen Ausgaben wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen erzeugen kann, die die Umgebungen beeinflussen, mit denen sie interagieren (Art. 3 Abs. 1 KI-VO-E).

Die einschlägigen Techniken sind in Annex I der KI-VO-E aufgelistet, und zu ihnen gehören auch die oben (Abschn. 1 und 2) beschriebenen selbstlernenden Algorithmen. Auch die weiteren Voraussetzungen der Definition sind erfüllt. Setzen Versicherer im Verhältnis zum Kunden Algorithmen ein, definieren sie Ziele, wie etwa den Abschluss eines Vertrages, die Einschätzung eines Risikos oder die Bewertung eines Schadens. Diese Ziele sind von Menschen gesetzt, auch wenn Versicherer juristische Personen sind. Ihnen sind die Ziele, die ihre Organe und Erfüllungsgehilfen setzen, im Verhältnis zum Versicherungsnehmer zuzurechnen. Die KI kann Ausgaben im beschriebenen Sinne erzeugen, und diese beeinflussen die Umgebungen, mit denen das KI-System interagiert. Umgebung in diesem Sinne ist weit zu verstehen, wie Erwägungsgrund 6 des KI-VO-E zeigt. Dort wird hervorgehoben, dass Einflüsse physischer und digitaler Dimension erfasst sein sollen.

Fallen demnach die hier angesprochenen KI-Anwendungen der Versicherungswirtschaft unter die geplante Verordnung, werden insbesondere die harmonisierten Transparenzregeln des Art. 52 KI-VO-E anzuwenden sein, da sie für KI-Systeme gelten sollen, die mit natürlichen Personen interagieren. Ob dazu auch eine Interaktion mit natürlichen Personen zählt, die als Organe oder Erfüllungsgehilfen einer juristischen Person handeln, sagt die KI-VO-E nicht ausdrücklich. Sinn und Zweck der Transparenzregeln der Verordnung ist es, eine Täuschung natürlicher Personen zu verhindern (s. Erwägungsgrund 70 KI-VO-E). Dieser Zweck ist zum einen nicht auf den rechtsgeschäftlichen Verkehr und zum anderen nicht darauf beschränkt, dass diese natürlichen Personen für sich selbst handeln. Daher dürften die Transparenzregeln der Verordnung für Versicherer unabhängig davon eingreifen, ob diese mit Privatpersonen oder Unternehmen interagieren. Nach Art. 52 Abs. 1 KI-VO-E müssen Versicherer daher darauf hinweisen, wenn sie mittels einer KI mit ihren Verhandlungs- oder Vertragspartnern in einen Austausch treten.¹⁴⁸

¹⁴⁶ Europäische Kommission, 21.04.2021, Proposal for a Regulation of the European Parliament and of the Council laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts, COM(2021) 206 final (im Folgenden KI-VO-E), abrufbar unter: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELLAR:e0649735-a372-11eb-9585-01aa75ed71a1> [Stand: 26.07.2021]; Kurzüberblick bei Behr, BB (2021), Heft 20, I.

¹⁴⁷ So auch Spindler, CR (2021), 361, 362; näher Sesing/Tschech, MMR (2022), 24, 28 ff.

¹⁴⁸ Eine Pflicht, ein Angebot ohne KI-Einsatz zu machen, besteht jedoch nicht; kritisch deshalb Spindler, CR (2021), 361, 368, 374.

Sollten sie biometrische Kategorisierungssysteme verwenden (Art. 3 Abs. 35 KI-VO-E), also solche, die natürliche Personen anhand ihrer biometrischen Daten¹⁴⁹ bestimmten Kategorien wie Geschlecht, Alter, Haarfarbe, Augenfarbe, Tätowierungen, ethnischer Herkunft oder sexueller oder politischer Ausrichtung zuordnen, muss auch das transparent gemacht werden. Denkbar ist das u. U., wenn bei Personenversicherungen anhand von erfassten Verhaltensdaten auf den Gesundheitszustand oder anhand der Fahrweise auf das Unfallrisiko geschlossen wird. Der genannte Katalog der Kategorien dürfte nicht abschließend, sondern beispielhaft gemeint sein.

Strengere Regeln gelten für hoch-riskante KI-Anwendungen (Art. 6 ff. KI-VO-E)¹⁵⁰; das Management der damit verbundenen Risiken – auch Bias-Risiken, s. Art. 10 Abs. 2 lit. f, 5 KI-VO-E –, die Daten-Governance, die technische Dokumentation, Mechanismen zur Nachverfolgbarkeit und menschliche Aufsicht sind hier detailliert vorgegeben. Die Regeln finden auf Versicherungsprodukte Anwendung, wenn biometrische Kategorisierungssysteme zur biometrischen Fernidentifizierung natürlicher Personen in Echtzeit oder nachträglich eingesetzt werden (Art. 6 Abs. 2 KI-VO-E i. V. m. Annex III), wie es etwa bei „Pay-as-you-live“-Versicherungen der Fall sein könnte, etwa dann, wenn eine Lebensversicherung verfolgen würde, ob der Versicherungsnehmer raucht, aber auch wenn eine Reisgepäckversicherung automatisch durch Verlassen des Hauses mit Gepäck aktiviert wird.

Die Transparenzregeln in Art. 52 KI-VO-E dienen zumindest auch dazu, Versicherungsnehmer zu schützen. Die Regeln für hoch-riskante KI sollen voreingenommene und diskriminierende Ergebnisse vermeiden (Erwägungsgrund 33 KI-VO-E) und so auch Versicherungsnehmer schützen. Sämtliche der genannten Regeln können daher grds. von der Versicherungsaufsicht überwacht werden. Allerdings wird hier zugleich eine Zuständigkeit einer nach Art. 59 KI-VO-E zu schaffenden nationalen, branchenübergreifenden KI-Aufsichtsbehörde gegeben sein. Die spezifische Aufsicht über die Technologie wird dann sinnvollerweise ausschließlich in der Hand dieser Behörde liegen, da diese das umfassende Know-how auch zum Bias-Management haben wird.¹⁵¹ Trifft sie Feststellungen und Maßnahmen im Rahmen der Technologieaufsicht, kann die BaFin dann u. U. im Rahmen der Legalitätsaufsicht nach §§ 298, 294 VAG auch versicherungsaufsichtsrechtliche Konsequenzen ziehen. Die Zusammenarbeit der Behörden muss entsprechend geregelt werden.

Für die KI-Aufsicht nach dem KI-VO-E sieht die Kommission bislang einen regulatorischen Sandkasten vor (Art. 53 KI-VO-E), um Innovationen nicht durch Überregulierung zu bremsen. Er bringt allerdings keine Absenkung der KI-aufsichtsrechtlichen Standards mit sich,¹⁵² sondern ermöglicht es, KI-Systeme vor ihrem Inverkehrbringen oder der Inbetriebnahme unter Aufsicht zu testen. Die Mitgliedstaaten werden in Erwägungsgrund 71 KI-VO-E ermutigt, allerdings nicht verpflicht-

¹⁴⁹ Das sind nach Art. 3 Abs. 33 KI-VO-E durch spezifische Technologien gewonnene Merkmale der Physis, der Psyche oder des Verhaltens, die eine eindeutige Identifizierung ermöglichen.

¹⁵⁰ Einen guten Überblick gibt *Spindler*, CR (2021), 361, 366 ff.

¹⁵¹ Vgl. allg. zur Vermeidung doppelter Aufsichtsstrukturen *GDV* (2021), Positionspapier des Gesamtverbandes der Deutschen Versicherungswirtschaft zum Vorschlag der EU-Kommission für eine Verordnung zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz, 13.07.2021, S. 8f.

¹⁵² *Spindler*, CR (2021), 361, 371.

tet, ihn einzurichten. Tun sie das, sind die entsprechenden Vorgaben des KI-VO-E zu beachten, d.h. die Sandboxes sind in gewissem Maße harmonisiert. Die deutsche Versicherungsaufsicht bietet bisher keine solche Erleichterung für innovative Geschäftsmodelle. Art. 53 Abs. 3 KI-VO-E legt ausdrücklich fest, dass ein regulatorischer Sandkasten für das KI-System Aufsichts- und Eingriffskompetenzen anderer zuständiger Behörden nicht berührt. Allerdings ist die BaFin hinzuzuziehen, wenn für versicherungsrelevante KI-Systeme andere Behörden einen solchen privilegierter Aufsichtsbereich einrichten (Art. 53 Abs. 2 KI-VO-E).

7 Ergebnisse und Fazit

Setzen Versicherungsunternehmen im Verhältnis zu ihren Kunden, Versicherungsnehmern und den Begünstigten von Versicherungsleistungen KI in Gestalt selbstlernender Algorithmen ein, die mithilfe von Daten Lösungen für wiederkehrende Entscheidungsprobleme finden können, kann die KI eine systematische Abweichung zwischen der Modellierung und der Wirklichkeit entwickeln, einen Bias. Die Ursachen von Bias sind vielfältig und können in der Datengrundlage, in der Zielsetzung und den Schlussfolgerungen der KI liegen. Beim Einsatz von KI ist ein Bias niemals ganz vermeidbar, man kann ihn durch verschiedene Maßnahmen in verschiedenen Phasen der Entwicklung und Nutzung der KI aber möglichst geringhalten, also minimieren.

KI wird typischerweise eingesetzt, wenn eine Vielzahl gleichartiger Entscheidungen zu treffen sind, etwa die Entscheidung über den Vertragsabschluss, die Berechnung des Tarifs oder die Regulierung des Schadens. Ein Bias wird daher regelmäßig große Breitenwirkung haben, sich also auf eine Vielzahl von Entscheidungen auswirken. Da ein Bias juristisch gesprochen zu einer verbotenen Diskriminierung führen kann, drohen damit massenhafte Rechtsverstöße.

Zivilrechtlich knüpfen sich an die Verstöße gegen die verschiedenen Diskriminierungsverbote Ansprüche des Versicherungsnehmers gegen den Versicherer auf Unterlassung, Beseitigung und Schadensersatz.

Erstens ist eine zentrale Weichenstellung hierbei die Frage, ob die Erklärung, welche dem Versicherungsnehmer zugeht, nach dem Empfängerhorizont als Willenserklärung des Versicherers anzusehen ist. Ist dies der Fall, entfallen die vielfach diskutierten Zurechnungsprobleme im Rahmen von Willenserklärungen und KI sowie Pflichtverletzungen und KI.

Zweitens spielt die Erklärbarkeit von KI vor allem für den Versicherer selbst eine zentrale Rolle.¹⁵³ Die Erklärbarkeit kann sich im Rahmen von Kausalitäts- und Beweisfragen entscheidend darauf auswirken, ob eine rechtswidrige Diskriminierung vorliegt und ob die Datenverarbeitung aufgrund einer Einwilligung zulässig ist. Erklärbarkeit erhöht daher auch die Rechtssicherheit für die Versicherer.

Drittens ergibt sich, dass der Versicherer stets sicherzustellen hat, dass die KI nur zulässige Informationen im Sinne des GenDG erhebt sowie versicherungsauf-

¹⁵³ Das betonen auch *Salchegger/Thogmartin*, VW (2020), Heft 11, 64 ff.; *Nguyen*, VW (2019), Heft 9, 52 ff.

sichtsrechtliche und datenschutzrechtliche Vorgaben gewahrt werden. Dabei ist insbesondere auch zu verhindern, dass die KI durch ihr selbstlernendes Verhalten diese Verbote und Vorgaben nicht umgeht.

Viertens ist für das Verschulden und damit die Haftung des Versicherers für die Auswirkungen bias-belasteter Entscheidungen entscheidend, ob und inwieweit Bias vermeidbar ist. Die Unvermeidbarkeit von Bias hat zur Folge, dass jeder Versicherer, der KI in der hier diskutierten Weise gegenüber Kunden, Versicherungsnehmern und den Begünstigten von Versicherungsleistungen einsetzt, vorhersehen kann, dass Diskriminierungen geschehen können; welche Diskriminierungen im Einzelnen drohen, ist für ihn nicht voraussehbar. Er kann diese Gefahr minimieren, indem er Bias-Management betreibt. Vermeiden kann er sie nach heutigem Stand der Technik nur, indem er auf die Nutzung von selbstlernender KI im direkten Umgang mit dem Kunden, also etwa beim Vertragsabschluss oder bei der Schadenregulierung, verzichtet. Daher sind Diskriminierungen durch eine KI dem Versicherer zurechenbar: Sie sind vorhersehbar und – durch Nichtnutzung von KI – vermeidbar.

Das Versicherungsaufsichtsrecht stellt über die Govenancevorgaben sicher, dass die Diskriminierungsverbote eingehalten werden. Sonstige konkrete Vorgaben zur Vermeidung von Bias gibt es nicht, aber allgemeine Regeln, die eine konsistente IT-Strategie verlangen. Diese umfasst die IT-Governance und das IT-Risikomanagement und ist der Aufsichtsbehörde mitzuteilen. Bei KI-basierten Versicherungsprodukten sind Diskriminierungsrisiken zudem im Rahmen der Produktgovernance zu bewerten, zu beobachten und zu minimieren. Werden aus praktischen oder rechtlichen Gründen Aufgaben im Zusammenhang mit KI-basierten Maßnahmen ausgegliedert, sind die allgemeinen Vorgaben für Ausgliederungen zu beachten. Verstößen Versicherer gegen diese aufsichtsrechtlichen Vorgaben oder gegen die verschiedenen Diskriminierungsverbote, können die Aufsichtsbehörden eingreifen. Angesichts der potenziellen Breitenwirkung von Bias sind behördliche Eingriffe auch regelmäßig geboten.

Neben dem Versicherungsaufsichtsrecht setzt das Datenschutzrecht der KI-Nutzung durch Versicherer im Verhältnis zu den genannten Personenkreisen Schranken durch Diskriminierungsverbote, aber auch durch Govenancevorgaben (Folgenabschätzung, Risikomanagement), die hier nur angedeutet wurden. Sie begründen neben Schadensersatzansprüchen Durchsetzungskompetenzen der Datenschutzbehörden.

Mit dem Vorschlag der EU-Kommission für einen Artificial Intelligence Act konkretisieren sich Überlegungen einer behördlichen Aufsicht über KI. Diese spezifische KI-Aufsicht wird auch Versicherungsunternehmen erfassen, die KI nutzen. Sie wird auch das hier erörterte Problem des Bias angehen. Es erscheint sachgerecht, eine technologiespezifische Aufsicht neben die Versicherungsaufsicht – und sonstige Arten der Wirtschaftsaufsicht – zu stellen. Die Möglichkeiten der einzelnen Zweige der Wirtschaftsaufsicht und damit auch der Versicherungsaufsicht, technologisch „aufzurüsten“ (Stichwort: RegTech) und KI wirksam zu kontrollieren, sind begrenzt.

Aus unserer Sicht ergibt sich aus dem Gesagten eine Reihe von Fragestellungen, die gelöst werden müssen. Dabei muss es das Ziel sein, dem Einsatz von KI in der Versicherung einen rechtsicheren Rahmen zu geben. Denn die Rechts-

unsicherheit wird von den Versicherern als einer der Haupthinderungsgründe in Bezug auf den Einsatz von KI genannt, neben technischen Herausforderungen und Akzeptanzfragen.¹⁵⁴

Die Haftungsfrage speziell im Hinblick auf Bias als eines unvermeidbaren Risikos muss geklärt werden; hier wurde für eine umfassende Zurechnung möglicher Diskriminierungen zum Versicherungsunternehmen plädiert. Auch über eine Anpassung des Diskriminierungsrechts an die technische Entwicklung ist nachzudenken.¹⁵⁵ Zu klären ist zudem, ob Erkenntnisse der Informatik zur (Un-)Fairness von KI auch für die rechtliche Bewertung von Bedeutung sein können und möglicherweise die rechtlichen Grenzen der Diskriminierung verschieben. Ebenfalls im Zusammenhang mit der Haftung steht die Erklärbarkeit von KI, die für die Feststellung von Kausalzusammenhängen eine Rolle spielen kann. Hier sind technische Möglichkeiten zu entwickeln, die ihrerseits auf die juristischen Anforderungen rückwirken können. Andererseits stellt sich aber auch die Frage, ob durch eine erklärbare KI die oben aufgeworfenen Beweisprobleme zufriedenstellend gelöst werden und welche neuen rechtlichen Herausforderungen sich bei einer erklärbaren KI stellen.

Zu klären ist von Seiten der Informatik auch, was Bias-Vermeidung im Sinne eines angemessenen Bias-Managements im Detail erfordert. Davon ausgehend stellt sich die juristische Frage, welche Maßnahmen davon mit welcher Detailliertheit durch welche aufsichtsrechtlichen Instrumente vorgegeben werden sollen. Verwaltungsorganisatorisch ist zu fragen, in die Hand welcher Aufsichtsbehörden die Überwachung der Maßnahmen gelegt wird. Hier muss ein stimmiges Zusammenwirken von Technologieaufsicht, Datenschutzaufsicht und Fachaufsicht erreicht werden.

Der Umgang mit Bias-Risiken beim Einsatz von KI in der Versicherung ist nach alldem eine nur interdisziplinär von der Informatik und der Rechtswissenschaft zu lösende Fragestellung.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

¹⁵⁴ *EIOPA*, Big Data Analytics in motor and health insurance (2019), S. 47.

¹⁵⁵ Hierzu etwa *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz (2019, S. 230 ff.); *Orwat*, Diskriminierungsrisiken durch Verwendung von Algorithmen (2019, S. 97 ff.).

Literatur

- Allen, R., Masters, D.: Artificial Intelligence: the right to protection from discrimination caused by algorithms, machine learning and automated decision-making. *ERA Forum* **20**, 585–598 (2020)
- Ammann, T.: Der Einsatz Künstlicher Intelligenz in der Finanz- und Versicherungswirtschaft, Markttrend und regulatorische Herausforderungen. *CR* **36**, 633–640 (2020)
- Ashri, R.: The AI-Powered Workplace: How Artificial Intelligence, Data, and Messaging Platforms Are Defining the Future of Work (2020)
- BaFin: IT-Kompetenz in der Geschäftsleitung, BaFin passt Entscheidungsmaßstäbe für Bestellung von IT-Spezialisten zu Geschäftsleitern an. *BaFin Journal*, Heft 12, 15–18 (2017)
- BaFin: Big Data trifft auf künstliche Intelligenz, Herausforderungen und Implikationen für Aufsicht und Regulierung von Finanzdienstleistungen (2018)
- BaFin: Big Data und künstliche Intelligenz: Prinzipien für den Einsatz von Algorithmen in Entscheidungsprozessen, Prinzipienpapier BDAI, 15.06.2021 (2021)
- Barocas, S., Hardt, M., Narayanan, A.: *Fairness and Machine Learning, Limitations and Opportunities* (2019)
- Bartel, K., Khuat, L.V., Wehling-Mauntel, U.: § 7 B. Versicherungsrecht. In: Chibanguza, K., Kuß, C., Steege, H. (Hrsg.) *Künstliche Intelligenz, Recht und Praxis automatisierter und autonomer Systeme*, S. 1049–1064. (2022)
- Bartuschka, W.: Künstliche Intelligenz, Chatbots, RobotProcess Automation – neue Technologien als Fluch oder Segen für gute Compliance? *CB* **7**, 340–345 (2019)
- Basedow, J.: Strikte Haftung und „nackte“ Pflichtversicherungen, EU-Konzepte für die digitale Welt? *EuZW* **32**, 1–2 (2021)
- Beck, S.: Diskriminierung durch Künstliche Intelligenz? *ZRP* **52**, 185 (2019)
- Behr, N.: „Je größer das Risiko, desto strenger die Regeln.“ – Wie die EU-Kommission künstliche Intelligenz regulieren möchte. *BB* **76**(20), I (2021)
- Bilski, N., Schmid, T.: Verantwortungsfindung beim Einsatz maschinell lernender Systeme. *NJOZ* **19**, 657–661 (2019)
- Bork, R.: Allgemeiner Teil des Bürgerlichen Gesetzbuchs, 4. Aufl. (2016)
- Brand, O., Baroch Castellvi, M. (Hrsg.): *Versicherungsaufsichtsgesetz, Handkommentar* (2018)
- Brink, S., Wolff, H.A. (Hrsg.): *BeckOK Datenschutzrecht*, 33. Ed. (2020). Zitiert als BeckOK *Datenschutzrecht/Bearbeiter*
- Bürkle, J.: § 9 Geschäftsorganisation und Compliance von Versicherungsunternehmen. In: Bähr, G.W. (Hrsg.) *Handbuch des Versicherungsaufsichtsrechts*, S. 181–222. (2011)
- Bürkle, J.: Gefährliche Mutanten im Versicherungssektor?, Eine Nachbetrachtung zu den verschärften Zulassungsregeln für Insurtechs. *VW* **76**(5), 84–87 (2021)
- Buxmann, P., Schmidt, H.: Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens. In: Buxmann, P., Schmidt, H. (Hrsg.) *Künstliche Intelligenz, Mit Algorithmen zum wirtschaftlichen Erfolg*, S. 3–19. (2019)
- Chen, J., Kallus, N., Xiaojie, M., Svacha, G., Udell, M.: Fairness under unawareness: assessing disparity when protected class is unobserved. In: *FAT* '19: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*, S. 339–348, DOI 10.1145/3287560.3287594. (2019)
- Chouldechova, A.: Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data* **5**(2), 153–163 (2017)
- Däubler, W., Bertzbach, M. (Hrsg.): *Allgemeines Gleichbehandlungsgesetz, Handkommentar*, 4. Aufl. (2018)
- Denga, M.: Deliktische Haftung für künstliche Intelligenz, Warum die Verschuldenshaftung des BGB auch künftig die bessere Schadensausgleichsordnung bedeutet. *CR* **34**, 69–78 (2018)
- Dressel, J., Farid, H.: The accuracy, fairness, and limits of predicting recidivism. *Sci. Adv.* **4**(1), eao5580 (2018)
- Eidenmüller, H., Wagner, G.: *Law by Algorithm* (2021)
- EIOPA: *Big Data Analytics in Motor and Health Insurance, A Thematic Review* (2019)
- EIOPA: *Artificial Intelligence Governance Principles: Towards Ethical and Trustworthy Artificial Intelligence in the European Insurance Sector, A Report from EIOPA's Consultative Expert Group*, 17.06.2021 (2021)
- Epping, V., Hillgruber, C. (Hrsg.): *BeckOK Grundgesetz*, 47. Ed. (2021). Zitiert als BeckOK *GG/Bearbeiter*
- Erdmann, K.U., Diehl, F.S., Schradin, H.R. (Hrsg.): *BeckOK VAG*, 13. Ed. (2021). Zitiert als BeckOK *VAG/Bearbeiter*
- Ernst, M.-A., Rogler, J. (Hrsg.): *Berufsunfähigkeitsversicherung, Handkommentar* (2018)

- Fader, P.S., Hardie, B.G.S.: Probability models for customer-base analysis. *J. Interactive Mark.* **23**(1), 61–69 (2009)
- Feldman, M., Friedler, S.A., Moeller, J., Scheidegger, C., Venkatasubramanian, S.: Certifying and removing disparate impact. In: *KDD '15: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, S. 259–268. (2015)
- Friedler, S.A., Scheidegger, C., Venkatasubramanian, S., Choudhary, S., Hamilton, E.P., Roth, D.: A comparative study of fairness-enhancing interventions in machine learning. In: *FAT* '19: Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*, S. 329–338. (2019)
- Friedman, B., Nissenbaum, H.: Bias in computer systems. *ACM Trans. Inf. Syst.* **14**(3), 330–347 (1996)
- Gausling, T.: KI und DSGVO im Spannungsverhältnis. In: Graf Ballestrem, J., Bär, U., Gausling, T., Hack, S., von Oelffen, S. (Hrsg.) *Künstliche Intelligenz, Rechtsgrundlagen und Strategien in der Praxis*, S. 11–53. (2020)
- Gausling, T.: Künstliche Intelligenz im digitalen Marketing, Datenschutzrechtliche Bewertung KI-gestützter Kommunikations-Tools und Profiling-Maßnahmen. *ZD* **9**, 335–341 (2019)
- GDV: Positionspapier des Gesamtverbandes der Deutschen Versicherungswirtschaft ID-Nummer 6437280268-55 zum Vorschlag der EU-Kommission für eine Verordnung zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz (2021). (Gesetz über künstliche Intelligenz), 13.07.2021
- Ghosh, A., Kleinberg, R.: Inferential privacy guarantees for differentially private mechanisms. In: Papadimitriou, C.H. (Hrsg.) *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, S. 9:1–9:3. (2017)
- Graewe, D.: New Corporate Governance: bessere Unternehmenskontrolle ohne Menschen. *BB* **75**(11), I (2020)
- Graf von Westphalen, F.: Definition der Künstlichen Intelligenz in der Kommissionsmitteilung COM (2020) 64 final – Auswirkungen auf das Vertragsrecht. *BB* **75**, 1859–1863 (2020)
- Graf von Westphalen, F.: Produkthaftungsrechtliche Erwägungen beim Versagen Künstlicher Intelligenz (KI) unter Beachtung der Mitteilung der Kommission COM(2020) 64 final. *VuR* **35**, 248–254 (2020)
- Green, B.: The false promise of risk assessments: Epistemic reform and the limits of fairness. In: *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, S. 594–606. (2020)
- Groß, J.: Braucht Künstliche Intelligenz Ethik? *CB* **8**, 155–159 (2020)
- Grünberger, M.: Reformbedarf im AGG: Beweislastverteilung beim Einsatz von KI. *ZRP* **54**, 232–235 (2021)
- Gsell, B., Krüger, W., Lorenz, S., Reymann, C. (Hrsg.): beck-online.GROSSKOMMENTAR, BGB (2021). Zitiert als BeckOGK BGB/Bearbeiter
- Gsell, B., Krüger, W., Lorenz, S., Reymann, C., Looschelders, D. (Hrsg.): beck-online.GROSSKOMMENTAR, AGG, (2021). Zitiert als BeckOGK AGG/Bearbeiter
- Günther, J., Böglmüller, M.: Künstliche Intelligenz und Roboter in der Arbeitswelt. *BB* **72**, 53–58 (2017)
- Hacker, P.: Verhaltens- und Wissenszurechnung beim Einsatz von Künstlicher Intelligenz. *RW* **9**, 243–288 (2018)
- Härtel, I.: Digitalisierung im Lichte des Verfassungsrechts – Algorithmen, Predictive Policing, autonomes Fahren. *LKV* **29**, 49–60. (2019)
- Haseloff, L., Friehoff, L.: Zurechnung von Erklärungen eines Chat-Bots oder intelligenten Agenten beim Abschluss von Versicherungsverträgen. *VersR* **71**, 1363–1369 (2020)
- Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2. Aufl. (2009)
- Hau, W., Poseck, R. (Hrsg.): BeckOK BGB, 58. Ed. (2021). Zitiert als BeckOK BGB/Bearbeiter
- Ilveto, C.: Metric learning for individual fairness. In: Roth, A. (Hrsg.) *1st Symposium on Foundations of Responsible Computing (FORC 2020)*, LIPICs, Bd. 156, S. 2:1–2:11. (2020)
- Intveen, M.: Automatisierte Softwareentwicklung unter Einsatz Künstlicher Intelligenz, Anforderungen und Fallstricke bei der jetzigen und zukünftigen Vertragsgestaltung. *ITRB* **20**, 70–75 (2020)
- Jobst, S.: Konsequenzen einer unmittelbaren Grundrechtsbindung Privater. *NJW* **73**, 11–16 (2020)
- Jung, P.: Privatversicherungsrechtliche Gefahrengemeinschaft und Treuepflicht des Versicherers. *VersR* **54**, 282–287 (2003)
- Käde, L., von Maltzan, S.: Die Erklärbarkeit von Künstlicher Intelligenz, Entmystifizierung der Black Box und Chancen für das Recht. *CR* **36**, 66–72 (2020)
- Kainer, F., Förster, L.: Autonome Systeme im Kontext des Vertragsrechts. *ZfPW* **6**, 275–305 (2020)
- Kamishima, T., Akaho, S., Asoh, H., Sakuma, J.: Fairness-aware classifier with prejudice remover regularizer. In: Flach, P.A., de Bie, T., Christianini, N. (Hrsg.) *Machine Learning and Knowledge Discovery*

- in Databases, European Conference, ECML PKDD 2012, Bristol, UK, September 2012, Proceedings, Part II, S. 35–50. (2012)
- Kaulbach, D., Bähr, G.W., Pohlmann, P. (Hrsg.): Versicherungsaufsichtsgesetz, Mit Finanzdienstleistungsaufsichtsgesetz, Verordnung (EU) Nr. 1094/2010 (EIOPA-Verordnung) und Versicherungs-Vergütungsordnung: Kommentar, 6. Aufl. (2019)
- Keßler, O.: Intelligente Roboter – neue Technologien im Einsatz, Voraussetzungen und Rechtsfolgen des Handelns informationstechnischer Systeme. *MMR* **20**, 589–594 (2017)
- Kolleck, A., Orwat, C.: Mögliche Diskriminierung durch algorithmische Entscheidungssysteme und maschinelles Lernen – ein Überblick. Hintergrundpapier, Bd. 24. (2020)
- Kreutzer, R.T., Sirrenberg, M.: Understanding Artificial Intelligence, Fundamentals, Use Cases and Methods for a Corporate AI Journey (2020)
- Kühling, J., Buchner, B. (Hrsg.): Datenschutz-Grundverordnung, Bundesdatenschutzgesetz, 3. Aufl. (2020)
- Langheid, T., Wandt, M. (Hrsg.): Münchener Kommentar zum Versicherungsvertragsgesetz, 2. Aufl. Bd. 1. (2016). Zitiert als MüKo VVG/Bearbeiter
- Linardatos, D.: Autonome und vernetzte Aktanten im Zivilrecht, Grundlinien zivilrechtlicher Zurechnung und Strukturmerkmale einer elektronischen Person (2021)
- Linke, C.: Die elektronische Person, Erforderlichkeit einer Rechtspersönlichkeit für autonome Systeme? *MMR* **24**, 200–204 (2021)
- Lohia, P.K., Natesan Ramamurthy, K., Bhide, M., Saha, D., Varshney, K.R., Puri, R.: Bias mitigation post-processing for individual and group fairness. In: ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), S. 2847–2851. (2019)
- Lücke, O.: Der Einsatz von KI in der und durch die Unternehmensleitung. „Lieutenant Commander Data“ on bord oder natural intelligence still needed? *BB* **74**, 1986–1994 (2019)
- Lüttringhaus, L.: Auf wackligem Fundament, Rechtsfragen der Künstlichen Intelligenz in der Versicherungsindustrie. *VW* **75**(6), 88–91 (2020)
- Martini, M.: Blackbox Algorithmen – Grundfragen einer Regulierung Künstlicher Intelligenz (2019)
- Martins, D.M.L.: Enabling Non-technical Users to Query and Purchase Data (2019)
- Micheler, E., Whaley, A.: Regulatory technology: Replacing law with computer code. *EBOR* **21**, 349–377 (2020)
- Möslin, F., Omlor, S.: Die europäische Agenda für innovative Finanztechnologien (FinTech). *BKR* **18**, 236–243 (2018)
- Neuner, J.: Das BVerfG im Labyrinth der Drittwirkung. *NJW* **73**, 1851–1855 (2020)
- Ngyuen, M.: Warum Roboter Entscheidungen fällen, die Versicherer nicht verstehen. *VW* **74**(9), 52–55 (2019)
- Noack, U.: Organisationspflichten und -strukturen kraft Digitalisierung. *ZHR* **183**, 105–144 (2019)
- Obermeyer, Z., Powers, B., Vogeli, C., Mullainathan, S.: Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**(6464), 447–453 (2019)
- Orwat, C.: Diskriminierungsrisiken durch Verwendung von Algorithmen, eine Studie erstellt mit einer Zuwendung der Antidiskriminierungsstelle des Bundes (2019)
- Pessach, D., Shmueli, E.: Algorithmic fairness (2020)
- Präve, P.: Das Gendiagnostikgesetz aus versicherungsrechtlicher Sicht. *VersR* **60**, 857–862 (2009)
- Prölss, E.R., Dreher, M. (Hrsg.): Versicherungsaufsichtsgesetz, mit Nebengesetzen, 13. Aufl. (2018)
- Prölss, E.R., Martin, A. (Hrsg.): Versicherungsvertragsgesetz, mit Nebengesetzen, Vertriebsrecht und Allgemeinen Versicherungsbedingungen, 31. Aufl. Bd. 14. (2021)
- Reusch, P.: Künstliche Intelligenz und Produkthaftung. *K&R* **22**(Beilage zu Heft 7/8), 20–21. (2019)
- Säcker, F.J., Rixecker, R., Oetker, H., Limperg, B. (Hrsg.): Münchener Kommentar zum Bürgerlichen Gesetzbuch, 8. Aufl. Bd. 1. (2018). Zitiert als MüKo BGB/Bearbeiter
- Säcker, F.J., Rixecker, R., Oetker, H., Limperg, B. (Hrsg.): Münchener Kommentar zum Bürgerlichen Gesetzbuch, 8. Aufl. Bd. 2. (2019). Zitiert als MüKo BGB/Bearbeiter
- Salchegger, M., Thogmartin, D.: Erklärungsbedürftig. *VW* **75**(11), 64–67 (2020)
- Salimi, B., Howe, B., Suciu, D.: Database repair meets algorithmic fairness. *ACM SIGMOD Rec.* **49**(1), 34–41 (2020)
- Schulze, R. (Hrsg.): Bürgerliches Gesetzbuch, Handkommentar, 10. Aufl. (2019)
- Sesing, A., Tschsch, A.: AGG und KI-VO-Entwurf beim Einsatz von Künstlicher Intelligenz, Einschätzung aus der Perspektive des (Anti-)Diskriminierungsrechts. *MMR* **25**, 24–30 (2022)
- Specht, L., Mantz, R. (Hrsg.): Handbuch europäisches und deutsches Datenschutzrecht, Bereichsspezifischer Datenschutz in Privatwirtschaft und öffentlichem Sektor (2019)

- Specht, L., Herold, S.: Roboter als Vertragspartner?, Gedanken zu Vertragsabschlüssen unter Einbeziehung automatisiert und autonom agierender Systeme. *MMR* **21**, 40–44 (2018)
- Spindler, G.: Der Vorschlag der EU-Kommission für eine Verordnung zur Regulierung der Künstlichen Intelligenz (KI-VO-E), Ansatz, Instrumente, Qualität und Kontext. *CR* **37**, 361–374 (2021)
- Spindler, G.: Roboter, Automation, künstliche Intelligenz, selbst-steuernde Kfz – Braucht das Recht neue Haftungskategorien?, Eine kritische Analyse möglicher Haftungsgrundlagen für autonome Steuerungen. *CR* **31**, 766–776 (2015)
- Steege, H.: Algorithmenbasierte Diskriminierung durch Einsatz von Künstlicher Intelligenz, Rechtsvergleichende Überlegungen und relevante Einsatzgebiete. *MMR* **22**, 715–721 (2019)
- Surminski, M.: Mehr Innovationen durch Sonderregeln für Insurtechs? *ZfV* **72**, 98 (2021)
- Swedloff, R.: Regulating algorithmic insurance, 04.03.2019 (2019). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3346753. Zugegriffen: 26.07.2021
- Wagner, G.: Verantwortlichkeit im Zeichen digitaler Techniken. *VersR* **71**, 717–741 (2020)
- von Walter, A.: Gedrosselte Intelligenz? Künstliche Intelligenz und der Datenschutz. *K&R* **22**(Beilage zu Heft 7/8), 21–24. (2019)
- Wandt, M.: *Versicherungsrecht*, 6. Aufl. (2016)
- Weingart, R.: Vertragsschluss durch Verhandlungsagenten, Wenn Software eigenständig im Rechtsverkehr auftritt. *CR* **36**, 701–708 (2020)
- Wilczek, F.: The unity of intelligence. In: Brockman, J. (Hrsg.) *Possible Minds, Twenty-Five Ways of Looking at AI*, S. 64–75. (2019)
- Zech, H.: Künstliche Intelligenz und Haftungsfragen. *ZfPW* **5**, 198–219 (2019)
- Zimmer, M., Mitzner, K.: Data Science und Künstliche Intelligenz in der Versicherungsbranche: Neue (?) Methoden für Versicherer. *BI-Spektrum* Heft 5, 36–41 (2019)