ALGORITHMENKONTROLLE

Positionspapier des Verbraucherzentrale Bundesverbands (vzbv)

02. Mai 2019

Impressum

Verbraucherzentrale
Bundesverband e.V.

Team
Digitales und Medien

Rudi-Dutschke-Straße 17
10969 Berlin

digitales@vzbv.de

Bundesverband der Verbraucherzentralen und Verbraucherverbände

Verbraucherzentrale Bundesverband e.V.

INHALT

I. ZUSAMMENFASSUNG

In immer mehr Lebensbereichen werden (selbstlernende) algorithmenbasierte[1] Entscheidungssysteme (Algorithmic Decision Making System, im Folgenden ADM-System) eingesetzt, um automatisiert Entscheidungen über Verbraucherinnen und Verbraucher[2] zu treffen. Eingeschlossen sind auch Entscheidungen, die überwiegend auf einer automatisierten Verarbeitung beruhen – wie beim Kreditscoring, bei dem sehr häufig rein formal der Bankangestellte am Ende die Entscheidung trifft. ADM-Systeme können große Auswirkungen auf Individuen und die Gesellschaft haben, stellen aber oft für Außenstehende eine Blackbox dar. Dennoch müssen auch in einer Welt selbstlernender Algorithmen die Einhaltung und Durchsetzung rechtlicher Regelungen sichergestellt sein, um deren Chancen zu nutzen und die Risiken einzudämmen. Solche können etwa hinsichtlich möglicher Diskriminierungen oder Irreführungen von Verbrauchern bestehen.

Hierfür sollte ein staatlich legitimiertes Kontrollsystem eingerichtet werden, mit dem Ziel, die Risiken von relevanten ADM-Prozessen zu minimieren. Dieses sollte sich durch Vielschichtigkeit auszeichnen und nicht aus einer einzigen Institution bestehen. Es sollte mehrere Elemente (zum Beispiel Verpflichtungen seitens der Betreiber, Zertifizierungsmöglichkeiten, Aufsichtsbehörden etc.) umfassen, deren Zusammenwirkung eine angemessene Kontrolle sicherstellen kann.
Nicht bei allen ADM-Systemen besteht im gleichen Umfang und Tiefe Regulierungsbedarf. Voraussetzung dafür, dass ein ADM-System einer regulatorischen Maßnahme unterworfen wird, sollte sein, dass das ADM-System gesellschaftlich relevante Auswirkungen hat, verbunden mit einem Schadenspotenzial für Individuen und/oder die Gesellschaft. Wird in einem ersten Schritt die Relevanz eines ADM-Systems festgestellt („Ob“), kann in einem zweiten Schritt die Bestimmung angemessener, fallspezifischer Regulierungsinstrumente („Wie“) erfolgen. Dabei sollte nach dem Ansatz verfahren werden, dass ein steigendes Risikopotenzial mit höheren Eingriffstiefen der regulatorischen Instrumente einhergeht (vgl. Abbildung 1). In diesem Sinne sollten Betreibern relevanter ADM-Prozesse je nach Risiko Verpflichtungen auferlegt werden, wie beispielsweise Transparenzvorgaben, Protokollierungspflichten, ADM-Folgeabschätzung, Audits / Zertifizierungen. Ungeachtet dessen sollten Aufsichtsbehörden als Teil des Kontrollsystems darüber hinaus stets die Befugnis haben, ADM-Systeme vollständig einzusehen und detailliert zu überprüfen („Inhaltskontrolle“). Betriebs- und Geschäftsgeheimnisse werden hierdurch nicht gefährdet, weil ein umfassendes staatlich legitimiertes Kontrollsystem zur vertraulichen Behandlung der Informationen verpflichtet ist.

---

[1] Ein Algorithmus ist zunächst nur eine festgelegte Handlungsanweisung, die auch „analog“ festgelegt und ausgeführt werden kann. Ein Beispiel sind die Straßenverkehrsordnung oder Gesetzesbücher. Vgl. Interview mit Sebastian Stiller in: Algorithmen treffen ins Mark der Macht. in: Politik & Kommunikation (2017), URL: https://www.politik-kommunikation.de/ressorts/artikel/algorithmen-treffen-ins-mark-der-macht-93003943 [Zugriff: 12.04.2019] sowie Zweig, Katharina: Was ist ein Algorithmus? in: AlgorithmWatch, URL: https://algorithmwatch.org/publication/arbeitspapier-was-ist-ein-algorithmus/ [Zugriff: 12.04.2019].

[2] Die gewählte männliche Form bezieht sich immer zugleich auf weibliche und männliche Personen. Wir bitten um Verständnis für den weitgehenden Verzicht auf Doppelbezeichnungen zugunsten einer besseren Lesbarkeit des Textes.

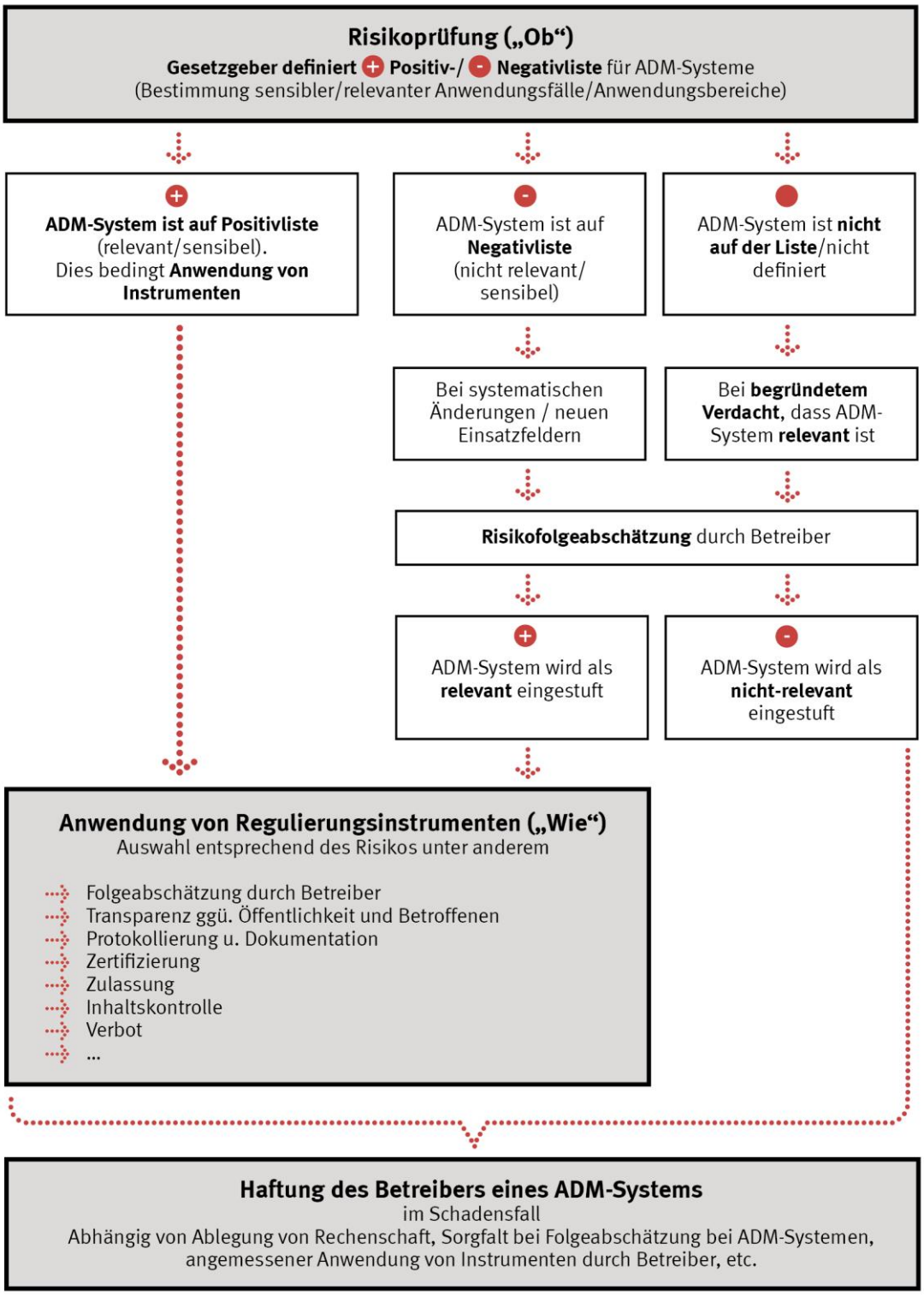

Abbildung 1: Risikoprüfung und Anwendung von Regulierungsinstrumenten innerhalb des Kontrollsystems

Im Folgenden werden die wichtigsten Forderungen des vzbv, welche für die Etablierung eines effektiven Kontrollsystems erforderlich sind, aufgeführt:

1. RISIKOPRÜFUNG VON ADM-SYSTEMEN ETABLIEREN

Der Gesetzgeber sollte mit Hilfe von Kriterien und einem Prüfschema definieren, welche ADM-Systeme mit einem hohen Schadenspotenzial für Individuen und/oder die Gesellschaft verbunden sein können. Die jeweiligen Anwendungsbereiche/-Fälle sollten einer konkretisierenden, aber nicht abschließenden Positiv- und Negativliste zugeordnet werden.

2. TRANSPARENZ SCHAFFEN

Transparenz über den Einsatz und die Arbeitsweise von ADM-Systemen sollte auf zwei Ebenen hergestellt werden. Zum einen gegenüber individuell Betroffenen und zum anderen gegenüber der Öffentlichkeit. Hierzu sollten Kennzeichnungs-, und Informationspflichten, Auskunftsrechte sowie die Verpflichtung der Veröffentlichung einer Risikofolgenabschätzung eingeführt werden, die entsprechend der Zielgruppe „individuell Betroffene“ und „Öffentlichkeit“ unterscheiden.

3. INHALTSKONTROLLE ERMÖGLICHEN

In besonders sensiblen Anwendungsfeldern (zum Beispiel im Gesundheitsbereich, beim automatisierten Fahren) sollte die Markteinführung von ADM-Systemen einer ex-ante Kontrolle, also einer Zulassung unterworfen werden.
Eine ex-post Kontrolle muss bei relevanten ADM-Systemen durch das Kontrollsystem fortlaufend möglich sein. Die Anwender von relevanten ADM-Systemen sollten verpflichtet sein, kontinuierlich selbst sicherzustellen, dass die von ihnen eingesetzten ADM-Systeme mit dem bestehenden Rechtsrahmen im Einklang stehen. Die von den Anwendern vorgenommenen Prüfungen müssen insbesondere für die zuständige Aufsichtsbehörde überprüfbar und nachvollziehbar sein. Im Falle eines Rechtsverstoßes sollten die Möglichkeiten der Aufsichtsbehörden von der Anpassung des ADM-Systems bis hin zum Verbot als Ultima Ratio reichen.

4. ANWENDUNGSBEREICH AUTOMATISIERTER ENTSCHEIDUNGEN IM EINZELFALL ERWEITERN

Der Anwendungsbereich für automatisierte Entscheidungen im Einzelfall (Artikel 22 DSGVO) sollte auf Entscheidungen ausgedehnt werden, die nicht nur auf einer ausschließlichen, sondern auch auf einer überwiegend automatisierten Verarbeitung von Daten beruhen. Um Sicherungsmechanismen einzuziehen und Fehler und Risiken von ADM-Systemen zu reduzieren, sollten rechtliche Vorgaben an diese Systeme dahingehend gemacht werden, dass die zugrundeliegenden und genutzten Daten für die Entscheidungsfindung erheblich sind. Die automatisierte Verarbeitung der Daten muss mit anerkannten mathematisch-statistischen Verfahren erfolgen. Die Prognosetauglichkeit, Validität und Reliabilität des mathematisch-statistischen Verfahrens sollte wissenschaftlich nachgewiesen werden können.

5. INSTITUTIONELLE AUSGESTALTUNG

Der Gesetzgeber sollte für die zuständigen Aufsichtsbehörden erforderliche Befugnisse (zum Beispiel Auskunfts-, Einsichts- und Zugangsrechte) etablieren, damit diese ADM-Systeme überprüfen und bewerten sowie Rechtsverstöße sanktionieren können. Darüber hinaus sollte eine Unterstützungseinheit etabliert werden, die die zuständigen sektorspezifischen Aufsichtsbehörden bei der Kontrolle von ADM-Systemen mit technisch-methodischer Expertise unterstützen kann.

6. NACHVOLLZIEHBARKEIT DURCH GESTALTUNG VON ADM-SYSTEMEN SICHERSTELLEN

Verbindliche Standards für die technische Gestaltung und Protokollierung, Dokumentation und Beschreibung von ADM-Systemen sind erforderlich, um diese einer Kontrolle zugänglich zu machen (Nachvollziehbarkeit-by-Design). Betreiber von relevanten ADM-Systemen sollten technische Schnittstellen vorhalten müssen, so dass die zuständigen Aufsichtsbehörden über diese, auf das System zugreifen können, um es auf ihre Rechtmäßigkeit sowie technische und methodische Fehler hin überprüfen zu können.

7. HAFTUNGSREGELN ANPASSEN

Betreiber von relevanten ADM-Systemen sollten für Schäden im Sinne einer Gefährdungshaftung bei bestimmungsgemäßer Verwendung durch den Verbraucher haftbar sein. Bei der Anpassung der Produkthaftungsrichtlinie müssen Schäden erfasst werden, die ADM-Systeme verursachen. Zurechnungs- und Beweisregelungen des Produkthaftungsrechts müssen neu justiert werden.

II. EINLEITUNG

Der zunehmende Einsatz von ADM-Systemen wirft etliche neue Fragen auf. Sie werden vermehrt eingesetzt, um Prozesse steuern und Entscheidungen über Verbraucherinnen und Verbraucher vorzubereiten und/oder zu treffen und sind daher in den letzten Jahren zunehmend in den Fokus der Politik und des Gesetzgebers gerückt.

Ein ADM-System umfasst weitaus mehr als den reinen Programmcode oder Algorithmus: „Algorithmische Entscheidungsfindung bezeichnet den Gesamtprozess von der Datenerfassung über die Datenanalyse bis hin zur Deutung und Interpretation der Ergebnisse und der Ableitung einer Entscheidung oder einer Entscheidungsempfehlung aus den Ergebnissen.“[3] Kennzeichnend für ADM-Systeme ist, dass sie eine algorithmische Komponente (Regelsystem) enthalten, die zu einem bestimmten Sachverhalt auf Basis einer Eingabe (Input) eine Entscheidung (Output) trifft und diese in Form eines (Zahlen-)Wertes ausgibt. Dies umfasst auch „lernende“ Systeme, die Entscheidungsre-

[3] Vgl. Vieth, Kilian; Wagner, Ben: Teilhabe, ausgerechnet (2017), URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/teilhabe-ausgerechnet [Zugriff: 16.04.2019].

geln mittels maschinellem Lernen aus Daten ableiten und diese im Zeitverlauf anpassen können.[4] Die unter dem Schlagwort Künstliche Intelligenz (KI) diskutierten Systeme fallen in der Regel unter diese Arbeitsdefinition.[5] ADM-Systeme können große Auswirkungen auf Individuen und die Gesellschaft haben, die sowohl Chancen bieten als auch Risiken bergen.[6] Es ist davon auszugehen, dass die Anzahl der Lebensbereiche und Anwendungskontexte, in denen diese Systeme eingesetzt werden, weiter zunimmt[7]. Hiervon sind Fragen der Lebensgestaltung, Teilhabemöglichkeiten, Konsumentscheidungen und Autonomie jedes Einzelnen sowie das gesellschaftliche Zusammenleben insgesamt betroffen. Damit verbunden werden grundlegende Fragen aufgeworfen, etwa zu Diskriminierung von Personen und Gruppen oder zu Autonomieverlust und Fremdbestimmung von Verbrauchern.
Beispiele für den Einsatz von ADM-Systemen sind die (personalisierte) Preissetzung von Gütern und Diensten[8], KI-basierte Gesundheitsratgeber[9], die Zuteilung von Sitzplätzen in Flugzeugen[10], die Bestimmung individueller Kreditausfallrisiken, Smart-

---

[4] Vgl. die Definition von ADM-Systemen von Zweig und Krafft: „Algorithmische Entscheidungssysteme (Algorithmic Decision Making Systems - ADM-Systeme, die) enthalten eine algorithmische Komponente, die - basierend auf der Eingabe - eine Entscheidung bzgl. eines Sachverhaltes trifft, d. h., die einen einzigen Wert berechnet. Wenn der Algorithmus von Experten erarbeitet wurde, spricht man von einem Expertensystem. Daneben gibt es solche, die das Regelsystem mit Hilfe von maschinellem Lernen aus Daten selbstständig ableiten.“Zweig, Katharina; Krafft, Tobias: Wie Gesellschaft algorithmischen Entscheidungen auf den Zahn fühlen kann, in: Mohabbat-Kar, Resa/Thapa, Basanta E.P/Parycek, Peter (Hrsg.): (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft, S. 471–492, Berlin, Kompetenzzentrum Öffentliche IT.

[5] Sie basieren in der Regel auch auf ADM-Systemen, da KI in der Regel eine algorithmische Komponente enthält, auf der das Regelsystem basiert, nach dem Entscheidungen getroffen werden. Das Regelsystem wird im Falle von KI oft mittels maschinellem Lernen aus Daten abgeleitet (etwa Machine Learning durch Neuronale Netze), kann aber auch fixe Regelsysteme enthalten, die Experten „manuell“ festgelegt haben. Vgl. dies. in: Stanford Encyclopedia of Philosophy (wie Anm. 5) sowie Zweig, Katharina; Krafft, Tobias: Transparenz und Nachvollziehbarkeit algorithmenbasierter Entscheidungsprozesse - Ein Regulierungsvorschlag aus sozioinformatischer Perspektive. Gutachten im Auftrag des Verbraucherzentrale Bundesverband, S. 9.

[6] Der Einsatz von ADM-Systemen kann die Teilhabe erhöhen, etwa über die Kostensenkung personalisierter Angebote und Dienste (vgl. Automatisierte Finanzberatung - Wenn der Algorithmus das Vermögen verwaltet. in: Frankfurter Allgemeine Zeitung (2016), URL: https://www.faz.net/aktuell/finanzen/fonds-mehr/automatisierte-finanzberatung-wenn-der-algorithmus-das-vermoegen-verwaltet-14384953.html [Zugriff: 16.04.2019]. Auch die Konsistenz von Entscheidungen kann verbessert werden, wenn ADM-Systeme immer nach den gleichen festgelegten Kriterien entscheiden und menschliche Fehler durch persönliche Wahrnehmung und Präferenzen reduziert werden. Andererseits können über die festgelegten Kriterien hinaus oft keine weiteren Kriterien einbezogen werden, so dass fraglich ist, ob Betroffene „menschliche“ Einzelfallentscheidungen erwarten können. Die mit ADM-Systemen assoziierten Risiken können unter anderem Sicherheitsrisiken, Gefährdung der Privatsphäre, Steigerung der Informationsasymmetrie zwischen Verbrauchern und Unternehmen, eingeschränkte materielle und soziale Teilhabe von Individuen und Gruppen (z. B. Diskriminierung) Fehlentscheidungen aufgrund von Bias sowie Manipulation bzw. unbewusste Beeinflussung individueller Entscheidungen, etwa durch Empfehlungssysteme, umfassen.

[7] Eine anschauliche Zusammenstellung von Beispielen für den zunehmenden und vielfältigen Einsatz von ADM-Systemen finden sich u.a. in AlgorithmWatch: Atlas der Automatisierung: Automatisierung und Teilhabe in Deutschland (2019), URL: https://atlas.algorithmwatch.org/ [Zugriff: 16.04.2019] sowie in Lischka, Konrad; Klingel, Anita. Bertelsmann Stiftung: Wenn Maschinen Menschen bewerten. Arbeitspapier. in: Bertelsmann Stiftung (2017), URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/wenn-maschinen-menschen-bewerten/ [Zugriff: 04.03.2019].

[8] Welt.de: Wer ein iPhone hat oder bei Chanel einkauft, zahlt bei Sixt mehr (2019), URL: https://www.welt.de/wirtschaft/article190490795/Sixt-Share-Wer-ein-iPhone-hat-zahlt-beim-Carsharing-mehr.html [Zugriff: 22.03.2019]

[9] Vgl. Ada Health GmbH, URL: https://ada.com/de/app/ [Zugriff: 30.04.2019]

[10] Zweig, Katharina; Krafft, Tobias: Transparenz und Nachvollziehbarkeit algorithmenbasierter Entscheidungsprozesse - Ein Regulierungsvorschlag aus sozioinformatischer Perspektive. Gutachten im Auftrag des Verbraucherzentrale Bundesverband, S. 5.

Home-Anwendungen, digitale Assistenzsysteme[11], Portfoliomanagement für Finanzanleger[12] sowie das autonome Fahren.

Der vzbv hat im Jahr 2017 das Thesenpapier „Algorithmenbasierte Entscheidungsprozesse“ veröffentlicht.[13] Das Ziel war, die verbraucherpolitische Diskussion zu diesem Thema weiter voranzutreiben. Das hier vorgelegte Positionspapier ist Ausfluss dieser Diskussion sowie zweier Gutachten, die der vzbv 2018 in Auftrag gegeben hat.[14]

Das Ziel muss sein, auch in einer Welt selbstlernender Algorithmen die Einhaltung und Durchsetzung rechtlicher Regelungen sicherzustellen: Etwa hinsichtlich des Allgemeinen Gleichstellungsgesetzes, dem Lauterkeitsrecht und dem Schutz personenbezogener Daten von Verbrauchern. Dafür müssen Aufsichtsinstitutionen in die Lage versetzt werden, ADM-Systeme auf ihre Rechtmäßigkeit hin zu überprüfen, um Rechtsverstöße sanktionieren zu können. Ebenso gilt es, die Entscheidungssouveränität der Verbraucher zu gewährleisten, Verbrauchervertrauen in ADM-Systeme durch Transparenz zu stärken sowie Wettbewerb und Innovation zu fördern.[15]
Das ist nur möglich, wenn ADM-Systeme kontrollierbar sind. Deshalb adressiert dieses Papier folgende Leitfragen und versucht, hierauf Antworten zu geben:

- Welche Rahmenbedingungen sind erforderlich, damit ADM-Systeme für Öffentlichkeit, Verbraucher, Aufsichtsbehörden angemessen nachvollziehbar und transparent sind?

- Wie sollte ein Kontrollsystem ausgestaltet sein, so dass eine effektive Einsicht und Überprüfung gesellschaftlich relevanter ADM-Systeme möglich ist?

- Welche Kontroll- und Eingriffsbefugnisse sind erforderlich, um Risiken von ADM-Systemen zu minimieren?

- Welche Grenzen[16] sollten für die Anwendung von ADM-Systemen gelten und wo sollten ADM-Systeme nicht eingesetzt werden dürfen?

[11] Etwa „Siri“ von Apple, „Alexa“ von Amazon oder der „Google Assistant“

[12] Frankfurter Allgemeine Zeitung (wie Anm. 6)

[13] Verbraucherzentrale Bundesverband: Algorithmenbasierte Entscheidungsprozesse. Thesenpapier des vzbv, URL: https://www.vzbv.de/pressemitteilung/keine-diskriminierung-durch-blackbox-algorithmen [Zugriff: 16.04.2019]

[14] Martini, Mario: Grundlinien eines Kontrollsystems für algorithmenbasierte Entscheidungsprozesse. Gutachten im Auftrag des Verbraucherzentrale Bundesverbandes (2018) Zweig, Katharina; Krafft, Tobias (wie Anm. 10)

[15] Dies sind keine widersprüchlichen Ziele, sondern sie greifen oft ineinander (beispielweise wenn Transparenz über die Verwendung von Nutzerdaten den Wettbewerb befördert, weil Nutzer sich alternativen datenschutzsensibleren Angeboten zuwenden).

[16] Beispielsweise Auflagen, Genehmigungsvorbehalt oder im letzten Schritt Verbot der Anwendung.

III. RELEVANZ VON ADM-SYSTEMEN („RISIKOPRÜFUNG“)

Die Risikoprüfung von ADM-Systemen sollte in mehreren Schritten erfolgen. Zunächst sollte der Gesetzgeber geeignete Regulierungsschwellen definieren. Maßstab hierfür sollte sein, ob ein ADM-System gesellschaftlich relevante Auswirkungen hat, verbunden mit einem Schadenspotenzial für Individuen und/oder die Gesellschaft. Dies sollte auf der Grundlage eines Prüfschemas ermittelt werden. Kriterien in dem Prüfschema könnten beispielsweise die Grundrechtssensibilität, die politische und ökonomische Macht der Betreiber des ADM-Systems, Abhängigkeit der Verbraucher für den Zugang zu einem Gut, Dienst oder Markt, Risiken der Diskriminierung oder die Anzahl der Betroffenen sein.[17]

Zur Konkretisierung und aus Gründen der Rechtssicherheit könnte der Gesetzgeber für bestimmte Sektoren/Anwendungen auf Basis dieses Prüfschemas eine konkretisierende, aber nicht abschließende Positiv- und Negativliste entwickeln.[18] Über die Positiv- und Negativliste werden ADM-Systeme entsprechend ihres Risikos als relevant beziehungsweise nicht-relevant kategorisiert. So kann es beispielsweise zweckmäßig sein, dass der Gesetzgeber festlegt, dass für bestimmte Märkte/Anwendungsfälle, die an sich ein hohes Risikopotenzial aufweisen, alle ADM-Systeme immer auf ihr spezielles Risikopotenzial hin geprüft werden oder gar einer Zulassung oder Vorabprüfung unterworfen werden müssen. Dies kann ADM-Systeme umfassen, die existenzielle Folgen für Verbraucher haben können (Kreditvergabe, Versicherungskonditionen), die mit Risiken für Leben und Gesundheit von Verbrauchern verbunden sind (beispielsweise autonomes Fahren, medizinische Anwendungen) und/oder die bereits einer Regulierung unterliegen (ADM-Systeme im Hochfrequenzhandel).

Eine konkretisierende Positiv- und Negativliste im Hinblick auf eine gestufte Regulierungskaskade birgt jedoch das Risiko, dass auf Systemanpassungen und neu entwickelte Anwendungen nicht oder nur verzögert reagiert werden kann. Das kann zum Beispiel der Fall sein, wenn eine Anwendung ursprünglich als risikoarm und somit auf der Negativliste geführt wurde und aufgrund von Systemanpassungen und/oder eines anderen Einsatzzwecks das Risiko im Nachhinein gestiegen ist. In diesen Fällen sollte bei begründetem Verdacht, dass ein ADM-System gesellschaftlich relevante Auswirkungen mit einem Schadenspotenzial für Individuen und/oder die Gesellschaft hat, die Anbieter/Anwender eines solchen Systems verpflichtet werden, eine Folgenabschätzung der Risiken zu erstellen. Entsprechende Regelungen bestehen bereits zum Beispiel im Bereich des Datenschutzes (vgl. Art. 35 DSGVO) und im Gefahrstoffrecht.[19] Aufgrund der Folgenabschätzung kann das ADM-System doch als relevant eingestuft

---

[17] Vgl. Martini, Mario (2018) (wie Anm. 14), S.44 ff. sowie den Entwurf eines Instruments zur Bestimmung des Wirkungspotenzials von digitalen Entscheidungssystemen von Vieth, Kilian; Wagner, Ben (2017) (wie Anm. 3)

[18] Martini, Mario (2018) (wie Anm. 14), S. 52 und 69.

[19] Ebd. Fn. S. 18. zu der Frage inwieweit Dienstanbietern, die grundrechtssensible oder potenziell gefährdende Algorithmen einsetzen eine thematisch umfassende Risikofolgeabschätzung abverlangt werden kann: „Hierbei kann die Rechtsordnung dem Anbieter einer Softwareanwendung durchaus auch die Ermittlung bisher unerkannter Gefahren für alle denkbaren Rechtsgüter überantworten; vgl. dazu die entsprechenden Regelungen im Gefahrstoffrecht: Art. 5 ff. VO (EU) Nr. 1272/2008.“

werden. Die Folgenabschätzung sollte entsprechend der o.g. Kriterien eine Abschätzung der Risiken beispielsweise für Grundrechte und mögliche Diskriminierung umfassen. Die Verpflichtung zur Veröffentlichung der Folgenabschätzung könnte dann eine Maßnahme zur Herstellung von Transparenz gegenüber der Öffentlichkeit sein (vgl. Abschnitt IV, 1.1 dieses Positionspapiers).

> Der Gesetzgeber sollte mit Hilfe einer Risikoprüfung definieren, welche ADM-Systeme mit einem hohem Schadenspotenzial für Individuen und/oder die Gesellschaft verbunden sein können. Die jeweiligen Anwendungsbereiche/-Fälle sollten in einer konkretisierenden, aber nicht abschließenden Negativliste und Positivliste aufgeführt werden. Erforderlich hierfür ist die Erarbeitung von Kriterien. Bei begründetem Verdacht, dass ein ADM-System Risiken für die Gesellschaft und/oder für Individuen birgt, sollten die Anbieter/Anwender verpflichtet werden, eine Folgenabschätzung der Risiken durchzuführen.

IV. REGULIERUNGSINSTRUMENTE

Kommt die Risikoprüfung zu dem Ergebnis, dass ein ADM-System relevant ist, bedarf es spezifischer Regulierungsinstrumente, um die ermittelten Risiken zu minimieren. Die Eingriffstiefe der Regulierungsinstrumente sollte dabei im Verhältnis zu der Höhe des prognostizierten Schadenspotenzials/Risikos eines ADM-Systems stehen: Ein hohes Schadenspotenzial geht mit einer größeren Eingriffstiefe einher.

Entsprechend ihres Risikos werden die ADM-Systeme in verschiedene Regulierungsklassen eingestuft. Damit könnten entsprechende Maßnahmen zur Schaffung von Transparenz und Nachvollziehbarkeit zum Einsatz kommen, die sukzessive tiefere Eingriffstiefen aufweisen.[20] Die Betreiber von ADM-Systemen niedriger Regulierungsklassen (relativ geringes Risiko) könnten beispielsweise verpflichtet werden, eine Folgeabschätzung zu veröffentlichen. Zudem könnte eine Kontrolle des Systems unter anderem mittels einer Black-Box Analyse erfolgen. Dabei würde die Aufsichtsbehörde keinen Einblick in das System an sich nehmen, aber über eine vom Betreiber bereitgestellte Schnittstelle testen, wie das System auf eingegebene Daten reagiert, beispielsweise ob bestimmte Gruppen systematisch diskriminiert werden.[21] Riskantere ADM-Systeme könnten in weiteren Stufen sukzessive zusätzlich eingriffsintensiveren Maßnahmen unterworfen werden: Etwa eine Pflicht zur Gewährung von Zugang zu Trainings- und Eingabedaten oder Möglichkeiten für Aufsichtsbehörden, die vom Betreiber im Rahmen der Folgeabschätzung angegebenen statistischen Qualitätsmaße selber zu

[20] Vgl. Zweig, Katharina; Krafft, Tobias (wie Anm. 10), S. 31ff. sowie Wischmeyer, Thomas: Regulierung intelligenter Systeme 143 (2018), in: Archiv des öffentlichen Rechts, Heft 1, S. 1–66., S. 63.

[21] Zweig, Katharina; Krafft, Tobias (wie Anm. 10) S. 36 sowie Diakopoulos, Nicholas: Algorithmic Accountability: On the Investigation of Black Boxes (2014), URL: https://www.cjr.org/tow_center_reports/algorithmic_accountability_on_the_investigation_of_black_boxes.php [Zugriff: 17.04.2019].

verifizieren.[22] In einer höheren Regulierungsklasse könnte zum Beispiel eine Einsichtnahme und die technisch-statistische Prüfung der verwendeten Datensätze erfolgen und das Lernverfahren etwa über einen Code Audit nachvollzogen werden.[23]

Bei der Bestimmung der Maßnahmen sollte zudem berücksichtigt werden, inwieweit die Verantwortlichen für ein ADM-System bereits selber Schritte ergriffen haben, um das mit dem ADM-System verbundene Risiko zu verringern. Dies könnten technische, rechtliche, organisatorische, oder informationelle Maßnahmen sein.[24]

1. TRANSPARENZ

Transparenz über den Einsatz und die Arbeitsweise von ADM-Systemen sollte auf zwei Ebenen hergestellt werden. Zum einen gegenüber individuell Betroffenen und zum anderen gegenüber der Öffentlichkeit. Verbraucher können zum Beispiel betroffen sein durch den fortschreitenden Verlust menschlicher Autonomie (durch beispielsweise Entscheidungsassistenten). Gesellschaftliche Folgen können zum Beispiel durch Gruppendiskriminierungen drohen (Bewerberauswahl oder Wohnungsvergabe). Deshalb müssen Kennzeichnungs-, und Informationspflichten sowie Auskunftsrechte eingeführt werden, die entsprechend der Zielgruppe „individuell Betroffene“ und „Öffentlichkeit“ unterscheiden.

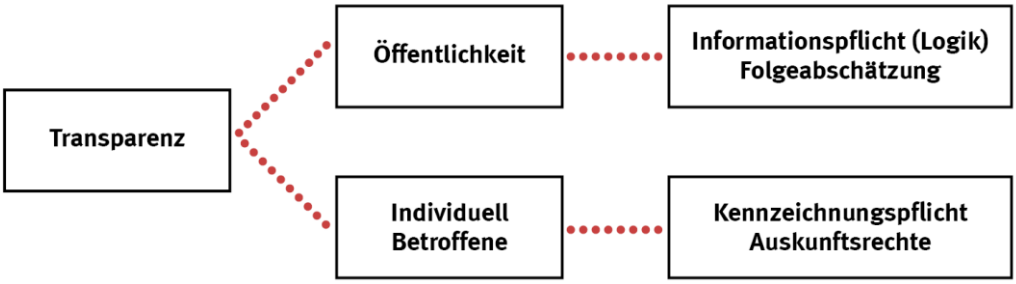

Abbildung 1: Übersicht Transparenz gegenüber Öffentlichkeit und Betroffenen

1.1 Transparenz gegenüber der Öffentlichkeit

Informationspflichten

Informationspflichten sind eine wesentliche Grundvoraussetzung, um Transparenz herzustellen und Vertrauen in Systeme zu schaffen, die auf algorithmischen Entscheidungen basieren. Nur so können sich die interessierte Öffentlichkeit und potentiell Betroffene ein Bild über die Konzeption, Reichweite und mögliche Auswirkungen einer Anwendung machen und diese in Grundzügen bewerten. Diese Informationen sind darüber hinaus für Verbraucher erforderlich, die vor der Wahl stehen, ob sie einen Dienst verwenden wollen. Auch helfen ihnen diese Informationen, um die Folgen einer möglichen datenschutzrechtlichen Einwilligung abschätzen zu können.

---

[22] Zweig, Katharina; Krafft, Tobias (wie Anm. 10) S. 37ff.

[23] Gesellschaft für Informatik: Technische und rechtliche Betrachtungen algorithmischer Entscheidungsverfahren. Gutachten der Fachgruppe Rechtsinformatik der Gesellschaft für Informatik e.V. im Auftrag des Sachverständigenrats für Verbraucherfragen, Berlin, URL: www.svr-verbraucherfragen.de/wp-content/uploads/GI_Studie_Algorithmenregulierung.pdf [Zugriff: 07.03.2019] S. 64ff.

[24] Martini, Mario (2018) (wie Anm. 14), S. 64f..: „An die grobe Risikoklassenzuordnung kann sich – auf einer *zweiten Stufe* – eine Feinjustierung anschließen: Mit Hilfe organisatorischer, technischer und rechtlicher Maßnahmen kann der Verantwortliche die Grundrechtssensibilität der Softwareanwendung verringern.“

Daher müssen Verantwortliche entsprechend der DSGVO über das Bestehen einer (voll-)automatisierten Entscheidungsfindung informieren und aussagekräftige Informationen über die involvierte Logik sowie die Tragweite und die angestrebten Auswirkungen einer derartigen Verarbeitung auf potentiell Betroffene bereitstellen.

Problematisch ist jedoch, dass diese Vorgaben auf Artikel 22 DSGVO Bezug nehmen, der nach vorherrschender Meinung lediglich Fälle einer *vollständig* automatisierten Entscheidungsfindung umfasst, die der betroffenen Person gegenüber rechtliche Wirkung entfaltet oder eine andere erhebliche Auswirkung auf ihn hat. Über ADM-Systeme, die lediglich einer Entscheidungsvorbereitung dienen, müsste somit grundsätzlich nicht in diesem Detailgrad (über die üblichen Informationspflichten bei der Verarbeitung personenbezogener Daten hinaus) informiert werden. Hier bedarf es einer Erweiterung des Anwendungsbereichs von Artikel 22 DSGVO (vgl. hierzu Abschnitt 3.2).

Auch ist fraglich, inwieweit sich die genannten Informationspflichten der DSGVO auch auf ADM-Systeme beziehen, die dem Betroffenen gegenüber *keine* rechtliche Wirkung entfalten oder eine andere erhebliche Auswirkung auf ihn hat. So schränken Artikel 13 und 14 DSGVO den Bezug zu Artikel 22 DSGVO dahingehend ein, dass die Informationen „zumindest in diesen Fällen“ (des Artikel 22) gegeben werden müssen, was auf einen weiten Anwendungsbereich schließen lässt. Auch Erwägungsgrund 60 betont, dass die betroffene Person darauf hingewiesen werden sollte, dass Profilbildung generell stattfindet und welche Folgen dies hat.

> Die datenschutzrechtlichen Informationspflichten (Art. 13 Abs. 2 lit. f und Art. 14 Abs. 2 lit. g DSGVO) sollten dahingehend konkretisiert werden, dass Betroffene stets unterrichtet werden sollten, wenn betreffende ADM-Prozesse vorgenommen werden und welche Folgen diese haben – selbst wenn diese keine rechtliche Wirkung für die Betroffenen entfalten oder andere erhebliche Auswirkungen auf sie haben. Des Weiteren sollten sich die Informationspflichten – analog zur geforderten Ausweitung des Artikel 22 DSGVO – auch auf Entscheidungen beziehen, die auch überwiegend – und nicht nur vollständig – auf einer automatisierten Verarbeitung beruhen.

Folgenabschätzung

Die Folgenabschätzung i.S.d. Art. 35 Abs. 1 DSGVO umfasst ausschließlich Informationen zu den Folgen bezüglich des Schutzes personenbezogener Daten und keine umfassende Risikoanalyse eines ADM-Systems.[25] Bei ADM-Systemen mit einem signifikanten Risikopotenzial ist es aber sachgerecht und zumutbar, dem Anbieter/Anwender eine umfassende Folgenabschätzung abzuverlangen zur Einschätzung des mit einem

[25] Ebd. S. 17-18 „Sonstige Schutzgüter – wie etwa das Vermögen, das Eigentum oder die körperliche Integrität – muss der Verantwortliche demgegenüber nicht zwangsläufig und unmittelbar in seinen Prüfradar integrieren […]. Das Regelungskonzept der DSGVO springt dadurch zu kurz.“

System verbundenen Risikos[26]. Entsprechende Regelungen bestehen bereits im Gefahrstoffrecht[27] Die Folgenabschätzung sollte eine Abschätzung der Risiken für Vermögen, Eigentum, körperliche Integrität und Diskriminierung umfassen.

Eine Bewertung von ADM-Systemen kann nur erfolgen, wenn die grundlegenden Eigenschaften und potenziellen Risiken öffentlich zugänglich sind. Dafür muss etwa bekannt sein, auf welche allgemeine Datengrundlage ein Machine-Learning-basiertes ADM-System trainiert wurde, welche Datenkategorien einbezogen werden, auf welche Kriterien das System optimiert. Sie sollte neben Informationen zu den zugrundeliegenden Daten und Logik des Modells auch Qualitätsmaße und Fairnessmaße zu den Daten und Modellgüte enthalten, etwa zu Bias oder (statistischen) Fehlerquoten[28] (insgesamt oder für bestimmte Teilgruppen), die ein System bei der Vorhersage/Kategorienbildung aufweist.

Diese Informationen sollten in Form einer ADM-Folgeabschätzung in einem öffentlichen Register veröffentlicht werden.

> Anbieter von relevanten ADM-Systemen, sollten eine umfassende Folgenabschätzung erstellen und veröffentlichen müssen.

1.2 Transparenz gegenüber individuell Betroffenen

Bei bestimmten ADM-Systemen können Transparenzanforderungen gegenüber Betroffenen erforderlich sein, um die Entscheidungssouveränität von Verbrauchern zu wahren und bewusste Konsumentscheidungen zu ermöglichen.

Kennzeichnungspflichten

Bei relevanten ADM-Systemen, die wichtige Entscheidungen über Verbraucher treffen oder vorbereiten, sollte eine Kennzeichnungspflicht eingeführt werden, die Betroffene darauf hinweist, dass Entscheidungen automatisiert durch ein ADM-System getroffen werden.

---

[26] Ebd., S. 17f., Vgl. ebenso den Entwurf einer umfassenden Folgeabschätzung vorgelegt von Diakopoulos, Nicholas u. a.: Principles for Accountable Algorithms and a Social Impact Statement for Algorithms, URL: http://www.fatml.org/resources/principles-for-accountable-algorithms [Zugriff: 17.04.2019].

[27] Martini, Mario (2018) (wie Anm. 14) S. 17-18, zur Frage inwieweit es sachgerecht ist, „Diensteanbietern, die grundrechtssensitive oder potenziell gefährdende Algorithmen einsetzen, eine thematisch umfassende Risiko- bzw. Folgenabschätzung abzuverlangen, bevor sie ihre Softwareanwendungen am Markt einsetzen“ (S.18) „Die Rechtsordnung kann dem Anbieter einer Softwareanwendung dabei durchaus auch überantworten, bisher unerkannte Gefahren für alle betroffenen Rechtsgüter zu ermitteln; vgl. dazu bspw. die entsprechenden Regelungen im Gefahrstoffrecht: Art. 5 ff. VO (EU) Nr. 1272/2008“ (Fn 53, S.18).

[28] So kann es sein, dass ein ADM-System für bestimmte soziale Gruppen höhere Fehlerquoten aufweist, als für andere, wodurch diese sytematisch benachteiligt werden können. Vgl. Angwin, Julia u. a. https://www.facebook.com/propublica: Machine Bias. in: ProPublica (2016), URL: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [Zugriff: 17.04.2019] für den Fall individueller Rückfälligkietsprognosen. Eine Beschreibung von Fairnessmassen liefern Zweig, Katharina; Krafft, Tobias: Fairness und Qualität algorithmischer Entscheidungen, in: Mohabbat-Kar, Resa/Thapa, Basanta E.P/Parycek, Peter (Hrsg.): (Un)berechenbar? Algorithmen und Automatisierung in Staat und Gesellschaft, S. 204–227, Berlin, Kompetenzzentrum Öffentliche IT.

Mittlerweile sind zum Beispiel einige digitale Sprachassistenten, die selbständig Anrufe tätigen können, so hoch entwickelt, dass viele Menschen am Telefon nicht erkennen, dass es sich bei dem Anrufer um eine Maschine handelt[29]. Dabei können diese Systeme in Echtzeit Emotionen von Menschen erkennen und auswerten[30]. Menschen sollten darauf aufmerksam gemacht werden müssen, dass sie mit einem ADM-System interagieren und dies in eine Entscheidung einfließt. Nur so können sich Verbraucher beispielsweise für oder gegen einen Dienst entscheiden und/oder vor der Inanspruchnahme zusätzliche Informationen einholen und sich gegen Irreführungen wehren.[31]

> Es sollte eine Kennzeichnungspflicht für relevante ADM-Systeme eingeführt werden. Die Kennzeichnung sollte dabei durch einen eindeutigen Hinweis zu Beginn der Interaktion mit dem Verbraucher erfolgen beziehungsweise durch verständliche Symbole sichergestellt werden.

Auskunftsrechte

Auskunftsrechte sind für Verbraucher zentral, um algorithmenbasierte Entscheidungen nachvollziehen und individuell überprüfen zu können. Nur so können sie ihre – beispielsweise in der DSGVO festgelegten – Rechte wahrnehmen und eine Entscheidung fundiert anfechten. Wesentlich ist dabei, dass Verbraucher verständlich, relevant und konkret informiert werden und ihnen das Ergebnis für den Einzelfall erläutert wird (anders als bei der Informationspflicht, bei der die Funktionsweise eines ADM-Systems allgemein dargelegt wird).

Insbesondere besteht durch die strukturellen Eigenarten von ADM-Systemen das Risiko von Fehlschlüssen der Scheinkausalität, wodurch ein besonderes Schutzbedürfnis ausgelöst wird.[32] Daher bedarf es einer Pflicht zur Erläuterung der Entscheidungsergebnisse vor allem dann, wenn die Entscheidung eine solche grundrechtssensible Tragweite hat, “vor allem, aber nicht nur, wenn der Entscheidung in persönlichkeitsrechtlich sensiblen Feldern rechtliche Wirkung zukommt“[33]. Diejenigen, die ADM-Systeme oder deren Ergebnisse einsetzen, berufen sich nicht selten darauf, dass sie über Ergebnisse eines ADM-Systems ihre eigene freie Willens- und Vertragsentscheidung definieren. Aber auch dann muss weiterhin die Kontrolle rechtskonformen Handelns möglich sein. Das Grundrecht auf allgemeine Handlungsfreiheit wird begrenzt durch verbotene Diskriminierung (AGG), durch andere Grundrechte und durch die – auch vertragsspezifischen - Vorgaben der Rechtsordnung.

---

[29] Spiegel Online: Nach Wirbel um Sprachassistent: Google Duplex gibt sich jetzt zu erkennen (2018), URL: https://www.spiegel.de/netzwelt/web/google-duplex-gibt-sich-jetzt-am-telefon-zu-erkennen-a-1215453.html [Zugriff: 17.04.2019]

[30] Somers, Meredith: Emotion AI, explained. in: MIT Sloan School of Management (2019), URL: https://mitsloan.mit.edu/ideas-made-to-matter/emotion-ai-explained [Zugriff: 17.04.2019]; Stimme: Die Seele auf der Zunge. in: Die Zeit (2019), URL: https://www.zeit.de/2019/07/stimme-biometrie-messbarkeit-emotionen-persoenlichkeit [Zugriff: 17.04.2019]

[31] Mittlerweile betreffen die meisten Verbraucherbeschwerden bei der U.S. Aufsichtsbehörde für Telekommunikation FCC sogenannte "Robocalls" also Anrufe durch Maschinen, die Verbraucher automatisiert anrufen, teils in betrügerischer Absicht. Federal Communications Commission: Stop Unwanted Robocalls and Texts (2019), URL: https://www.fcc.gov/consumers/guides/stop-unwanted-robocalls-and-texts [Zugriff: 14.03.2019]

[32] Martini, Mario (2018) (wie Anm. 14) S. 13

[33] Ebd. S. 13

Bei der Kontrolle durch Betroffene selbst kommt es darauf an, nicht nur die verwendeten Daten auf zum Beispiel Richtigkeit und Vollständigkeit zu überprüfen, sondern auch, ob die verwendeten Korrelationen in einem zutreffenden kausalen Sachzusammenhang zum Individuum stehen können. Vereinfacht müssen Verbraucher erkennen, welche Informationen und damit verbundenen Annahmen wesentlich und damit prägend für eine Entscheidung waren. Vor allem dann, wenn das Ergebnis vom selbst Erwarteten abweicht, können auch nur dann die Anfechtungsrechte nach Art. 22 DSGVO sinnvoll wahrgenommen werden

Ein sinnvoller Ansatz könnten „counterfactual explanations“ sein, bei denen Betroffenen im Falle einer negativen Entscheidung die wichtigsten Gründe für die Entscheidung genannt werden. Diese können auch bei komplexen Systemen (etwa Neuronalen Netzwerken)erfolgen[34].

> Wenn eine Entscheidung (oder eine überwiegend auf einer automatisierten Verarbeitung beruhende Entscheidung) eines ADM-Systems in grundrechtssensiblen Feldern (zum Beispiel Leib/Leben, Eigentum, Recht auf Schutz personenbezogener Daten) eine rechtliche Wirkung oder eine andere erhebliche Auswirkung auf den Betroffenen zur Folge hat, sollte dieser zusätzlich zum datenschutzrechtlichen Auskunftsrecht nach Artikel 15 DSGVO das Recht erhalten, dass ihm der Verantwortliche das Ergebnis für den Einzelfall erläutern und ihm die zugrundeliegenden Daten, sowie ihre Gewichtung bei der Berechnung konkret in einer nachvollziehbaren Form offenlegen muss.

2. INHALTSKONTROLLE

2.1 Ex-ante Kontrolle in sensiblen Anwendungsfeldern

Es gibt Anwendungen, in denen ADM-Systeme große – teilweise irreversible - Schäden für Verbraucher verursachen können, etwa durch nicht adäquate Modelle, Datengrundlage oder falsche Grundannahmen. Die Schäden können dabei unterschiedlicher Natur sein, etwa finanziell, immateriell oder physisch. Beispiele umfassen Anwendungen, die potenziell schwerwiegende Grundrechtsverletzungen, Risiken für Leben und Gesundheit von Verbrauchern nach sich ziehen können (beispielsweise wie Pflegeroboter oder Mobilitätsanwendungen). Eine Zulassung ist in regulierten Märkten für manche sensiblen Anwendungen, Dienste und Produkte bereits übliche Praxis[35].

> In besonders sensiblen Anwendungsfeldern sollte die Markteinführung von ADM-Systemen einer Zulassung oder Vorabprüfung durch Aufsichtsinstitutionen unterworfen werden.

---

[34] Wachter, Sandra; Mittelstadt, Brent; Russell, Chris: Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR (2017), in: SSRN Electronic Journal, URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3063289## [Zugriff: 17.04.2019]

[35] Vgl. Busch, Christoph: Algorithmic Accountability. ABIDA Gutachten 01IS15016A. in: ABIDA - Assessing Big Data (2018), URL: http://www.abida.de/en/node/397

2.2 Ex-post Kontrolle

Bei relevanten ADM-Systemen muss die Möglichkeit einer fortlaufenden Kontrolle durch das Kontrollsystem sichergestellt werden. So sollten die Betreiber von relevanten ADM-Systemen, kontinuierlich prüfen müssen, dass das System in Einklang mit dem geltenden Recht steht. Dies gilt insbesondere für „lernende“ Systeme, deren Entscheidungsregeln kontinuierlichen Änderungen unterworfen sind.

Dies beinhaltet auch, dass insbesondere die zuständigen Aufsichtsbehörden jederzeit die Möglichkeit haben müssen, ein ADM-System etwa hinsichtlich seiner Rechtmäßigkeit hin zu überprüfen. Dies umfasst etwa die Überprüfung der Trainingsdaten und verwendeten Lernverfahren, das finale Regelmodell sowie die verwendeten Inputdaten und die Outputdaten.

Sollte bei einer Überprüfung eine Einsicht in den Programmcode erforderlich sein, sollte dieser in einer Form zugänglich gemacht werden, die vollständig funktionstüchtig ist und - etwa über entsprechende Protokollierung, Dokumentation und Kommentierung bewertbar ist.[36]

> Der Gesetzgeber sollte festlegen, dass die Anwender von relevanten ADM-Systemen kontinuierlichen Kontrollpflichten unterliegen, um selbst sicherzustellen, dass die von ihnen eingesetzten ADM-Systeme mit dem bestehenden Rechtsrahmen im Einklang stehen. Die von den Anwendern vorgenommenen Kontrollen müssen insbesondere für die zuständige Aufsichtsbehörde überprüfbar und nachvollziehbar sein.

Anpassungen von Bestandteilen eines ADM-Systems

Wenn die Datenbasis, der Algorithmus oder andere Bestandteile, die einem ADM-System zugrunde liegen, so gestaltet sind, dass die Ergebnisse gegen rechtliche Vorgaben verstoßen, müssen diese Bestandteile angepasst werden. Dies kann beispielsweise der Fall sein, wenn durch eine Verzerrung der Datenbasis oder durch die Entscheidungskriterien des Algorithmus die Ergebnisse des ADM-Prozesses systematisch bestimmte Personengruppen im Sinne des Allgemeinen Gleichbehandlungsgesetzes (AGG) benachteiligen. So ist es möglich, dass ADM-Systeme, die Unternehmen bei der Vorauswahl von Stellenbewerbern einsetzen, gegen Bestimmungen des AGG verstoßen, beispielsweise, wenn Frauen systematisch benachteiligt werden.[37]

---

[36] Zweig, Katharina; Krafft, Tobias (wie Anm. 10), S. 40

[37] Vgl. Heise.de: Amazon: KI zur Bewerbungsprüfung benachteiligte Frauen (2018), URL: https://www.heise.de/newsticker/meldung/Amazon-KI-zur-Bewerbungspruefung-benachteiligte-Frauen-4189356.html [Zugriff: 17.04.2019] Durch das gezielte Ausspielen von Stellenanzeigen können beispilesweise Frauen bestimmte Stellenagenboten systemtisch vorenthalten werden. Vgl. dazu Di Sherwin, Galen. American Civil Liberties Union: How Facebook Is Giving Sex Discrimination in Employment Ads a New Life (2018), URL: https://www.aclu.org/blog/womens-rights/womens-rights-workplace/how-facebook-giving-sex-discrimination-employment-ads-new [Zugriff: 17.04.2019]. Ähnlich können durch zielgruppenspezifische Onlinewerbung ethische Minerheiten auf dem Wohnungsmarkt gezielt diskriminiert werden, vgl. U.S. Department of Housing and Urban Development: HUD Charges Facebook with Housing Discrimination Over Company's Targeted Advertising Practices (2019), URL: https://www.hud.gov/press/press_releases_media_advisories/HUD_No_19_035 [Zugriff: 29.03.2019]

> Aufsichtsbehörden sollten für den Fall, dass die Ergebnisse eines ADM-Systems gegen geltendes Recht verstoßen, Betreibern Verpflichtungen (zum Beispiel Anpassung der Datenbasis, Kriterien) auferlegen können, um das ADM-System rechtskonform anzupassen. Erfolgt diese Anpassung nicht, müssen zumindest die gegen geltendes Recht verstoßenden Bestandteile aus dem Verkehr gezogen werden.

Verbote als Ultima Ratio

Ein gesetzliches Verbot des Einsatzes bestimmter ADM-Systeme kann als Ultima Ratio ein gerechtfertigtes Mittel sein. So verbietet das Bundesdatenschutzgesetz, dass ausschließlich Adressdaten zur Erstellung eines Score-Wertes (beispielsweise zur Schätzung der Zahlungsausfallwahrscheinlichkeit eines Verbrauchers) verwendet werden dürfen. Auch kann die Börsenaufsichtsbehörde die Nutzung einer algorithmischen Handelsstrategie für den Hochfrequenzhandel an Börsen untersagen, um Verstöße gegen börsenrechtliche Vorschriften zu verhindern und um Missstände im Handel zu beseitigen. Dies kann der Fall sein, wenn ADM-Systeme der Marktmanipulationen dienen; etwa indem durch große Auftragsvolumina und Stornierungen Handelssysteme verlangsamt werden oder ein falsches Bild von Nachfrage und Angebot vorgetäuscht wird.[38]

> Eine Untersagung oder ein gesetzliches Verbot des Einsatzes bestimmter ADM-Systeme kann in bestimmten Fällen als Ultima Ratio ein gerechtfertigtes Mittel sein.

2.3 Kontrollverfahren

Kontrollalgorithmen zur Ergebniskontrolle

Die Prüfung der Rechtmäßigkeit von ADM-Systemen kann beispielsweise die Analyse großer Datenmengen, die Prüfung der Gewichtung von Faktoren in komplexen multidimensionalen Modellen sowie eine Input-Output-Analyse beinhalten. Zudem ist eine Prüfung erforderlich, „ob die behauptete Beziehung zwischen Sachverhalt und Ergebnis mit dem tatsächlichen Entscheidungsverhalten übereinstimmt oder nicht“[39] – also ob das dem ADM-System zugrunde liegende Entscheidungsmodell angemessen gestaltet und valide ist. Dabei kann, aufgrund der Komplexität der Materie und involvierten Datenmengen, der Einsatz von Kontrollalgorithmen die Effizienz und Effektivität der Überprüfung erheblich steigern. So können Kontrollalgorithmen systematisch nach auffälligen Mustern in der Datenbasis und den Ergebnissen eines ADM-Systems suchen, die beispielsweise Aufschluss über eine ungerechtfertigte Diskriminierung im Sinne des AGG geben können.

> Der Gesetzgeber muss rechtlich absichern, dass Kontrollalgorithmen bei der Prüfung der Rechtmäßigkeit von ADM-Systemen zum Einsatz kommen können.

---

[38] Vgl. Strafbare Marktmanipulation: Raub durch Hochfrequenzhandel. in: Frankfurter Allgemeine Zeitung (2015), URL: https://www.faz.net/aktuell/finanzen/fonds-mehr/hochfrequenzhandel-an-boersen-anfaellig-fuer-marktmanipulation-13368033.html [Zugriff: 17.04.2019] sowie Bundesanstalt für Finanzdienstleistungsaufsicht - BaFin: Algorithmischer Handel und Hochfrequenzhandel, URL: https://www.bafin.de/DE/Aufsicht/BoersenMaerkte/Handel/Hochfrequenzhandel/high_frequency_trading_artikel.html [Zugriff: 17.04.2019].

[39] Martini, Mario (2018) (wie Anm. 14), S. 27.

Test- und Auditverfahren für ADM-Systeme

Test- und Auditverfahren können dazu dienen, Fehler oder Biases in ADM-Systemen und deren Daten Grundlage zu identifizieren. Diese können bewusst oder unbewusst Eingang in das System gefunden haben und in ihrer Konsequenz auch Rechtsverstöße darstellen. Etwa, wenn sie dazu führen, dass die Entscheidungen, die auf Basis des ADM-Systems getroffen werden, Verstöße gegen Diskriminierungsverbote oder Irreführungen von Verbrauchern darstellen. Bei Testergebnissen ist nicht immer eindeutig ob sie wirklich einen Fehler eines ADM-Systemen aufzeigen, was ihre Beweisfunktion einschränkt. Es bedarf einer Festlegung welchen Testverfahren und Ergebnissen welche Beweislast zukommt.[40] Dafür könnten beispielsweise technisch-statistische Standards für die Qualität der Testverfahren und Ergebnisse festgelegt werden. Diese sollten je nach Anwendungsbereich spezifisch erfolgen.[41] Zudem ist nicht immer klar, welche Rechtsfolgen ausgelöst werden, wenn Tests Rechtsvorschriften feststellen.[42] Eine Überprüfung eines ADM-Systems kann auch in Form von Auditverfahren nach bestimmten Standards erfolgen, die wiederum festlegt werden müssen[43]. Auch die Gestaltung der Auditstandards sollten dem spezifischen Anwendungsbereich des ADM-Systems Rechnung tragen.

> Damit im Rahmen einer Überprüfung Tests und Audits ihre Beweisfunktion ausfüllen können, sollte der Gesetzgeber Rahmenbedingungen schaffen, die den Einsatz von technisch-statistischen Tests und Audits rechtlich absichern. Flankierend müssen technisch-statistische Standards für die Qualität der Testverfahren sowie Verfahrensstandards für Audits erarbeitet werden.

Zertifizierungen

Zertifizierungsverfahren können ein wirkungsvolles Instrument sein, um sicherzustellen, dass ADM-Systeme rechtliche Rahmenbedingungen oder anderweitig definierte Standards einhalten[44]. Damit können Zertifikate die (Rechts-)Sicherheit sicherstellen und als Qualitätsmerkmal Anwendern und Verbrauchern Orientierung bieten. In besonders risikobehafteten Anwendungsbereichen oder -fällen kann eine verpflichtende Zertifizierung sinnvoll sein. Diese würde sicherstellen, dass ADM-Systeme ein Mindestmaß an Sicherheit und Qualität verbürgen.

Die zentrale Herausforderung für alle Zertifizierungsverfahren ist die Definition von Standards, anhand derer eine Auditierung erfolgen kann. Die Standards müssen dabei ausreichend konkret und hoch genug sein, um in aussagekräftige Zertifikate zu münden. Zudem zeigt die Erfahrung, dass durch das Entstehen einer Vielzahl an Zertifikaten im Markt diese ihre Orientierungsfunktion für Verbraucher einbüßen, zumal wenn sie auf intransparenten Standards basieren.

---

[40] Beispielsweise hinsichtlich der statistischen Signifikanz. Zum rechtlichen Rahmen von Testverfahren für ADM-Systeme vgl. Gesellschaft für Informatik (wie Anm. 23) S. 146ff.

[41] Beispielsweise ist es denkbar, dass an Anwendungen im medizinischen Bereich, die Gesundheit und Leben von Patienten betreffen, andere Ansprüche gestellt werden sollten, als an Anwendungen, die Kaufempfehlungen aufbereiten.

[42] Gesellschaft für Informatik (wie Anm. 23)

[43] Vgl. Standards zur Sicherstellung von Qualität im Bereich Software Engineering unter Software quality assurance. in: Wikipedia (2019), URL: https://en.wikipedia.org/w/index.php?oldid=880438786 [Zugriff: 17.04.2019]

[44] Gesellschaft für Informatik (wie Anm. 23) S. 8

> Der Gesetzgeber sollte Rahmenbedingungen für die Etablierung von Zertifizierungsverfahren von ADM-Systemen schaffen. In besonders risikobehafteten Anwendungsbereichen oder -fällen muss die Verpflichtung zur Zertifizierung möglich sein.

In-Camera Verfahren zur Wahrung von Geschäftsgeheimnissen

Im Falle einer Überprüfung von ADM-Systemen durch Aufsichtsbehörden stellt sich in der Regel die Frage nach der Wahrung von Geschäftsgeheimnissen nicht, da diese der Geheimhaltungspflicht unterliegen. In diesem Zusammenhang ist zu diskutieren, ob und wie die Überprüfung auf andere unabhängige, zum Stillschweigen verpflichtete Experten als Teil des Kontrollsystems übertragen werden kann.[45] Zur legitimen Wahrung von Geschäftsgeheimnissen könnte bei der Überprüfung von ADM-Systemen in Anlehnung an die in anderen Prozessen (zum Beispiel Patentstreitigkeiten) bewährten In-Camera-Verfahren[46] angewendet werden. Dabei würde der Betreiber eines ADM-Systems im Falle der Überprüfung sensitive Informationen nicht der Öffentlichkeit oder anderen Nicht-Berechtigten zur Verfügung stellen, sondern einer „Kammer“ von Experten, die zur Geheimhaltung verpflichtet wären. Diese würden dann die Prüfung des ADM-Systems vornehmen.

> Der Gesetzgeber sollte In-Camera-Verfahren zur Einsichtnahme und Überprüfung von ADM-Systemen etablieren, die legitime Geheimhaltungsinteressen der Betreiber oder sensitive Informationen dritter Betroffener wahren, aber gleichzeitig eine fundierte, unabhängige und umfassende Prüfung der Rechtskonformität eines ADM-Systems gewährleisten.

3. MATERIELLRECHTLICHER ANPASSUNGSBEDARF

Erweiterung des Allgemeinen Gleichbehandlungsgesetzes (AGG)

ADM-Systeme können dazu führen, dass bestimmte Personengruppen systematisch unterschiedlich behandelt werden. Die unterschiedliche Behandlung könnte beispielsweise auf unterschiedlicher (gruppenspezifischer) Ausprägung mancher Variablen in den zugrundeliegenden Daten basieren. So werden Frauen typischerweise andere Konsummuster aufweisen als Männer. Diese können beispielsweise herangezogen werden, um Rückschlüsse auf unterschiedliche Zahlungsbereitschaft oder Risikoaffinität von Frauen und Männern zu ziehen und entsprechende Ungleichbehandlungen vorzunehmen.

Das *Allgemeine Gleichbehandlungsgesetz* (AGG) soll Personen vor Diskriminierung aufgrund ihrer Rasse, ethnischen Herkunft, des Geschlechts, der Religion oder Weltanschauung, einer Behinderung, des Alters oder der sexuellen Identität schützen[47]. Sein

---

[45] Vgl. den Abschnitt „Außerbehördlicher Kontrollmechanismus: Übertragung der Aufsichtstätigkeit an Beliehene“ im Kapitel „Institutionelle Ausgestaltung“ unten.

[46] vgl. § 99 Abs. 1 S. 2 VwGO; vgl. auch § 30 VwVfG

[47] Vgl. §1 AGG

Anwendungsbereich ist auf die Bereiche Arbeit, Bildung, Soziales und Massengeschäfte begrenzt. Dadurch sind weder alle Lebensbereiche, in denen algorithmenbasierte Diskriminierung vorkommen kann, noch alle Formen von Ungleichbehandlungen (etwa unterschiedliche Bepreisung auf Grund von gruppenspezifischer Zahlungsbereitschaft).[48] Es könnte erwogen werden, den Anwendungsbereich auf alle Ungleichbehandlungen auszuweiten, die auf einer algorithmenbasierten Datenauswertung oder einem automatisierten Entscheidungsverfahren beruhen. Aus den Erfahrungen mit dem AGG könnte jedoch eine entsprechend breite Erweiterung des Anwendungsbereichs die Gefahr rechtsmissbräuchlicher Entschädigungsansprüche bergen.[49] Insofern gilt es, mit der Etablierung eines staatlich legitimierten Kontrollsystems zunächst einmal die Einhaltung und Durchsetzung der bestehenden Regelungen des AGG im Rahmen von ADM-Systemen sicherzustellen.

3.1 Automatisierte Entscheidung im Einzelfall

Eine zentrale rechtliche Grundlage für die Regulierung von ADM-Systemen – soweit personenbezogene Daten verarbeitet werden – stellt Artikel 22 DSGVO dar. Demnach hat der Betroffene das Recht, keiner ausschließlich auf einer automatisierten Verarbeitung beruhenden Entscheidung unterworfen zu werden, die ihm gegenüber rechtliche oder andere erhebliche Folgen hat – es sei denn, es gibt eine gesetzliche Erlaubnis oder sie ist für die Erfüllung eines Vertrags notwendig oder der Betroffene hat explizit eingewilligt.

Problematisch ist jedoch, dass der Anwendungsbereich des Artikel 22 DSGVO auf Entscheidungen begrenzt ist, die „*ausschließlich*“ auf einer automatisierten Datenverarbeitung beruhen. So ist beispielsweise unklar, inwieweit die Kreditvergabe unter Verwendung eines Scorewertes als eine solche „ausschließlich auf einer automatisierten Verarbeitung beruhenden Entscheidung“ zu betrachten ist, da die Entscheidung, ob ein Kredit gewährt wird, häufig zumindest formal von dem entsprechenden Bankmitarbeiter getroffen wird. Die Artikel-29-Datenschutzgruppe[50] betont, dass eine menschliche Mitwirkung nur gegeben sei, wenn es sich dabei nicht lediglich um eine symbolische Geste handele und die Person über die Autorität und Kompetenz verfüge, die Entscheidung zu ändern.[51] Diese Entscheidungskompetenz bei der Kreditvergabe hat ein Bankmitarbeiter nach Ansicht des vzbv in den allermeisten Fällen nicht. Somit zeigt dieses Beispiel, dass die Vorschriften zu automatisierten Entscheidungen und zur Profilbildung stark auslegungsfähig sind. Auch eine große Zahl anderer algorithmischer Entscheidungsverfahren dürfte von Artikel 22 DSGVO nicht erfasst sein.[52] Insofern besteht hier Regelungsbedarf, um Rechtssicherheit zu schaffen.

---

[48] Vgl. Martini, Mario (2018) (wie Anm. 14), S. 23

[49] Vgl. ebd., S. 24

[50] Die Artikel-29-Datenschutzgruppe war eine unabhängige Arbeitsgruppe, die sich mit dem Schutz personenbezogener Daten und der Privatsphäre beschäftigt hat. Seit Inkrafttreten der DSGVO wurde sie abgelöst vom Europäischen Datenschutzausschuss: EDPB - European Data protection Board: Artikel-29-Datenschutzgruppe (2019), URL: https://edpb.europa.eu/our-work-tools/article-29-working-party_de [Zugriff: 17.04.2019]

[51] WP29 Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01), 13.02.2018, Seite 21

[52] Martini, Mario (2018) (wie Anm. 14), S. 19-20

> Artikel 22 DSGVO sollte auf Entscheidungen ausgedehnt werden, die nicht nur auf einer ausschließlichen, sondern auch auf einer überwiegend automatisierten Verarbeitung von Daten beruhen.

Problematisch ist ferner, dass Artikel 22 DSGVO keine Qualitätsanforderungen an automatisierte Entscheidungen festschreibt.[53] Solche Qualitätsanforderungen sollten sicherstellen, dass Fehler und Risiken von ADM-Systemen reduziert werden und dass Verbraucher darauf vertrauen können, dass ADM-Systeme tatsächlich auf Basis valider Annahmen und Modelle arbeiten.

Die Idee, Qualitätsanforderungen für automatisierte Entscheidungen zu definieren, ist nicht neu. So hat der Europäische Gerichtshof im Rechtsgutachten 1/15 vom 26. Juli 2017 für das EU-Parlament hinsichtlich der automatisierten Analyse von Fluggastdaten betont, dass die Modelle und Kriterien, auf denen die automatisierte Verarbeitung der Daten beruht, *spezifisch*, *zuverlässig* sein müssen und *nicht diskriminieren* dürfen. Außerdem müssten die Datenbanken, mit denen die PNR-Daten abgeglichen werden, *zuverlässig* und *aktuell* sein. Um Diskriminierung zu verhindern, müsse ferner die *Zuverlässigkeit und Aktualität* dieser Modelle und Kriterien sowie der verwendeten Datenbanken, *unter Berücksichtigung statistischer Daten und der Ergebnisse der internationalen Forschung regelmäßig überprüft* werden.[54]

Auch die – eigentlich zivilrechtliche – Vorschrift des § 31 BDSG stellt Anforderungen für automatisierte Entscheidungen im Bereich des Kreditscorings. Um Willkür zu verhindern macht § 31 BDSG die Zulässigkeit der Nutzung eines Scorewerts im Geschäftsverkehr beispielsweise neben der Einhaltung des Datenschutzrechts davon abhängig, ob *die* zur Berechnung des Scorewerts genutzten *Daten unter Zugrundelegung eines wissenschaftlich anerkannten mathematisch-statistischen Verfahrens nachweisbar* für die Berechnung der Wahrscheinlichkeit des bestimmten Verhaltens *erheblich* sind.

> Um Sicherungsmechanismen einzuziehen und Fehler und Risiken von ADM-Systemen zu reduzieren, sollten materiell-rechtliche Qualitätsvorgaben an diese Systeme festgeschrieben werden. Insbesondere sollten die für eine überwiegend oder ausschließlich auf einer automatisierten Verarbeitung beruhenden Entscheidung genutzten Daten unter Zugrundelegung eines anerkannten mathematisch-statistischen Verfahrens verarbeitet werden und für die Entscheidungsfindung erheblich sein. Die Prognosetauglichkeit, Validität und Reliabilität des mathematisch-statistischen Verfahrens sollte wissenschaftlich nachgewiesen werden können.

[53] Zwar werden Anforderungen in Erwägungsgrund 71 DSGVO genannt, allerdings sind diese nur abstrakt, oberflächlich und unzureichend angelegt.

[54] Rechtsgutachten 1/15 des EuGH vom 26. Juli 2017 für das EU-Parlament Rn 172, 174

3.2 Benennung einer verantwortlichen Person durch Betreiber von ADM-Systemen

Betreiber von relevanten ADM-Systemen sollten eine verantwortliche Person benennen, die eine äquivalente beziehungsweise komplementäre Rolle zu Aktuaren[55] oder Datenschutzbeauftragten spielt. Diese verantwortliche Person könnte - in Ergänzung zu dem Datenschutzbeauftragten - etwa sicherstellen, dass ein ADM-System im Einklang mit geltendem Recht (beispielsweise AGG, Gesetz gegen den unlauteren Wettbewerb (UWG)) gestaltet und betrieben wird.

> Der Gesetzgeber sollte Unternehmen verpflichten, eine verantwortliche Person zu benennen, deren Verantwortung es ist, sicherzustellen, dass ADM-Systeme im Einklang mit dem geltenden Recht gestaltet und betrieben werden.

V. INSTITUTIONELLE AUSGESTALTUNG

Obwohl viele der materiell-rechtlichen Regelungen, auf die sich eine Algorithmenkontrolle stützen würde, auf EU-Ebene angesiedelt sind[56] und es für viele Sektoren auch auf EU-Ebene zuständige (Regulierungs-) Institutionen gibt, obliegt die Umsetzung und Durchsetzung gesetzlicher Regelungen, die auf europäischem Recht basieren in der Regel nationalen (auf Landes- und Bundesebene angesiedelten) Aufsichtsbehörden (Datenschutzaufsichtsbehörden, Bundesnetzagentur, BAFin etc.).

1. INSTITUTIONEN ZUR AUFSICHT / ÜBERPRÜFUNG

Die bundesdeutsche Aufsichtsstruktur ist vielfältig und für bestimmte Sektoren führen spezifische Aufsichtsbehörden die Aufsicht auf Landes- und Bundesebene durch. So gibt es etwa das Bundeskartellamt, die BaFin (v.a. für den Hochfrequenzhandel[57]), die Antidiskriminierungsstelle des Bundes und datenschutzrechtliche Aufsichtsbehörden.[58] In der Debatte um die institutionelle Ausgestaltung des Kontrollsystems zeichnet sich ab, dass die Schaffung einer einheitlichen Aufsichtsbehörde, die sektorübergreifend ADM-Systeme beaufsichtigt, als ungeeignet angesehen wird.[59] Für bestimmte Anwendungen der digitalen Wirtschaft gibt es jedoch keine einheitliche und einzige Aufsichtsbehörde. So gibt es zum Beispiel nicht eine Behörde, die die Aufsicht über Plattformen innehat, sondern unterschiedliche Behörden sind entsprechend der betroffenen Aspekte (zum Beispiel Datenschutzrecht, Kartellrecht etc.) zuständig. Insofern könnte es erforderlich sein, die Frage der Zuständigkeit im Hinblick auf eine Überprüfung des

---

[55] Vgl. § 141 Abs. 5 VAG sowie: Was ist ein Aktuar? in: Deutsche Aktuarvereinigung - DAV, URL: https://aktuar.de/aktuar-werden/was-ist-ein-aktuar/Seiten/default.aspx [Zugriff: 17.04.2019].

[56] Auch für noch zu schaffende künftige Regelungen kann davon ausgegangen werden, dass diese sinnvollerweise auf EU-Ebene verortet werden.

[57] Bundesanstalt für Finanzdienstleistungsaufsicht - BaFin (wie Anm. 38)

[58] So auch Martini, Mario (2018) (wie Anm. 14), S. 30f. f.w.N.

[59] Ebd.; S. 30ff.

ADM-Systems zu erörtern und gegebenenfalls neu zu ordnen (beispielsweise im Hinblick auf ADM-Systeme in bestimmten Anwendungsbereichen, etwa digitale Sprachassistenten).

> Der Gesetzgeber sollte prüfen, ob die Zuständigkeiten und Befugnisse bestehender Aufsichtsbehörden im Hinblick auf eine Überprüfung von ADM-Systemen angemessen sind. Sollten Schutzlücken identifiziert werden, etwa, weil relevante Anwendungsbereiche von ADM-Systemen keiner direkten Aufsicht unterworfen sind, sollte die Aufsichtslandschaft neu geordnet werden, um diese Schutzlücken zu schließen.

Bestehende sektorspezifisch zuständige Aufsichtsbehörden/Institutionen sollten die Kontrolle in ihrem Anwendungsbereich weiterhin durchführen. Ihre inhaltliche, bereichsspezifische Expertise ist Voraussetzung für eine Analyse, die umfassend die Risikopotenziale eines ADM-Systems für Verbraucher und Gesellschaft in den Blick nehmen kann.

Als Teil des Kontrollsystems müssen die Aufsichtsbehörden über die Möglichkeit verfügen, selber zu prüfen, inwieweit ADM-Systeme mit den rechtlichen Vorgaben in Einklang stehen. Hierfür müssen sie über die erforderlichen Befugnisse verfügen. Blaupausen für derartige behördliche Befugnisse zur inhaltlichen Kontrolle bestehen in verschiedenen Rechtsgebieten. Beispielsweise regelt Art. 58 DSGVO die Untersuchungsbefugnisse der Datenschutzaufsicht, § 32e Abs. 5 und Abs. 6 GWB regelt die Sektoruntersuchungen durch das Bundeskartellamt. Die Kontrolle des Hochfrequenzhandels durch Finanzaufsichtsbehörden basiert auf § 6 Abs. WpHG. § 3 Abs. 4 Nr. 5 BörsG n.F. i.V.m. § 7 Abs. 3 S. 1 BörsG).

> Der Gesetzgeber sollte für die einschlägigen Aufsichtsbehörden erforderliche Befugnisse (zum Beispiel Auskunfts-, Einsichts- und Zugangsrechte) beschließen, damit diese ADM-Systeme bewerten und überprüfen können.

2. EINRICHTUNG EINER EINHEIT ZUR UNTERSTÜTZUNG SEKTORSPEZIFISCHER AUFSICHTSBEHÖRDEN

Die technisch-methodische Expertise, um die Entscheidungslogiken von ADM-Systemen nachzuvollziehen und diese auf ihre Rechtmäßigkeit und gesellschaftlichen Auswirkungen hin zu kontrollieren, kann bei bestehenden Aufsichtsbehörden nicht vorausgesetzt werden. Deshalb ist die Idee „eine technisch versierte *Unterstützungseinheit* aus der Taufe zu heben, welche die Aufsichtsbehörden bei der Vorbereitung und Durchführung ihrer Befugnisse unterstützt“[60] eine angemessene Lösung. Damit würde die hoheitliche Durchsetzung gesetzlicher Rahmenbedingungen weiterhin sektoral bei den bisher zuständigen Behörden liegen und die „Unterstützungseinheit“ könnte die unterschiedlich zuständigen Behörden auf der vertikalen Ebene bei der Kontrolle von ADM-Systemen unterstützen.

[60] Ebd. S. 33. Vgl. auch Gesellschaft für Informatik (wie Anm. 23), S. 9.

> Die Bundesregierung sollte eine Unterstützungseinheit etablieren, die die zuständigen sektorspezifischen Aufsichtsbehörden bei der Kontrolle von ADM-Systemen mit technisch-methodischer Expertise unterstützen kann.

3. ÜBERTRAGUNG DER AUFSICHTSTÄTIGKEIT AUF BELIEHENE

Eine Prüfung von ADM-Systemen etwa zur Zulassung oder ex-post, die die Geheimhaltungsinteressen der Betreiber oder dritter Betroffener wahrt, müsste nicht per se durch die zuständigen Aufsichtsbehörden selber erfolgen, sondern könnte durch den Gesetzgeber auch auf Beliehene übertragen werden[61] (etwa der TÜV). Diese könnten beispielsweise auch Zertifizierungsverfahren durchführen (siehe nächster Abschnitt).

4. REPRÄSENTATION DER VERBRAUCHERINTERESSEN

Um sicherzustellen, dass bei der Überprüfung von ADM-Systemen die Interessen der betroffenen Verbrauchergruppen angemessen berücksichtigt werden, sollten Beiräte bei den sektorspezifisch zuständigen Prüfungskommissionen eingerichtet werden. In diesen Beiräten sollten Verbrauchervertreter sowie andere zivilgesellschaftliche Organisationen vertreten sein, um sicherzustellen, dass den Interessen der Verbraucher bei der Prüfung angemessen Rechnung getragen wird.

VI. GESTALTUNG VON ADM-SYSTEMEN

1. DOKUMENTATIONSPFLICHTEN DES BETREIBERS

Das Kontrollsystem muss auch komplexe ADM-Systeme nachvollziehen können. Dies ist nur gewährleistet, wenn die Prozessschritte eines ADM-Systems entsprechend dokumentiert werden. Die Dokumentationspflicht muss dabei den gesamten Prozess der Entscheidungsfindung bzw. Entscheidungsvorbereitung eines ADM-Systems umfassen (etwa das Trainingsmodell und Trainingsdaten, Entscheidungen über die Kriterien, nach denen ein Modell optimiert wurde[62]). Dies ist angemessen, da Entscheidungen beziehungsweise Fehler, die zur Rechtswidrigkeit des gesamten ADM-Systems führen können, während aller Phasen der Entwicklung eines ADM-Systems vorkommen können[63] (etwa wenn verzerrte Trainingsdaten dazu führen, dass das finale ADM-System geschützte Gruppen im Sinne des AGG systematisch diskriminiert).

> Verbindliche Regeln, Normen und Standards für die technische Gestaltung sowie die Dokumentation und Beschreibung von ADM-Systemen sind erforderlich, um diese einer Kontrolle zugänglich zu machen (Nachvollziehbarkeit-by-Design). Programmabläufe und Entscheidungen im Designprozess eines ADM-Systems sollten durch Unternehmen so dokumentiert werden müssen, dass gewährleistet ist, dass ADM-Systeme nachvollziehbar sind.

---

[61] Martini, Mario (2018) (wie Anm. 14), S. 32ff.

[62] Beispielsweise bei der automatischen Vergabe von Sitzplätzen an Flugzeugpassagiere gibt es Hinweise darauf, dass die Platzvergabe nicht immer nach reinen Sicherheitsaspekten optimiert wird, wenn gemeinsam buchenden Reisenden getrennte Plätze zugewiesen werden. Vgl. Zweig, Katharina; Krafft, Tobias (wie Anm. 10)

[63] Zweig, Katharina: 2. Arbeitspapier: Überprüfbarkeit von Algorithmen. Arbeitspapier AlgorithmWatch, URL: https://algorithmwatch.org/zweites-arbeitspapier-ueberpruefbarkeit-algorithmen/

2. NORMEN UND STANDARDS ZUR PROZESSGESTALTUNG

Normen und Standards zur Prozessgestaltung und für das Qualitätsmanagement von ADM-Systemen sollten eingeführt werden, um sicherzustellen, dass diese von vornherein rechtliche Vorgaben sowie gegebenenfalls weitere ethische Anforderungen erfüllen[64]. Die Erarbeitung über nationale, europäische und internationale Normungsinstitute sollte vorangetrieben werden.[65]

> Es müssen (Prozess-)Standards und Normen eingeführt werden, um sicherzustellen, dass ADM-Systeme von vornherein rechtliche Anforderungen (Legality-by-Design) sowie gegebenenfalls weitere ethische Standards erfüllen.

3. STANDARDISIERTE SCHNITTSTELLEN

Für eine effektive Kontrolle zur Durchführung technisch-statistischer Tests im Rahmen der Überprüfung eines ADM-Systems ist es erforderlich, dass die Betreiber eine technische Schnittstelle vorhalten, über die die zuständige Aufsichtsbehörde jederzeit Zugang zu dem System erhalten kann. Hierüber können mögliche technische oder methodische Fehler identifiziert werden. Mit diesem Zugang könnten beispielsweise sogenannte Input-Output-Tests durchgeführt werden, mit denen geprüft werden kann, ob ein ADM-System systematisch Gruppen, die durch das AGG geschützt sind, benachteiligt.[66]. Im Falle Machine-Learning-basierter ADM-Systeme sollten die Schnittstellen einen Zugang zu den Trainingsdaten und dem verwendeten Lernverfahren bereitstellen, mit dem ein Algorithmus die Entscheidungsregeln/Systeme lernt (also das Trainingsmodell und den Programmcode umfassen). Zur Überprüfung der Entscheidungen an sich (etwa hinsichtlich möglicher Diskriminierung) bedarf es eines Zugangs zu dem eigentlichen Entscheidungsprozess. Also ein Zugang zu fertig trainiertem Modell/Code sowie den genutzten Input- und Outputdaten.

> Betreiber von relevanten ADM-Systemen sollten technische Schnittstellen vorhalten müssen, so dass die zuständigen Aufsichtsbehörden über diese auf das System zugreifen können, um es auf ihre Rechtmäßigkeit sowie technische und methodische Fehler hin überprüfen zu können.

---

[64] Beispielhaft erwäht sei die DIN EN ISO 13485, die Anforderungen an ein Qualitätsmanagementsystem für Unternehmen festlegt, die Medizinprodukte und zugehörige Dienstleistungen bereitstellen. Vgl. DIN EN ISO 13485 Medizinprodukte - Qualitätsmanagementsysteme - Anforderungen für regulatorische Zwecke. in: DIN - Deutsches Institut für Normung (2016), URL: https://www.din.de/de/mitwirken/normenausschuesse/named/normen/wdc-beuth:din21:244078306 [Zugriff: 30.04.2019]. Eine Norm, die direkt ethische Standards festlegt ist die DIN EN ISO 14155, die im Kontext der klinischen Prüfung medizinischer Produkte Probanden schützen soll; vgl. DIN EN ISO 14155 Klinische Prüfung von Medizinprodukten an Menschen - Gute klinische Praxis (ISO/DIS 14155:2018). in: DIN - Deutsches Institut für Normung (2018), URL: https://www.din.de/de/mitwirken/normenausschuesse/nafuo/entwuerfe/wdc-beuth:din21:289646651 [Zugriff: 30.04.2019]..

[65] Zu Scraping Audits die über technische Schnittstellen durchgeführt werden können vgl. Gesellschaft für Informatik (wie Anm. 23), 58ff.

[66] Vgl. Zweig, Katharina; Krafft, Tobias (wie Anm. 10) S. 40ff.

VII. HAFTUNG

Wenn Verbraucher durch ADM-Systeme geschädigt werden, muss die Kompensation des Schadens durch eine angemessene Haftung der Verantwortlichen sichergestellt werden. Denkbare Schadensereignisse durch ADM-Systeme sind zum Beispiel, dass

- dem Verbraucher aufgrund einer fehlerhaften Einschätzung seiner Zahlungskraft ein Kredit verweigert wird und ihm dadurch ein Schaden entsteht, zum Beispiel er einen teureren Kredit als den zu Unrecht verweigerten in Anspruch nehmen muss oder eine günstigere von mehreren wirtschaftlichen Möglichkeiten wegen der Verweigerung nicht wahrnehmen kann,
- in Smart Homes Schäden durch Fehler in Bewässerungssystemen, Alarmanlagen oder in der Fußbodenheizung entstehen oder
- der Verbraucher wegen eines irreführenden Rankings auf einer Vergleichsplattform oder wegen einer fehlerhaften Empfehlung eines digitalen Assistenten ein Produkt oder eine Dienstleistung zu einem teureren Preis als dem günstigsten auf der Plattform erwirbt, obwohl der Eindruck erweckt wurde, der günstigste Preis sei angezeigt worden.

Ein grundlegendes Problem ergibt sich im Recht der Beweisführung. Der im Rahmen einer zivilrechtlichen Haftung vom Geschädigten zu führende Beweis, dass der Anwender eines ADM-Systems auch verantwortlich für den Schaden ist, ist in den meisten Fällen vor Gericht kaum möglich. Denn die genauen technischen Vorgänge, die Arbeitsschritte des ADM-Systems sind für den Geschädigten in aller Regel nicht zu erkennen.[67] Ein besonderes Augenmerk muss bei der Haftung also der Beweislast gelten. Angemessen erscheint hier eine von einem Fehler unabhängige Haftung für ADM-Systeme im Sinne einer Gefährdungshaftung bei bestimmungsgemäßer Verwendung durch den Verbraucher. Für die Haftung des Anbieters sollte es dann ausreichen, wenn ein ADM-System bei bestimmungsgemäßer Verwendung einen Schaden verursacht, der bei der Anwendung des jeweiligen ADM-Systems typischerweise zu erwarten ist. Diese Gefährdungshaftung wurde vor allem entwickelt für die Halter von Kraftfahrzeugen und Tieren, denen in jedem Fall eine gewisse „Betriebsgefahr“ inhärent ist.[68] ADM-Systeme können, insbesondere wenn sie selbstlernend sind, ein Eigenleben entwickeln, das an das autonome, vom Menschen nur beschränkt kontrollierte Handeln von Tieren erinnert. Beispielsweise sind die Gefahren des Einsatzes eines Pflegeroboters (Fehler, die zu Verletzungen oder zum Tod führen können) oder der Gesichtserkennung (Fehler bei der Erkennung und dadurch Verletzung von Persönlichkeitsrechten, aber auch Datenlecks und vollständige Überwachung) evident und lassen sich auch durch Vorsichtsmaßnahmen nicht gänzlich eliminieren. Angesichts der umfassenden Überwachungs- und Kontrollmöglichkeiten durch ADM-Systeme, die sich in ihrem vollen Ausmaß erst andeuten, kann es dabei nicht nur auf Gefahren für Leib und Leben

[67] So auch Martini (2018) – Gutachten für vzbv; S. 35f. m.w.N.

[68] Das bedeutet, dass bei dem Betrieb eines Kfz oder dem Spazierengehen mit einem Hund nach menschlichem Ermessen stets eine gewisse Gefahr von dem Kfz/Hund ausgeht, die nicht vollständig eingedämmt werden kann. Wer ein solches „Objekt“ also hält, soll grundsätzlich dafür verantwortlich sein, wenn der jeweils typische Schaden (Unfall, Biss) eintritt.

ankommen. Die Selbstbestimmung des Einzelnen und das Recht, nicht in totaler Überwachung zu leben, ist ein Rechtsgut von ausreichendem Gewicht für die Anwendung einer Gefährdungshaftung. Sie ist daher hier das einzig angemessene Mittel.

Eine solche Beweislastverteilung würde deswegen den jeweiligen Risikosphären entsprechen. Der Nutzer beziehungsweise Geschädigte müsste den Kausalzusammenhang zwischen dem bestimmungsgemäßen Gebrauch und dem Schaden darlegen. Der Anbieter würde dann grundsätzlich haften, ohne dass es auf sein konkretes Verschulden ankommt.

> Betreiber von relevanten ADM-Systeme sollten für Schäden im Sinne einer Gefährdungshaftung bei bestimmungsgemäßer Verwendung durch den Verbraucher haftbar sein.

In einigen Fällen, insbesondere in Bezug auf Smart Homes, kann eine angemessene Haftung auch durch eine Ausweitung des Produkthaftungsrechts erreicht werden. Insbesondere dann, wenn Produkte durch ADM-Systeme gesteuert werden, müssen ihre Hersteller für von ihnen ausgehende Gefahren jedenfalls verantwortlich bleiben.

> Als Ansatzpunkt hierfür bietet sich eine Evaluierung und Aktualisierung der europäischen Produkthaftungsrichtlinie an. Schäden, die durch ADM-Systeme verursacht werden, müssen dabei erfasst werden. Zurechnungs- und Beweisprobleme des Produkthaftungsrechts müssen neu justiert werden, damit Geschädigte ihre berechtigten Ansprüche tatsächlich durchsetzen können und dass Haftungsrecht seine Steuerungs- und Vorsorgefunktion für die Produktsicherheit erfüllen kann.

Darüber hinaus sind auch durch ADM-Systeme verursachte Schäden denkbar, die nicht einfach in Geld gemessen werden können. Insbesondere in Fällen, in denen Menschen ein „Score“ zugewiesen wird, um sie für den Anwender besser quantifizierbar zu machen (Kreditwürdigkeit, Vertrauenswürdigkeit etc.), kann der entstehende Schaden bei einer Fehleinschätzung (oder sogar einer richtigen Einschätzung) immateriell sein.

> Die Anwender von ADM-Systemen müssen für Fehlfunktionen haftbar gemacht werden können. Hier ist die Einführung eines immateriellen Schadensersatzanspruchs erforderlich, so wie er bereits in Art. 82 der DSGVO angedacht wird.

Problematisch für eine effektive Rechtsdurchsetzung sind aber insbesondere Szenarien bei dem vor allem kleine Schäden bei einer Vielzahl von Menschen entstehen, die nicht hoch genug sind, dass sich für den Einzelnen ein Einklagen lohnen würde (Streuschäden).

> Hier ist ein effektives Ausgleichssystem einzurichten, beispielsweise durch kollektive Klageinstrumente oder durch eine verbesserte und vereinfachte Abschöpfung von Unrechtsgewinnen.

VIII. ANWENDUNGSBEISPIELE

Zur Veranschaulichung, wie die Umsetzung des Kontrollsystems in der Praxis aussehen könnte, wird im Folgenden exemplarisch und vereinfacht anhand von Beispielen skizziert, wie eine Risikoprüfung und die darauffolgende Umsetzung von Maßnahmen aussehen könnten.

Geringes oder kein Risikopotential: Künstliche Gegenspieler in Games

Im Spielebereich finden hoch entwickelte ADM-Systeme zunehmend Anwendung. Dabei übernimmt das – in der öffentlichen Debatte häufig als Künstliche Intelligenz (KI) verortete – System die Rolle des Gegenspielers des menschlichen Spielers. Bekannte Beispiele sind Google‘s AlphaGo, dem es 2016 überraschend gelang, einen hochrangigen menschlichen professionellen Go-Spieler zu schlagen[69]. Ein weiteres Beispiel ist ein KI-Projekt von OpenAI, das ein Team von ehemaligen professionellen menschlichen Spielern in Dota 2 schlug[70] – einem Echtzeit-Multiplayer-Strategiespiel.

Der Einsatz künstlicher Spielpartner in Spielen lässt, zumindest in den bisherigen Erscheinungsformen und Funktionen, ein geringes Schadenspotenzial für Individuen und/oder die Gesellschaft erwarten. Aus diesem Grunde scheint es nicht nötig, Maßnahmen zum Schutz der Verbraucher zu ergreifen, wie etwa Kennzeichnungs- und Informationspflichten oder gar eine Einsichtnahme und Kontrolle des zugrundeliegenden ADM-Systems.

Geringes Risikopotenzial: Spamfilter

Sogenannte Spamfilter sind ADM-Systeme, deren Aufgabe es ist, massenhaft versandte, unerwünschte Nachrichten (häufig werblichen Inhalts) auszusortieren.

Die Anwendung von Spamfiltern lässt für die individuellen Nutzer überwiegend keine negativen Auswirkungen erkennen. Zudem wird es im Interesse des Betreibers des Filters liegen, dass dieser fehlerlos funktioniert, da sonst die Gefahr des Vertrauensverlustes und ein Wechsel der Nutzer zu einem anderen Anbieter droht. Somit sind weitergehende Maßnahmen wie etwa Transparenz zur Logik ihrer Funktionsweise des Filters gegenüber den Nutzern oder gar eine Einsichtnahme oder Kontrolle durch eine Aufsichtsinstitution nicht erforderlich.

Sollte sich allerdings der begründete Verdacht ergeben, dass der Spamfilter nicht sachgerecht funktioniert, etwa indem er (gegebenenfalls sogar systematisch) dringliche, wichtige Emails aussortiert und daraus ein Schaden entsteht könnte dies etwa eine Einsichtnahme und Überprüfung rechtfertigen.

Mittleres Risikopotential: Digitale Sprachassistenten

Digitale Sprachassistenten (etwa Alexa oder Siri) nehmen zunehmend Einzug in den Verbraucheralltag. Über Smartphone oder Smart Speaker können Verbraucher mittels

---

[69] DeepMind Technologies Limited: AlphaGo (2019), URL: https://deepmind.com/research/alphago/ [Zugriff: 18.03.2019]

[70] Steinlechner, Peter: KI besiegt mit eigener Heldenwahl ehemalige Profis. in: Golem.de (2018), URL: https://www.golem.de/news/dota-2-ki-besiegt-mit-eigener-heldenwahl-ehemalige-profis-1808-135873.html [Zugriff: 18.03.2019]

Spracherkennung mit den Assistenten Unterhaltungen führen, Websuchen vornehmen, ihr Smart Home steuern, einkaufen oder sich praktische Ratschläge - wie Routenbeschreibungen zur nächsten Tankstelle – holen.

Sie sind ein Beispiel für hochentwickelte ADM-Anwendungen, die einen großen Einfluss auf den Verbraucheralltag und das Verbraucherverhalten haben können, etwa indem sie Konsummuster beeinflussen und Konsumströme ganzer Gruppen beeinflussen können. Gerade bei Produktempfehlungen sollte sichergestellt sein, dass sie tatsächlich im Sinne des Verbrauchers sind und rechtskonform agieren. So könnte sich der gesamtwirtschaftliche Schaden leicht zu großen Summen aufsummieren, wenn ein Digitaler Sprachassistent Verbraucher systematisch auf Anbieter mit leicht erhöhten Preisen verweist. Dies könnte etwa aufgrund eines Fehlers oder absichtlich geschehen: Etwa um eigene Angebote gegenüber denen von Wettbewerbern zu bevorzugen[71], weil Anbieter den Betreiber des Sprachassistenten dafür bezahlen, bestimmten Angebote in den Markt zu pushen oder – etwa um im Rahmen einer Preisdifferenzierungsstrategie[72] - die Einnahmen durch Provision für die Vermittlung zu erhöhen.[73] Ebenso können sie auf wichtige Verbraucherentscheidungen einen relevanten Einfluss haben - etwa, wenn sie von Verbrauchern genutzt werden, um sich Altersvorsorgeprodukte oder Versicherungen empfehlen zu lassen. Lock-in Effekte erschweren Verbrauchern den Wechsel von digitalen Sprachassistenten.[74] Deshalb ist im Falle solcher Unregelmäßigkeiten nicht ohne weiteres mit einer Änderung des Anbieters zu rechnen. Damit ist zweifelhaft, ob der zu erwartende Marktdruck auf die Anbieter der Assistenten eine ausreichende (potenzielle) Sanktion darstellt, um rechtskonformes, verbraucherfreundliches Verhalten zu incentivieren. Digitale Assistenten sind aber nicht nur in der Form von „persönlichen Assistenten“ denkbar, die auf den Geräten der Nutzer installiert sind. So ist denkbar, dass sie von Unternehmen eingesetzt werden um die Interaktion mit Verbrauchern zu gestalten, etwa in der Beratung oder in Hotlines.

Digitale Assistenten sind ADM-Systeme, die eine enorme Relevanz für Verbraucher und Verbraucherentscheidungen haben können – nicht zuletzt in finanzieller Hinsicht. Nicht zu unterschätzen ist auch das Risiko einer Beeinflussung der Verbraucherpräferenzen. Durch die Gewöhnung an die Konsumempfehlungen des digitalen Assistenten könnten die Nutzerpräferenzen langsam in eine bestimmte Richtung gelenkt werden, ohne dass dies dem Verbraucher bewusst ist. Denkbar ist auch, dass Konsummuster zementiert werden, weil der Empfehlungsalgorithmus sich an früheren Konsumentscheidungen orientiert. Ebenso unklar sind derzeit noch die gesamtwirtschaftlichen Auswirkungen: Wenn Konsumentscheidungen an digitale Assistenten delegiert werden, entfällt der Verbraucher in seiner Rolle als Schiedsrichter im Wettbewerb. Es besteht

---

[71] Beispielhaft: European Commission: Antitrust: Commission Fines Google €2.42 Billion for Abusing Dominance as Search Engine by Giving Illegal Advantage to Own Comparison Shopping Service. in: European Commission (2017), URL: http://europa.eu/rapid/press-release_IP-17-1784_en.htm [Zugriff: 30.04.2019].

[72] Kenning, Peter; Wagner Gert. https://www.facebook.com/faz: Wo beginnt Diskriminierung? Die knifflige Sache mit dem Feilschen. in: Frankfurter Allgemeine Zeitung, URL: https://www.faz.net/aktuell/wirtschaft/diginomics/die-knifflige-sache-mit-dem-feilschen-16105620.html?printPagedArticle=true#pageIndex_0 [Zugriff: 30.04.2019]

[73] Vgl. Bundeskartellamt: Bundeskartellamt sieht verbraucherrechtlichen Handlungsbedarf bei Vergleichsportalen (2018), URL: https://www.bundeskartellamt.de/SharedDocs/Meldung/DE/Pressemitteilungen/2018/12_12_2018_Vergleichsportale.html [Zugriff: 30.04.2019]

[74] Erhöhte Wechselkosten entstehen zum Beispiel durch den Aufwand, den ein Verbraucher in das Training, beziehungsweise Personalisierung des Assistenten investiert hat und der verloren ginge. Zudem sind manche digitalen Assistenten an das Betriebssystem der Hardware des Nutzers geknüpft.

das Risiko, dass dies die Funktionsfähigkeit der Märkte, die Zuteilung knapper Ressourcen optimal zu koordinieren, beeinträchtigt. Schließlich setzt eine Prognose der Konsumpräferenzen auch eine weitreichende Bildung von Persönlichkeitsprofilen voraus – verbunden mit der Frage, wofür diese Persönlichkeitsprofile sonst noch genutzt werden.

Digitale Assistenten stehen gleichzeitig exemplarisch für ADM-Anwendungen, die nicht der direkten Zuständigkeit einer Aufsichtsbehörde unterworfen sind. Im Falle von Verdachtsmomenten wären für diese Anwendungen verschiedene Aufsichtsbehörden zuständig. Etwa das Bundeskartellamt im Falle von Verstößen gegen das Kartellrecht oder im Falle von Verstößen gegen das AGG die Antidiskriminierungsstelle des Bundes, die allerdings nicht mit Einsichts- oder Durchsetzungsbefugnissen ausgestattet ist.

Das Einhegen der potenziellen Risiken von digitalen Sprachassistenten kann auf Unternehmensseite eine Reihe von Maßnahmen umfassen: Eine Kennzeichnungspflicht (damit Verbraucher wissen, ob sie mit einer KI interagieren - etwa via Telefon -), Auskunfts- und Informationspflichten sowie die Erstellung und Veröffentlichung der Folgenabschätzung. Gleichzeitig sollten Behörden jederzeit die Möglichkeit haben, das System zu analysieren, etwa hinsichtlich einer unrechtmäßigen Preisdiskriminierung. Dafür müsste der Betreiber Dokumentationspflichten nachkommen und eine Schnittstelle zur Verfügung stellen, mit der die Aufsichtsbehörde Tests durchführen könnte, etwa um zu prüfen, ob ein systematischer Zusammenhang zwischen bestimmten Kundenmerkmalen (Alter Geschlecht etc.) und den Preisen besteht oder wie sich die Personalisierung des Angebots auf unterschiedliche Verbrauchergruppen auswirkt.

Eine ex-ante Kontrolle beziehungsweise Zulassung des Systems scheint – auch aufgrund der Vielfalt der Anwendungen - weder notwendig noch praktikabel, da auf Verdachtsmomente hin eine Einsichtnahme und Überprüfung durch eine Aufsichtsbehörde erfolgen kann. Dies schließt nicht aus, dass digitale Assistenten in bestimmten sensiblen Anwendungsbereichen einer Zulassung unterworfen werden sollten.[75]

Hohes Risikopotential: Roboter in der Pflege

Die Einsatzmöglichkeiten von ADM-Systemen in der Pflege sind bereits heute vielfältig. Die Überlegungen über den Einsatz von Roboter-Systemen in der Pflege werden neben potentiellen Kostenersparnissen vor allem vor dem Hintergrund der Optimierung von Versorgungsabläufen, der Entlastung von Pflegekräften und einer verbesserten Unterstützung des Pflegebedürftigen bei der Ermöglichung eines selbstbestimmten Alltags geführt. Mögliche Anwendungsbereiche von Pflegerobotern umfassen beispiels-

[75] Etwa digitale Assistenten zur Gesundheitsberatung, die Verbrauchern eine quasi-Diagnose erstellen. Im Falle der KI-„Diagnose"-App Ada Health interagiert der Nutzer mit der App per Sprachassistent, schildert Symptome. Die Anwendung gibt an zu welcher Wahrscheinlichkeit welche Krankheit vorliegt, i.d.R. verbunden mit dem Hinweis, dass der Nutzer zum Arzt gehen und das Ergebnis überprüfen lassen soll. Auch wenn rechtlich keine „Diagnose" erstellt wird, besteht das Risiko, dass Nutzer die Ergebnisse ohne eine ärztliche Meinung einzuholen, als Diagnose wahrnehmen. Vgl. etwa Ada Health GmbH (wie Anm. 9)

weise das Überwachen des Gesundheitszustandes der Pflegebedürftigen, die Darreichung und Überwachung der Einnahme von Medikamenten[76] oder Nahrung sowie Unterstützung bei der Körperhygiene, sowie persönliche Roboterassistenten, die für Hol- und Bringaufgaben als auch zur sozialen Interkation eingesetzt werden können.

Um optimal zu funktionieren, sind auch Pflegeroboter in vielen Anwendungsbereichen auf eine große Datenbasis angewiesen, die fortlaufend erweitert und systemintern ausgewertet wird, um die Leistungsfähigkeit zu verbessern und sich neuen Umständen anzupassen[77]. Als Beispiele hierfür sei der humanoide Serviceroboter Pepper genannt, der mit zahlreichen Sensoren ausgestattet und in der Lage ist, ältere Menschen an das Trinken zu erinnern und zum Bewegen zu motivieren. Ähnlich verhält es sich mit einem anderen Produkt: Care-O-bot, der von der Fraunhofer-Gesellschaft entwickelt wurde, hat den Zweck, Bewohner in stationären Pflegeeinrichtungen mit Getränken zu versorgen. Dabei merkt sich der Roboter, wie viele Getränke jeder zu sich genommen hat und serviert zum Beispiel Dehydrierten häufiger Getränke[78].

Bei der Bestimmung des Risikopotentials beim Einsatz von Pflegerobotern kommen verschiedene Dimensionen zum Tragen: Mögliche Gesundheitsgefährdungen der Pflegebedürftigen, etwa durch Verwechslung von Arzneimitteln, Nahrungsmitteln oder physische Verletzungen der Pflegebedürftigen durch Bewegungen der Roboter selbst. Gründe dafür könnten in der Sensorik der Pflegeroboter selber oder der Interpretation der Umgebungsdaten liegen. Der Einsatz von Pflegerobotern bedingt die Sammlung (unter anderem durch Mikrophone, Kameras), Verarbeitung und gegebenenfalls Weiterleitung von hochsensiblen persönlichen Daten der Betroffenen. Ein Missbrauch oder Verlust der Daten, etwa durch mangelnde Sicherheitsvorkehrungen, kann für Betroffene einen erheblichen Schaden bedeuten. Weitere ethische Probleme und immaterielle Schäden, die durch den Einsatz von Robotern verursacht werden können, sind vielschichtig. Sie reichen von Fragen der Selbstbestimmung und Autonomie der Pflegebedürftigen (etwa in Bezug auf die Medikamenteneinnahme) bis zu Fragen der Akzeptanz und des psychischen und physischen Wohlbefindens bei der Betreuung durch einen Roboter im Gegensatz zu der Betreuung durch Menschen.
Die soeben erwähnten Pflegeroboter stehen beispielhaft für ADM-Systeme, denen ein hohes Risikopotenzial - auch für Leben, Gesundheit und Eigentum von Menschen - zugeschrieben werden kann. Dies rechtfertigt das Ergreifen von Maßnahmen mit einer hohen Eingriffstiefe. Je nach Zweckrichtung des Pflegeroboters bestehen für die Marktzulassung unterschiedliche gesetzliche Vorschriften. Schaut man sich den gegenwärtigen Markt an, kommen bei der Mehrheit der Pflegeroboter mangels eines therapeutischen, diagnostischen oder medizinischen Einsatzzweckes nur die Regelungen aus dem Produktsicherheitsgesetz und nicht etwa darüber hinaus die strengeren Vorschriften des Medizinproduktegesetzes in Betracht[79]. Beim Medizinproduktegesetz ist je nach Risikoklasse ein Prüf- und Bewertungsverfahren durch eine unabhängige, staatlich autorisierte Instanz, wie zum Beispiel den Technischen Überwachungsverein (kurz: TÜV)

[76] Anderson, Michael; Anderson, Susan Leigh: Ethische Roboter für die Altenpflege, in: Otto, Philipp/Gräf, Eike (Hrsg.): 3TH1CS. Die Ethik der digitalen Zeit, S. 90–99, Berlin, iRights Media

[77] Kehl, Christoph: Robotik und assistive Neurotechnologien in der Pflege — gesellschaftliche Herausforderungen (2018), in: TAB Arbeitsbericht, Nr. 177, URL: http://www.tab-beim-bundestag.de/de/untersuchungen/u106002.html [Zugriff: 30.04.2019], S. 136 ff.

[78] Teamnews: Künstliche Intelligenz - Eine bessere Welt durch Pflegeroboter & Co (Ausgabe 2017)

[79] Kehl, Christoph (wie Anm. 77), S. 20.

vorgesehen, und erst nach dieser Bewertung hat der Hersteller die Erlaubnis die für alle Medizinprodukte erforderliche CE-Kennzeichnung anzubringen. Das Produktsicherheitsgesetz sieht demgegenüber lediglich vor, dass allein der Hersteller die Verantwortung dafür trägt, ein neues Produkt vor der Inbetriebnahme zu prüfen und zu entscheiden, ob darüber hinaus eine CE-Kennzeichnung überhaupt erforderlich ist. Weitere Sicherheitsanforderungen ergeben sich speziell aus § 8 Absatz 1 ProdSG, aus denen sich konkrete Gestaltungsanforderungen auch für Roboter in der Pflege ableiten lassen, etwa die Maschinenverordnung, die mit Verweis auf die Maschinenrichtlinie auf relevante ISO-Normen verweist. In diesem Zusammenhang ist positiv festzustellen, dass solche Maschinen, zu denen auch Roboter in der Pflege gehören, zumindest mit einer Nothaltmöglichkeit auszustatten[80] sind, womit der Grad an Maschinenautonomie theoretisch begrenzt ist. Allerdings kann dies in der Praxis immer dann zu Problemen führen, wenn der Pflegeroboter nicht primär von Pflegekräften, sondern Pflegebedürftigen bedient wird etwa, wenn demente oder verwirrte Patienten diesen Schalter betätigen, während der Roboter eine wichtige Aufgabe erfüllt. Dies kann zu gefährlichen Situationen führen, etwa beim Transport oder dem Heben einer Person. Solche Gefährdungsrisiken sind allein vom Hersteller zu ermitteln und zu dokumentieren. Eine solche Dokumentation ist allerdings in der Regel nicht einsehbar.

Der Fall von Pflegerobotern illustriert beispielhaft, wie unterschiedlich der Grad von Regulierung ist, je nach Funktion und Anwendung des konkreten Systems. Da Pflegeroboter stets einen direkten oder indirekten Einfluss auf die pflegerische Versorgung und damit auf Leben und Gesundheit des Pflegebedürftigen haben, erscheint es sachgerecht, dass derartige Systeme an sich nicht den verhältnismäßig leichten Zulassungsvorschriften des ProdSG unterfallen. Hier gibt es Verbesserungsbedarf im Sinne einer Regelung, die ein einheitliches und strenges Zulassungsverfahren für Roboter im Pflegekontext regelt. Dennoch gehen auch die Vorschriften des Medizinprodukterechts nicht weit genug. Trotz der Einschaltung unabhängiger Prüfinstitute im Rahmen des Prüf- und Bewertungsverfahrens unterliegen Medizinprodukte mit hohen Risikoklassen (IIb und III) anders als Medikamente keiner amtlichen Zulassung. Bei der bereits erwähnten CE-Zertifizierung, werden ausschließlich die technische Funktionsfähigkeit und die prinzipielle Produktsicherheit überprüft, nicht aber der Nutzen, den Patienten aus dem Produkt ziehen können. Es gibt EU-weit keine einheitlichen Standards. Daher können Hersteller auch gezielt jene benannten Stellen aufsuchen, deren Prüfungen besonders lax sind. Bei Arzneimitteln dagegen kann das Bundesinstitut für Arzneimittel und Medizinprodukte bei Gefahr die Zulassung entziehen, die Anwendungsgebiete einschränken, Warnhinweise anbringen und Ärzte durch Rote-Hand-Briefe warnen. Schon bevor das Arzneimittel auf den Markt kommt, ist die Behörde in den gesamten Zulassungsprozess involviert. Erst wenn durch kontrollierte Studien Wirksamkeit, Unbedenklichkeit (Sicherheit) und Qualität nachgewiesen sind, erfolgt der Marktzugang.[81]

Ein effektives ADM-Kontrollsystem sollte Roboter im Pflegekontext, aufgrund ihres hohen Schadenspotenzials, äquivalent behandeln: Es bedarf EU-weit einheitlicher Standards, eines strengen Zulassungsverfahrens sowie Einsichts-und Kontrollmöglichkeiten für die zuständigen Aufsichtsinstitutionen. Etwa indem für Aufsichtsinstitutionen die

[80] Maschinenrichtlinie 2006/42/EG, Anhang 1, 1.2.4.3 (2019), URL: www.maschinenrichtlinie.de/fileadmin/dokumente/2006-42-EG_maschinenrichtlinie_de.pdf

[81] Bundesinstitut für Arzneimittel und Medizinprodukte: Arzneimittel, URL: https://www.bfarm.de/DE/Buerger/Arzneimittel/Arzneimittelzulassung/_node.html [Zugriff: 30.04.2019]

Möglichkeit geschaffen wird, das System zu analysieren und zu testen, beispielsweise über eine technische Schnittstelle. Weitere adäquate Maßnahmen für ADM-Systeme in Anwendungsbereichen mit dermaßen hohem Risikopotential, sind die Veröffentlichung einer Folgenabschätzung vorab sowie Kennzeichnungs- und Auskunftspflichten gegenüber Betroffenen.