

Staatliche Verantwortungsübernahme beim hoheitlichen Einsatz intransparenter Algorithmen

Daniel Busche

Inhaltsübersicht

I.	Transparenz: Interpretierbarkeit und Erklärbarkeit eines Algorithmenmodells	47
II.	Verantwortungsübernahme als Problem der Zurechenbarkeit	49
1.	Einsatz zur Entscheidungsunterstützung	50
2.	Einsatz zur automatisierten Entscheidungserstellung	51
a)	Veranlassung der Entscheidung durch eine Behörde	51
b)	Zurechenbarkeit beim Handeln außerhalb der übertragenen Aufgabe	52
aa)	Maßstab beim Handeln Privater	52
bb)	Übertragung auf den Einsatz von Algorithmen	53
c)	Zwischenergebnis	54
III.	Staatliche Verantwortung für die Herstellung und Darstellung einer Entscheidung	54
1.	Transparenz als Problem der Entscheidungsherstellung	55
a)	Intransparenz als Normalfall	55
b)	Vergleich mit menschlicher Entscheidungsfindung	56
c)	Übertragung auf den Einsatz von Algorithmen	57
d)	Technische Alternativen zur Transparenz	58
e)	Anwendungsbeispiele für den möglichen Einsatz intransparenter Algorithmen	59
f)	Zwischenergebnis	60
2.	Transparenz als Problem der Entscheidungsdarstellung	61
a)	Verfassungsrechtlich verankerte Begründungspflicht	61
b)	Begründung einer Entscheidung beim Einsatz von Algorithmen	62
aa)	Anforderungen auf den Algorithmeneinsatz übertragbar	62
bb)	Absehen von einer Begründung im Einzelfall möglich	63

c)	Zwischenergebnis	64
IV.	Fazit	65

Bereits seit einigen Jahren wird vermehrt über den hoheitlichen Einsatz von Algorithmen bzw. „künstlicher Intelligenz“ diskutiert.¹ Das Thema ist nicht ganz neu,² hat aber in der öffentlichen Wahrnehmung durch den technischen Fortschritt im Bereich der künstlichen Intelligenz – insbesondere durch gesteigerte Möglichkeiten des maschinellen Lernens³ – an Relevanz gewonnen. Ein zentrales Thema dabei und in der allgemeinen Diskussion zu Algorithmen ist deren mögliche Intransparenz. In diesem Kontext werden Algorithmen häufig als „Blackbox“ bezeichnet.⁴ Dem liegt die Vorstellung zugrunde, dass sich nicht oder nur teilweise erklären lässt, warum ein bestimmter Input zu einem bestimmten Output führt. Das Innenleben komplexer Algorithmen sei für menschliche Betrachter:innen unverständlich.⁵ Mit dieser Beschreibung geht regelmäßig die Befürchtung einher, dass entsprechende Modelle „unvorhersehbar“ agieren und dadurch Schaden anrichten könnten, wenn sie dennoch eingesetzt werden.⁶ Auch wird der Verlust an menschlicher Einflussnahmemöglichkeit beklagt.⁷ So wird befürchtet, dass der bzw. die Einzelne zum Objekt algorithmisch gesteuerter Entscheidungen werde.⁸

1 Vgl. aus der Vielzahl der Beiträge *Guckelberger*, Öffentliche Verwaltung im Zeitalter der Digitalisierung, 2019, Rn. 405 ff. u. 559 ff.; *Hermstrüwer*, AöR 145 (2020), 479; *Herold*, Demokratische Legitimation automatisiert erlassener Verwaltungsakte, 2020; *Martini/Ruschemeyer/Hain*, VerwArch 112 (2021), 1; *Roth-Isigkeit*, AöR 145 (2020), 321; *Unger*, Demokratische Herrschaft und Künstliche Intelligenz, in: *Unger/v. Ungern-Sternberg* (Hrsg.), Demokratie und Künstliche Intelligenz, 2019, S. 113.

2 Eine Zusammenfassung der historischen Diskussion findet sich bei *Gräwe*, Die Entstehung der Rechtsinformatik, 2011, S. 35 ff.

3 *Russell/Norvig*, Artificial Intelligence – A Modern Approach, 4. Aufl., (Global Edition), 2021, S. 42 ff.

4 Begriffsprägend *Pasquale*, The Black Box Society – The Secret Algorithms that Control Money and Information, 2015; gleich auch bspw. *Martini*, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, 2019.

5 *Martini*, JZ 2017, 1017 (1018).

6 Vgl. bspw. *Danaher*, Philosophy & Technology 29 (2016), 245 (249); *Roth-Isigkeit* (Fn. 1), 335 aus einer haftungsrechtlichen Perspektive.

7 Vgl. dazu und zu weiteren angenommenen Gefahren des Algorithmeneinsatzes die Vielzahl an Nachweisen bei *Barczak*, DÖV 2020, 997 (998).

8 Zu dieser die Technikdiskussion schon seit lange prägenden Befürchtung *Golla*, DÖV 2019, 673 (675 ff.).

Vor diesem Hintergrund beschäftigt sich der Beitrag mit der Frage, wie bedeutsam Anforderungen an die Transparenz von Algorithmen wirklich sind, damit der Staat die Verantwortung für deren Einsatz übernehmen kann und darf. Für eine sinnvolle Begrenzung sind die Ausführungen dabei auf die Verwaltungstätigkeit beschränkt. Auch ist darauf hinzuweisen, dass die im Beitrag angesprochenen Aspekte keinesfalls abschließend zu verstehen sind. Das Problem der Transparenz von Algorithmen kann auch aus anderen Perspektiven und mit einem jeweils ganz anderen Schwerpunkt erörtert werden. Zu denken ist insbesondere an den Umgang mit Diskriminierungsgefahren⁹ oder aber auch speziell an das Verhältnis der Transparenz zu Informationszugangsrechten.¹⁰

Der Beitrag gliedert sich in drei Teile: Im 1. Schritt wird dargelegt, wie der Begriff der Transparenz zur Beschreibung von Algorithmen zu verstehen ist. Im 2. Schritt werden die Maßstäbe herausgearbeitet, nach denen durch oder mit Hilfe von Algorithmen erstellte Entscheidungen einem Hoheitsträger rechtlich zurechenbar sind. Im 3. Schritt ist als inhaltlicher Schwerpunkt zu erörtern, inwiefern Probleme im Bereich der Transparenz sich auf die staatliche Verantwortung für die Herstellung und Darstellung einer Entscheidung auswirken.

I. Transparenz: Interpretierbarkeit und Erklärbarkeit eines Algorithmenmodells

Der Begriff der Transparenz im Kontext von Algorithmen hat eine normative und eine deskriptive Seite. Er wird benutzt, um normative Anforderungen an Algorithmenmodelle zu formulieren, aber auch um entsprechende Algorithmenmodelle und ihre Eigenschaften zu beschreiben. An dieser Stelle geht es zunächst um die letztgenannte, d. h. die technische Perspektive. Es soll das eigene Verständnis der Transparenz eines Algorithmenmodells als tatsächliche Grundlage der sich anschließenden rechtlichen Bewertung offenlegen werden.

Um die Transparenz eines Algorithmenmodells einzuordnen, kommt es auf deren Interpretier- und Erklärbarkeit an.¹¹ Ein Algorithmenmodell wird als interpretierbar bezeichnet, wenn bereits durch die Untersuchung

9 Mit diesem Schwerpunkt *Hermstrüwer* (Fn. 1), 479 ff.

10 Dazu *Guckelberger*, *Verwaltung* (Fn. 1), Rn. 501 ff.

11 Dazu das Standardwerk im Bereich der künstlichen Intelligenz, *Russell/Norvig*, *AI* (Fn. 3), S. 729 f.

des Modells an sich verständlich wird, warum ein gegebener Input zu einem bestimmten Ergebnis führt und wie sich dieses Ergebnis bei einem anderen Input verändern würde. Die Erklärbarkeit bezeichnet hingegen allgemein die Möglichkeit zu verstehen, warum ein Algorithmus ein bestimmtes Ergebnis für einen gegebenen Input ausgibt.¹² Dabei können auch externe Methoden zum Einsatz kommen. Ein interpretierbares Algorithmenmodell ist zugleich immer auch erklärbar. Ein erklärbares Algorithmenmodell ist hingegen nicht zwangsläufig interpretierbar, da sich die Erklärbarkeit auch erst durch den Einsatz externer Methoden ergeben kann.¹³ Unter dem Schlagwort der „Explainable AI“¹⁴ wird aktuell intensiv daran geforscht, entsprechende Methoden zu entwickeln. Dadurch soll es ermöglicht werden, auch das Entscheidungsverhalten komplexer Algorithmenmodelle, insbesondere im Bereich des „Deep Learning“, ganz oder zumindest teilweise erklärbar zu machen.¹⁵

Legt man ein enges Begriffsverständnis zu Grunde, könnte man dazu tendieren, allein auf die Interpretierbarkeit abzustellen, um von einem transparenten Algorithmenmodell auszugehen.¹⁶ Nur die Interpretierbarkeit erlaubt einen Blick in das Modell selbst. Allerdings dürfte es darauf nicht entscheidend ankommen. Der wesentliche Vorteil eines interpretierbaren Modells ist es, dass sich das Entscheidungsverhalten im Detail nachvollziehen lässt und deshalb unerwünschte Folgen weitestgehend vermieden werden können. Die Interpretierbarkeit ist eine Möglichkeit, Erklärbarkeit herzustellen. Auch für die Frage der Transparenz sollte deshalb allgemein auf die Erklärbarkeit des Modells abgestellt und nicht per se die Interpretierbarkeit vorausgesetzt werden. Lässt sich – mit oder ohne den Einsatz externer Methoden – im Detail erklären, warum ein Algorithmus ein bestimmtes Ergebnis für einen gegebenen Input ausgibt, reicht das aus, um von einem transparenten Algorithmenmodell auszugehen. Freilich dürfte in der Praxis die Erklärbarkeit nach gegenwärtigem Stand regelmä-

12 Russell/Norvig, AI (Fn. 3), S. 729.

13 Russell/Norvig, AI (Fn. 3), S. 729 f.

14 Vgl. beispielsweise den Sammelband Samek et al. (Hrsg.), Explainable AI: Interpreting, Explaining and Visualizing Deep Learning, 2019 oder Castelvecchi, Nature 538 (2016), 20.

15 Samek/Müller, Towards Explainable Artificial Intelligence, in: Samek et al., Explainable AI (Fn. 14), S. 1 (5 ff.).

16 So scheinbar zunächst Martini, Blackbox (Fn. 4), S. 28, der von „einer lückenlosen Inspektion“ spricht, aber an späterer Stelle doch auf die Erklärbarkeit des konkreten Entscheidungsverhaltens auch unter Einsatz externer Methoden abstellt (S. 43 f.).

ßig die Interpretierbarkeit voraussetzen. Die Methoden im Bereich der „Explainable AI“ sind noch weit davon entfernt, das Entscheidungsverhalten komplexer Algorithmen insbesondere im Bereich des „Deep Learning“ im Detail erklären zu können. So beschränken sich die „Erklärungen“ regelmäßig auf die visuelle Hervorhebung, der für die Entscheidung wichtigen Bereiche eines künstlichen neuronalen Netzwerks oder der besonders entscheidungsrelevanten Eingabedaten.¹⁷ Solche Visualisierungen sind in ihrer Aussagekraft beschränkt und können eine Erklärbarkeit im Detail regelmäßig nicht gewährleisten.¹⁸ Für die Zukunft sind allerdings gewisse Fortschritte zu erwarten, sodass Transparenz nicht mehr zwingend Interpretierbarkeit voraussetzen wird.

Darüber hinaus ist zu berücksichtigen, dass die Transparenz nicht absolut, sondern nur graduell beurteilt werden kann. Die meisten Algorithmenmodelle bewegen sich auf einer Skala zwischen vollständiger Intransparenz und vollständiger Transparenz. Als ein besonders transparentes Algorithmenmodell lässt sich beispielsweise ein durch maschinelles Lernen erstellter Entscheidungsbaum einordnen. Anhand der Abzweigungen lässt sich nachvollziehen, warum ein bestimmtes Ergebnis für einen gegebenen Input ausgegeben wird.²⁰ Auf der anderen Seite der Skala befinden sich hingegen regelmäßig die Algorithmen im Bereich des „Deep Learning“. Beispielsweise in tiefen neuronalen Netzwerken werden die Informationen verteilt über sehr viele (und teilweise versteckte) Zwischenschichten mit einer extrem hohen Anzahl an Gewichten gespeichert.²¹ Auf Grund dieser Komplexität lässt sich das Entscheidungsverhalten solcher Algorithmen mit den gegenwärtigen Methoden nicht im Detail erklären.²²

II. Verantwortungsübernahme als Problem der Zurechenbarkeit

Zunächst sind die Maßstäbe herauszuarbeiten, nach denen eine Entscheidung auch beim Einsatz von Algorithmen einer Behörde zugerechnet werden kann. Sollte der Einsatz intransparenter Algorithmen bereits die Zurechenbarkeit ausschließen, muss zur Frage der sonstigen Verantwortung

17 Samek/Müller, Explainable AI (Fn. 15), S. 16.

18 Samek/Müller, Explainable AI (Fn. 15), S. 16 f.

19 Vgl. für einen diesbezüglichen Überblick die Abbildung 1.3 in Ertel, Grundkurs Künstliche Intelligenz, 5. Aufl., 2021, S. 9.

20 Russell/Norvig, AI (Fn. 3), S. 729.

21 Ertel, Grundkurs (Fn. 19), S. 342.

22 Ertel, Grundkurs (Fn. 19), S. 343.

des Staates ohnehin nicht mehr Stellung bezogen werden. Dabei wird im Hinblick auf die Art des Algorithmeneinsatzes zwischen dem Einsatz zur bloßen Entscheidungsunterstützung und zur automatisierten Erstellung der Entscheidung unterschieden.²³

Dem Einsatz von Algorithmen zur Entscheidungsunterstützung und dem Einsatz zur automatisierten Erstellung ist zunächst gemein, dass sie sich grundsätzlich auf die Erstellung einer Entscheidung beziehen. Auch der Einsatz zur Entscheidungsunterstützung zielt darauf ab, sich zumindest mittelbar auf das Entscheidungsergebnis auszuwirken. Denkbar ist beispielsweise die Erstellung eines Entscheidungsvorschlages oder aber der Einsatz zur Klärung einer bestimmten Sachfrage im Vorfeld der eigentlichen Entscheidung (z. B. die Erstellung einer Wahrscheinlichkeitsprognose, ob ein Gebäude einsturzgefährdet ist, oder die Ermittlung des Risikos, ob möglicherweise eine Steuerverkürzung vorliegt, als Anlass für weitere Ermittlungen²⁴). Eine automatisierte Erstellung liegt hingegen vor, wenn sich der Einsatz der Algorithmen unmittelbar auf das Entscheidungsergebnis auswirkt, ohne dass eine menschliche Ergebniskontrolle stattfindet (vgl. §§ 28 Abs. 2 Nr. 4, 37 Abs. 5, 39 Abs. 2 Nr. 3 VwVfG sowie § 35a VwVfG).²⁵

1. Einsatz zur Entscheidungsunterstützung

Beim Einsatz von Algorithmen zur bloßen Entscheidungsunterstützung wird die Entscheidung letztlich allein durch die zuständige menschliche Bearbeiterin oder den zuständigen menschlichen Bearbeiter getroffen. Eine solche Entscheidung ist insofern grundsätzlich unproblematisch der Behörde des Bearbeiters oder der Bearbeiterin zurechenbar, ohne dass es auf die Eigenschaften des eingesetzten Algorithmus überhaupt

23 Dem Einsatz zur Entscheidungsunterstützung dürfte praktisch die größere Relevanz zu kommen *Herberger*, NJW 2018, 2825 (2825).

24 Vgl. für die Anforderungen an entsprechende Risikomanagementsysteme im Besteuerungsverfahren § 88 Abs. 5 AO.

25 So auch für den automatisierten Erlass von Verwaltungsakten VG Saarlouis, Urt. v. 12.3.2008 – 11 K2 146/05, juris Rn. 23; *Berger*, in: Knack/Henneke (Hrsg.), VwVfG, 11. Aufl., 2020, § 35a Rn. 34; *Ritgen*, in: Bauer et al. (Hrsg.), VwVfG (E-Government), 2. Aufl., 2014, § 37 Rn 102; *Stelkens*, in: Stelkens/Bonk/Sachs (Hrsg.), VwVfG, 9. Aufl., 2018, § 35a Rn. 17f.; *Tiedemann*, in: BeckOK VwVfG, 54. Ed. (Stand: 1.1.2022), § 37 Rn. 50.

ankommt.²⁶ Problematisch könnte allenfalls sein, wenn der oder die Bearbeiter:in vollständig auf die eingesetzten Algorithmen vertraut und beispielsweise einen Entscheidungsvorschlag ohne eigene inhaltliche Prüfung übernimmt.²⁷ Allerdings betrifft auch eine solche Konstellation nicht die Frage der Zurechenbarkeit im Außenverhältnis, sondern bewirkt je nach konkreter Konstellation möglicherweise die Rechtswidrigkeit der Entscheidung (z. B. aufgrund einer Ermessensunterschreitung).

2. Einsatz zur automatisierten Entscheidungserstellung

Findet hingegen keine menschliche Ergebniskontrolle statt und wird die Entscheidung automatisiert vom Algorithmus getroffen, könnte sich dieses Ergebnis allerdings ändern. So wird teilweise davon ausgegangen, dass es für die Zurechenbarkeit einer automatisiert erstellten Entscheidung erforderlich ist, dass die Behörde das Entscheidungsverhalten des Algorithmus im Detail nachvollziehen bzw. bestimmen kann und insofern der eingesetzte Algorithmus transparent sein muss.²⁸

a) Veranlassung der Entscheidung durch eine Behörde

Die Frage der Zurechenbarkeit einer Entscheidung zu einer Behörde stellt sich regelmäßig bei der Prüfung, ob eine „Maßnahme der Behörde“ im Sinne des § 35 VwVfG vorliegt. Insofern kann auf die dort geltenden Maßstäbe zurückgegriffen werden. Nach den diesbezüglichen Grundsätzen dürfte es für die Zurechnung ausreichen, dass die Behörde den Algorithmus bewusst für eine konkrete Aufgabe einsetzt und damit die Entscheidung in ihrem Namen veranlasst.²⁹ Dafür spricht zunächst, dass es für die Zurechenbarkeit grundsätzlich auf die Perspektive der Adressat:innen, d. h. den objektiven Empfänger:innenhorizont ankommt.³⁰ Die Adres-

26 In diese Richtung auch *Herold*, Legitimation (Fn. 1), S. 193 f.

27 *Hermstrüwer* (Fn. 1), 492; *Tönsmeyer-Uzuner*, Expertensysteme in der öffentlichen Verwaltung, 2000, S. 64 f.; *Unger* (Fn. 1), S. 126 spricht vom „*automation bias*“.

28 So *Lazaratos*, Rechtliche Auswirkungen der Verwaltungsautomation auf das Verwaltungsverfahren, 1990, S. 276 f.

29 *Guckelberger*, Verwaltung (Fn. 1), Rn. 627 f.; *Statkens* (Fn. 25), § 35 Rn. 67 f.

30 *Ramsauer*, in: *Kopp/Ramsauer* (Hrsg.), *VwVfG*, 22. Aufl., 2021, § 35 Rn. 54 m. w. N.

sat:innen dürften allerdings regelmäßig nicht erkennen können, ob bei der Entscheidung ausnahmsweise ein für die Behörde intransparenter und nicht zuverlässig vorhersehbar agierender Algorithmus mitgewirkt und insofern die Entscheidung teilweise „autonom“ getroffen hat. Aus ihrer Sicht handelt es sich unabhängig von den Eigenschaften des eingesetzten Algorithmus um eine Entscheidung der Behörde.

Wie weit die Zurechenbarkeit im Außenverhältnis grundsätzlich reicht, zeigt auch der Vergleich mit dem behördlichen Einschalten von Privaten als Verwaltungshelfer:innen.³¹ Beim behördlichen Einsatz von Verwaltungshelfer:innen werden die von der Privatperson erstellten Bescheide der Behörde selbst dann zugerechnet, wenn die Behörde keine Möglichkeit hat, den Inhalt der Bescheide zu kontrollieren oder auch nur zu kennen.³² Erforderlich ist lediglich, dass der Verwaltungshelfer mit Wissen und Wollen der Behörde tätig wird und die Bescheide im Namen der Behörde ergehen, auch wenn deren Inhalt selbstständig bestimmt wird.³³ Diese Voraussetzungen wären auch beim Einsatz wirklich „autonomer“ und damit nicht vorhersehbar agierender Algorithmen erfüllt. Probleme im Bereich der Transparenz und selbst der Vorhersehbarkeit schließen die Zurechnung nicht aus, solange die Behörde den Algorithmus bewusst für eine konkrete Aufgabe einsetzt und damit die Entscheidung in ihrem Namen veranlasst.

b) Zurechenbarkeit beim Handeln außerhalb der übertragenen Aufgabe

aa) Maßstab beim Handeln Privater

Bei der Einschaltung Privater endet die Zurechnung allerdings, wenn es gänzlich an einer behördlichen Veranlassung für die konkrete Entscheidung fehlt, d. h., wenn der oder die Private außerhalb der ihm oder ihr übertragenen Aufgabe oder gänzlich selbstständig (z. B. Aufstellen von Verkehrszeichen an öffentlichen Straßen durch private Grundstückseigentümer:innen) tätig wird.³⁴ In einer solchen Konstellation liegt ein „Schein-“

31 Ähnlich Stelkens (Fn. 25), § 35 Rn. 67 f.

32 Dazu Guckelberger, Verwaltung (Fn. 1), Rn. 628.

33 BVerwG, Urt. v. 23.8.2011 – 9 C 2.11, BVerwGE 140, 245 (247); zustimmend Ramsauer (Fn. 30), § 35 Rn. 65.

34 Guckelberger, Verwaltung (Fn. 1), Rn 628; Stelkens (Fn. 25), § 35 Rn. 60.

bzw. „*Nichtakt*“³⁵ vor, der keinem Hoheitsträger als Rechtsakt zugerechnet werden kann und dem deshalb keine rechtliche Wirkung zukommt.³⁶

bb) Übertragung auf den Einsatz von Algorithmen

Insofern stellt sich die Frage, ob auch diese Grundsätze auf Algorithmen übertragbar sind. Dabei geht es um die Konstellation, dass der eingesetzte Algorithmus die Grenzen des für ihn vorgesehenen Aufgabenbereichs verlässt und eine Entscheidung trifft, die außerhalb der ihm zugewiesenen Aufgabe liegt. Einer solchen Situation dürfte kurz- bis mittelfristig keine große praktische Relevanz zukommen.³⁷ Allerdings sind Grenzfälle denkbar, bei denen der Aufgabenbereich beispielsweise aufgrund einer fehlerhaften Programmierung geringfügig überschritten wird.

In der Sache sprechen wesentliche Gründe gegen eine Übertragung.³⁸ Zunächst unterscheiden sich die analoge und die digitale Konstellation an dieser Stelle in einem zentralen Punkt. Beim Handeln einer Privatperson liegt keinerlei Veranlassung der Behörde vor. Auch wenn die Behörde den Privaten eine hoheitliche Aufgabe übertragen hat, maßt er sich andere, nicht übertragene hoheitliche Befugnisse lediglich bei Gelegenheit der ihm übertragenen Aufgabe an. Wenn hingegen ein von der Behörde eingesetzter Algorithmus eine Entscheidung trifft, ist diese Entscheidung auch durch die Behörde veranlasst, wenn sie außerhalb der übertragenen Aufgabe liegt. Die Behörde hat den Algorithmus entwickelt oder entwickeln lassen und sich dafür entschieden, ihn einzusetzen. Die Behörde nutzt den Algorithmus bewusst für eine effiziente Aufgabenerfüllung. Insofern ist ihr auch das möglicherweise problematische Verhalten des Algorithmus zuzurechnen.³⁹ Das gilt unabhängig davon, ob die Behörde einen solchen Algorithmus überhaupt hätte einsetzen dürfen.⁴⁰ Die potenzielle Unzulässig-

35 Zwischen „Schein-“ und „Nichtakt“ differenzierend *Münkler*, Der Nichtakt. Eine dogmatische Rekonstruktion, 2015, S. 160 ff.

36 *Stelkens* (Fn. 25), § 44 Rn. 5.

37 Algorithmen, die den ihnen zugewiesenen Aufgabenbereich selbstständig erweitern, dürften die aktuellen technischen Möglichkeiten weit übersteigen, vgl. *Russell/Norvig*, AI (Fn. 3), S. 730 f.; ähnlich auch *Wischmeyer*, AÖR 143 (2018), 1 (14 f.).

38 Mit diesem Ergebnis auch *Guckelberger*, Verwaltung (Fn. 1), Rn 628.

39 Ähnlich *Guckelberger*, Verwaltung (Fn. 1), Rn. 628.

40 In diese Richtung auch *Polomski*, Der automatisierte Verwaltungsakt, 1993, S. 85 ff.

sigkeit des Einsatzes kann zwar die Rechtswidrigkeit der Entscheidung begründen, aber auch rechtswidrig zustande gekommene Entscheidungen sind einer Behörde grundsätzlich zurechenbar.

Zum anderen spricht auch der Schutz der Bürger:innen für eine Zurechnung. So lassen sich gegenüber der eigenmächtig handelnden Privatperson gegebenenfalls zivilrechtliche Haftungsansprüche geltend machen. Algorithmen scheiden nach geltender Rechtslage hingegen als Haftungs-subjekte aus. Unabhängig von der Frage, ob in dieser Konstellation eine weitergehende Haftung rechtspolitisch wünschenswert wäre,⁴¹ dürften sich nach geltender Rechtslage negative Folgen der Entscheidung häufig überhaupt nur ausgleichen lassen, wenn die Entscheidung der Behörde zugerechnet wird. Die Zurechnung ist Voraussetzung dafür, dass dem Bürger oder der Bürgerin je nach Konstellation zumindest ein Folgenbeseitigungs-anspruch, ein Anspruch wegen enteignenden bzw. enteignungsgleichen Eingriff oder ein Anspruch aus Aufopferung zustehen könnte.⁴² Ein An-spruch aus Amtshaftung (§ 839 BGB) setzt hingegen eine schuldhafte Amtshandlung eines menschlichen Amtswalters oder einer menschlichen Amtswalterin voraus und kann aus diesem Grund häufig keinen angemes-senen Schadensausgleich gewährleisten.

c) Zwischenergebnis

Es zeigt sich, dass es für die Zurechnung einer Entscheidung zur Behörde auf die Transparenz des eingesetzten Algorithmenmodells nicht ankommt. Vielmehr reicht es aus, dass die Behörde einen Algorithmus bewusst ein-setzt und damit die Entscheidung veranlasst.

III. Staatliche Verantwortung für die Herstellung und Darstellung einer Entscheidung

Die Vorstellung des staatlichen Einsatzes intransparenter Algorithmen dürfte vielfach ein erhebliches Störgefühl hervorrufen und intuitiv auf Ab-

41 Dazu *Roth-Isigkeit* (Fn. 1), 331 ff. und *Martini/Ruschemeyer/Hain* (Fn. 1), 32 ff.

42 Vgl. zu diesen verschuldensunabhängigen Haftungsansprüchen bei der Ent-scheidungsautomation *Martini/Ruschemeyer/Hain* (Fn. 1), 4 ff.

lehnung stoßen.⁴³ Insofern könnte auch bezweifelt werden, dass der Staat die Verantwortung für den Einsatz entsprechender Algorithmen übernehmen darf. Angenommene Probleme dürften sich dabei sowohl auf die Herstellung als auch die Darstellung⁴⁴ einer hoheitlichen Entscheidung beziehen.

1. Transparenz als Problem der Entscheidungsherstellung

Damit ein Hoheitsträger die Verantwortung für eine Entscheidung übernehmen kann, muss er den Prozess der Entscheidungsherstellung in einem gewissen Umfang steuern und kontrollieren können. Beim Einsatz intransparenter Algorithmen wird diesbezüglich teilweise vor einem Verlust an menschlicher Einflussnahmemöglichkeit gewarnt.⁴⁵ Insofern könnte man den Einsatz intransparenter Algorithmen auf Grund der Gefahr des Verlustes der Steuerungsmöglichkeit für ausgeschlossen halten. Im Kontrast dazu möchte ich an dieser Stelle aufzeigen, dass die Steuer- und Kontrollierbarkeit der Entscheidungsherstellung nicht per se die Transparenz der eingesetzten Algorithmenmodelle voraussetzt. Je nach Kontext kann auf die Transparenz der eingesetzten Algorithmen verzichtet werden.⁴⁶

a) Intransparenz als Normalfall

Das Problem der Transparenz ist nicht neu, sondern prägt moderne Gesellschaften allgemein⁴⁷ und den Einsatz von Technik im Besonderen. Menschliche Anwender:innen vertrauen in vielfältiger Weise auf das Funktionieren technischer Einrichtungen, ohne dass ihnen deren Funktions-

43 Vgl. bspw. *Steger/Sonnabend*, Die unheimliche Macht der Algorithmen, Handelsblatt v. 9.5.2017.

44 Zu dieser grundlegenden begrifflichen Unterscheidung *Luhmann*, Recht und Automation in der öffentlichen Verwaltung, 1966, S. 51.

45 *Stelkens*, Der vollständig automatisierte Erlass eines Verwaltungsakts als Regelungsgegenstand des VwVfG, in: Hill/Kugelmann/Martini (Hrsg.), Digitalisierung in Recht, Politik und Verwaltung, S. 81 (101), der allerdings insofern wohl von sich selbstständig fortentwickelnden Algorithmen ausgeht; zustimmend *Guckelberger*, Verwaltung (Fn. 1), Rn. 592; vgl. auch *Unger* (Fn. 1), 118 ff.

46 Ähnlich *Wischmeyer* (Fn. 37), 8, 43 ff.

47 *Luhmann*, Die Kontrolle von Intransparenz, 2017, S. 96, spricht von der modernen Gesellschaft als „*Symphonie der Intransparenz*“.

weise im Detail bekannt ist oder auch nur bekannt sein kann. Dabei dürfte das Vertrauen der Anwender:innen häufig zumindest auf der Erwartung beruhen, dass eine kompetente Dritte oder ein kompetenter Dritter (Entwickler:in, Hersteller:in etc.) die Funktionsweise erklären kann und insofern das zuverlässige Funktionieren garantiert. Auf der anderen Seite stößt auch diese „potenzielle“ Erklärbarkeit vielfach an ihre Grenzen. So werden selbst in einem so sensiblen Bereich wie der Medizin Verfahren eingesetzt, deren ursächlicher Wirkmechanismus nicht bekannt ist, solange das zuverlässige Funktionieren empirisch nachgewiesen ist (z. B. die Narkose)⁴⁸. Speziell im Hinblick auf Algorithmen gilt, dass das Problem der Unverständlichkeit komplexer Algorithmen alles andere als neu ist. Die Diskussion dazu wird seit mehr als 50 Jahren geführt.⁴⁹ Die Wirkweise moderner Computerprogramme dürfte regelmäßig kein einzelner Mensch mehr im Detail erklären oder nachvollziehen können. Auch dürfte nahezu jedes komplexe Computerprogramm irgendeine vom Entwicklungsteam nicht vorhersehbare Fehlfunktion („Bugs“) aufweisen. Nichtsdestotrotz werden entsprechende Algorithmen und Programme vielfältig eingesetzt, ohne dass darin bisher ein unvertretbares Risiko gesehen wird oder daraus regelmäßig unverhältnismäßige Schäden resultieren würden.

b) Vergleich mit menschlicher Entscheidungsfindung

Auch im Bereich hoheitlicher Tätigkeit ist das Vertrauen auf eine „Blackbox“ als Entscheidungsträger kein neues Phänomen, sondern in gewisser Weise der Normalfall.⁵⁰ Bereits der Begriff entstammt ursprünglich nicht der Informatik, sondern der Psychologie, genauer dem Behaviorismus. Es dürfte sich um eine allgemein anerkannte Erkenntnis handeln, dass sich das menschliche Entscheidungsverhalten nicht im Detail erklären lässt.⁵¹ So ist anerkannt, dass bei der menschlichen Entscheidungsfindung vielfältige Gründe auf sozialer, psychologischer oder biologischer Ebene eine Rolle spielen können, die für einen Dritten oder eine Dritte als Beobachter:in nicht erkennbar sind und die teilweise auch der Entscheiderin oder dem Entscheider selbst nicht bewusst sein dürften.⁵² Konkret aus der

48 Eckle/Hucklenbruch/Todorovic, *Anaesthesia* 58 (2009), 1144 (1148).

49 Dazu Passig, *Merkur* Nr. 823 (2017), 16 (18 ff.).

50 Vgl. Wischmeyer (Fn. 37), 8, 43 ff.

51 Hermstrüwer (Fn. 1), 489.

52 Vgl. grundlegend in dieser Richtung Simon, *The Quarterly Journal of Economics* 69 (1955), 99 (104) oder Luhmann, *Intransparenz* (Fn. 47), S. 102; speziell

Perspektive der Rechtsanwendung ist insbesondere weitgehend anerkannt, dass das Vorverständnis des jeweiligen Rechtsanwenders oder der jeweiligen Rechtsanwenderin die Entscheidung maßgeblich prägt.⁵³ Im Hinblick auf eine konkrete Entscheidungsfindung gibt es keine Möglichkeit, die die Entscheidung objektiv determinierenden Faktoren letztverbindlich zu bestimmen. Insofern lässt sich auch nicht mit letzter Gewissheit sicherstellen, dass menschliche Amtswalter:innen die „richtige“ oder eine „vertretbare“⁵⁴ Entscheidung treffen werden und schon gar nicht, dass diese Entscheidung auf den „richtigen“ bzw. „vertretbaren“ Gründen beruht. Dennoch wird in der Verwaltungspraxis gerade darauf vertraut. Dieses Vertrauen ist nicht naiv oder unbegründet, sondern durch Erfahrungswissen belegt. Es lässt sich aus vergangenen Erfahrungen, die Erwartung generieren, dass menschliche Amtswalter:innen regelmäßig zumindest eine vertretbare Entscheidung treffen werden und hinreichende Gründe dafür angeben können. Bewusste oder grobe Fehlentscheidungen stellen hingegen eine seltene Ausnahme dar.

c) Übertragung auf den Einsatz von Algorithmen

Auf diesen Einsichten aufbauend lassen sich die Maßstäbe entwickeln, die für den Einsatz von Algorithmen gelten sollen. Der Vergleich mit menschlichen Amtswaltern zeigt zunächst, dass der Transparency im Bereich der Entscheidungsherstellung⁵⁵ kein absoluter Wert zukommt. Transparency ist auch bezüglich des Algorithmeneinsatzes nicht als Selbstzweck, sondern als ein mögliches Mittel, um ein bestimmtes Ziel zu erreichen, zu verstehen. Die Steuer- und Kontrollierbarkeit eines Algorithmus setzt voraus,

im Hinblick auf juristische Entscheidungen *Spamann/Klöhn*, The Journal of Legal Studies 45 (2016), 255 (255 ff.).

- 53 Dazu ausführlich das grundlegende Werk von *Esser*, Vorverständnis und Methodenwahl bei der Rechtsfindung, 1972, insb. S. 136 ff.
- 54 Die eine „richtige“ Entscheidung dürfte im Bereich des Rechts allenfalls als „Fiktion“ oder „regulative Idee“ existieren, vgl. *Stelkens*, Die Idee der einzigen richtigen, sich aus dem Gesetz ergebenden Entscheidung und ihre Bedeutung für die deutsche Rechtswissenschaft, in: *van Oostrom/Weth* (Hrsg.), Festschrift für Maximilian Herberger, 2017, S. 895 (895 ff.).
- 55 Im Hinblick auf möglicherweise bestehende Informationszugangsrechte kommt der Transparency hingegen eine rechtliche Bedeutung an sich zu.

dass mit hinreichender⁵⁶ Wahrscheinlichkeit erwartet werden kann, dass die Entscheidungen nach den vorgegebenen Maßstäben „richtig“ sind.⁵⁷ Dadurch soll sichergestellt werden, dass ein ausreichender menschlicher Einfluss auf das algorithmisch erstellte Entscheidungsergebnis verbleibt.⁵⁸ Diesbezügliche Defizite sind anzunehmen, wenn die Gefahr besteht, dass der Algorithmus sich im Rahmen seiner Entscheidungsfindung nicht an die vorgegebenen Maßstäbe hält und deshalb eine sachlich unrichtige oder sogar rechtswidrige Entscheidung erlassen wird.⁵⁹ In diesem Kontext kann der Transparenz der eingesetzten Algorithmen eine entscheidende Bedeutung zukommen, weil sich so gewährleisten lässt, dass die eingesetzten Algorithmen zuverlässig die gewünschte Entscheidung treffen. Im Umkehrschluss folgt daraus aber auch, dass auf die Transparenz verzichtet werden kann, wenn sich auf andere Art und Weise sicherstellen lässt, dass die gewünschten Entscheidungsergebnisse erzielt werden.⁶⁰

d) Technische Alternativen zur Transparenz

Die Transparenz ist nicht die einzige Möglichkeit, um das zuverlässige Funktionieren und damit die Steuer- und Kontrollierbarkeit eines Algorithmenmodells sicherzustellen. Als technische Alternativen sind insbesondere das „Blackbox-Testing“ sowie die „Redundanz“ eines Systems zu nennen.⁶¹

„Blackbox-Testing“ bezeichnet Testmethoden für Algorithmen, bei denen die interne Entscheidungslogik nicht im Detail nachvollzogen werden muss, um die Funktionsweise zu überprüfen.⁶² Vielmehr wird über eine

56 Welches Maß an Wahrscheinlichkeit insofern erforderlich ist, richtet sich nach dem jeweiligen Kontext des Algorithmeneinsatzes. Es dürfte insbesondere von der konkreten Entscheidungssituation, den Möglichkeiten, die Entscheidung zu revidieren, dem Ausmaß möglicher Schäden sowie insbesondere den zur Verfügung stehenden Alternativen abhängen.

57 Ähnlich *Guckelberger*, Verwaltung (Fn. 1), Rn. 592; *Martini*, in: *Kahl/Ludwigs* (Hrsg.), Handbuch des Verwaltungsrechts, Bd. 1, 2021, Rn. 88; *Stelkens*, Automatisierter Erlass (Fn. 45), S. 102.

58 *Stelkens*, Automatisierter Erlass (Fn. 45), S. 102.

59 Ähnlich *Martini* (Fn. 57), Rn. 88.

60 Auch *Russell/Norvig*, AI (Fn. 3), S. 672, weisen in diesem Kontext darauf hin, dass es entscheidend darauf ankommt, ob ein lernender Algorithmus zuverlässig funktioniert.

61 *Passig* (Fn. 49), 26 f.

62 *Passig* (Fn. 49), 27.

möglichst hohe Anzahl an Tests sichergestellt, dass es nicht oder mit einer sehr geringen Wahrscheinlichkeit zu unerwünschten Ergebnissen kommen wird. Letztlich beruhen entsprechende Testverfahren auf der Beobachtung von Korrelationen und verzichten auf eine kausale Erklärung im Detail. Das dürfte allerdings häufig ausreichen, sofern die Korrelation nur angemessen hoch ist. Zur weiteren Reduktion des Risikos lässt sich auf das Prinzip der „Redundanz“ eines Systems abstellen.⁶³ Sicherheitskritische Komponenten in technischen Systemen können mehrfach redundant konstruiert werden, d. h. die Fehlfunktion einer Komponente wird durch das Vorhandensein weiterer Komponenten mit der gleichen Aufgabe abgesichert.⁶⁴ Das Prinzip lässt sich auch bei der Entwicklung komplexer Algorithmen nutzen. So ist es vorstellbar, dass unterschiedliche Modelle nach unterschiedlichen Prinzipien für dieselbe Aufgabe entwickelt werden und eine entsprechende Entscheidung gegenüber dem Bürger nur getroffen wird, wenn alle Modelle zum jeweils gleichen Ergebnis kommen. Die Absicherung kann sich aber auch nur auf einen Teilbereich beziehen. Beispielsweise könnte im Hinblick auf den sogleich näher beschriebenen Algorithmus zur koordinierten Steuerung unterschiedlicher Ampelanlagen eine dezentrale Sicherung in den einzelnen Ampelanlagen dafür sorgen, dass die Anzeige von „feindlichem Grün“⁶⁵ auch bei einem Fehler des Steuerungsalgorithmus technisch ausgeschlossen wird.

e) *Anwendungsbeispiele für den möglichen Einsatz intransparenter Algorithmen*

Das je nach Kontext auf die Transparenz eines Algorithmenmodells verzichtet werden kann, verdeutlichen auch zwei konkrete Anwendungsbeispiele für den möglichen hoheitlichen Einsatz intransparenter Algorithmen:

Komplexe Algorithmen können zur intelligenten Verkehrssteuerung eingesetzt werden.⁶⁶ Dabei steuert ein Algorithmus beispielsweise die Verkehrsampeln in einem bestimmten Verkehrsabschnitt koordiniert und an

63 Vgl. [https://de.wikipedia.org/wiki/Redundanz_\(Technik\)](https://de.wikipedia.org/wiki/Redundanz_(Technik)), Stand: 24.2.2022.

64 *Passig* (Fn. 49), 26.

65 Feindliches Grün bezeichnet die Abgabe einander widersprechender Lichtzeichen durch eine Ampelanlage (vgl. BGH, Urt. v. 18.12.1986 – III ZR 242/85, NJW 1987, 1945 (1945 ff.)).

66 Vgl. für ein Pilotprojekt in Utah (US) Wang *et al.*, *Transportation Research Record* 2674 (2020), 172 (172 ff.).

die jeweilige Verkehrslage in Echtzeit angepasst. Das Entscheidungsverhalten eines solchen Algorithmus dürfte für einen menschlichen Anwender regelmäßig nicht im Detail nachvollziehbar, d. h. weder interpretier- noch erklärbar sein. Der Algorithmus wird gerade eingesetzt, weil die Aufgabe die menschliche Leistungsfähigkeit übersteigt und entsprechend komplexe Berechnungen von Menschen nicht ausgeführt werden können.⁶⁷ Sollte der Algorithmus allerdings seine vorgegebene Aufgabe, die Reduzierung der durchschnittlichen Fahrtzeit, erfüllen und gleichzeitig durch ausreichende Sicherheitsvorkehrungen der Eintritt von Schäden mit hinreichender Sicherheit ausgeschlossen sein, ist kein sachlicher Grund ersichtlich, warum allein die Intransparenz dazu führen soll, dass ein entsprechender Algorithmus nicht eingesetzt und das davon ausgehende Potenzial ungeutzt bleibt.

Ein weiteres Beispiel ist das EasyPASS-System⁶⁸ zur automatisierten Grenzkontrolle an Flughäfen. Dabei wird der Ausweis des Einreisenden eingescannt, das abgespeicherte biometrische Bild mit dem Live-Bild abglichen und im Falle der Übereinstimmung die Einreise ermöglicht. Für diese Aufgabe kommt höchstwahrscheinlich ein künstliches neuronales Netz zum Einsatz.⁶⁹ Obwohl die Entscheidungsstruktur des Algorithmus nicht im Detail nachvollzogen werden kann, ist ein solcher Einsatz unproblematisch, solange der Algorithmus zuverlässig funktioniert. Das liegt in dieser Situation auch daran, dass die Entscheidung durch einen menschlichen Amtswalter ohne größere Schäden korrigiert werden kann, indem die Einreise manuell bewilligt wird.

f) Zwischenergebnis

Ein gewisses Restrisiko unerwünschten Verhaltens wird sich auch durch diese und gegebenenfalls weitere Vorkehrungen nicht vollständig ausschließen lassen. Welches Maß an Risiko insofern hinnehmbar ist, dürfte von der konkreten Entscheidungssituation, den Möglichkeiten, die Entscheidung zu revidieren, dem Ausmaß möglicher Schäden sowie insbeson-

67 Wischmeyer, § 20 Regierungs- und Verwaltungshandeln durch KI, in: Ebers et al. (Hrsg.), KI und Robotik, 2020, S. 612 (Rn. 29 f.).

68 Vgl. https://www.easypass.de/EasyPass/DE/Was_ist_EasyPass/home_node.html, Stand: 24.2.2022.

69 Künstliche neuronale Netze können insbesondere im Bereich der Bilderkennung ihre Stärken ausspielen und werden deshalb in diesem Bereich ganz überwiegend eingesetzt, vgl. Ertel, Grundkurs (Fn. 19), S. 330.

dere den zur Verfügung stehenden Alternativen abhängen. Letztlich erfordert jede Entscheidung über den Einsatz von komplexen Algorithmen eine umfassende Abwägung im Hinblick auf den konkreten Einsatzkontext. Die Intransparenz eines Algorithmus ist dabei allerdings kein Ausschlusskriterium, sondern nur ein Kriterium von vielen.

2. Transparenz als Problem der Entscheidungsdarstellung

Aus der Perspektive der Bürger:innen ist die Transparenz eines Algorithmus insbesondere als Problem der Entscheidungsdarstellung zu erörtern.⁷⁰

a) Verfassungsrechtlich verankerte Begründungspflicht

Grundsätzlich hat jeder Bürger und jede Bürgerin einen im Grundrechtschutz sowie im Rechtsstaats- und Demokratieprinzip verfassungsrechtlich verankerten⁷¹ Anspruch darauf, eine Begründung zu erhalten, wenn er oder sie in seinen oder ihren subjektiven Rechten betroffen ist.⁷² Konkretisiert wird dieser Anspruch einfach-rechtlich insbesondere durch das Erfordernis, einen Verwaltungsakt gem. § 39 Abs. 1 VwVfG zu begründen. Aus Sicht der Bürger:innen kommen der Begründung dabei im Wesentlichen zwei Funktionen zu. Primär dient sie dazu, das Entscheidungsverhalten der Behörde nachzuvollziehen und dadurch auf seine Rechtmäßigkeit hin kontrollieren zu können. So ist die Begründung regelmäßig die notwendige Grundlage für die Bürger:innen, um einzuschätzen, ob das Einlegen eines Rechtsbehelfs gegen die getroffene Entscheidung erfolgsversprechend ist.⁷³ Darüber hinaus führt die Begründung dazu, dass die Bürger:innen eher geneigt sein werden, eine staatliche Entscheidung, deren Gründe sie nachvollziehen können, zu akzeptieren.⁷⁴ In dieser Hinsicht dient die Be-

70 Ähnlich auch Wischmeyer (Fn. 37), 54 ff.

71 Umfassend zu den verfassungsrechtlichen Grundlagen der Begründung *Kischel*, Die Begründung – zur Erläuterung staatlicher Entscheidungen gegenüber dem Bürger, 2003, S. 63 ff.

72 BVerwG, Urt. v. 7.5.1981 – 2 C 42.79, DVBl 1982, 198 (199); BVerfG, Beschl. v. 19.1.1999 – 1 BvR 2161/94, BVerfGE 99, 81 (189 ff.); BVerfG, Urt. v. 20.2.2001 – 2 BvR 1444/00, BVerfGE 103, 142 (160 f.); *Stelkens* (Fn. 25), § 39 Rn. 2 f.

73 *Stelkens* (Fn. 25), § 39 Rn. 1.

74 Ausführlich dazu *Kischel*, Begründung (Fn. 71), S. 52 ff.; *Stelkens* (Fn. 25), § 39 Rn. 1.

gründung der Herstellung von Legitimation durch die mögliche Akzeptanz der Betroffenen und stärkt das Vertrauen in staatliche Entscheidungen allgemein.⁷⁵

b) Begründung einer Entscheidung beim Einsatz von Algorithmen

Werden intransparente Algorithmen eingesetzt besteht diesbezüglich das Problem, dass das Entscheidungsverhalten nicht im Detail nachvollzogen werden kann. Die Details der Entscheidungsfindung können deshalb auch nicht gegenüber den Bürger:innen offenbart werden. Allerdings kann dieses Problem ganz grundsätzlich relativiert werden. So ist im Hinblick auf menschliche Entscheider:innen allgemein anerkannt, dass zwischen der Herstellung und der Darstellung einer Entscheidung zu unterscheiden ist.⁷⁶ Die Begründung bezieht sich dabei auf die Darstellung der Entscheidung gegenüber dem Bürger. Sie zielt nicht darauf ab, den „echten“ Gang der Entscheidungsfindung im Detail zu rekonstruieren. Diejenigen Gründe, die ex ante auf empirischer Ebene zu der betreffenden Entscheidung geführt haben, können und müssen nicht dargelegt werden.⁷⁷ Vielmehr ergibt sich aus der Begründung, inwiefern die Entscheidung aus dem Recht hergeleitet werden kann. In der Begründung legt der Entscheider oder die Entscheiderin gegenüber den Bürger:innen dar, inwiefern die Entscheidung nach seiner oder ihrer subjektiv „wahren“ und nicht nur vorgeschobenen⁷⁸ Ansicht rechtlich richtig ist.⁷⁹ Diese Grundsätze sind im Hinblick auf menschliche Amtswalter:innen entwickelt worden, lassen sich aber auch auf den Algorithmeneinsatz übertragen.

aa) Anforderungen auf den Algorithmeneinsatz übertragbar

Beim Einsatz von Algorithmen ist es erforderlich, aber auch ausreichend, dass die wesentlichen tatsächlichen und rechtlichen Gründe i. S. d. § 39 Abs. 1 VwVfG gegenüber den Bürger:innen angegeben werden, um

75 Tiedemann (Fn. 25), § 39 Rn. 1.

76 Vgl. Fn. 35.

77 Kischel, Begründung (Fn. 71), S. 358 f.; Wischmeyer (Fn. 37), 55, 59.

78 Stelkens (Fn. 25), § 39 Rn. 47.

79 Mit ausführlicher Begründung Kischel, Begründung (Fn. 71), S. 9 ff., 358 f.; Wischmeyer (Fn. 37), 56 f.

der Begründungspflicht zu genügen. Nur so dürften sich die weiter oben herausgearbeiteten Funktionen der Begründung erfüllen lassen.⁸⁰ Den Bürger:innen Einblick in den Algorithmus selbst, d. h. in dessen Quellcode zu gewähren, dürfte hingegen nicht dabei helfen, die Entscheidung nachvollziehen zu können.⁸¹ Für technische Laien ist der Quellcode eines Algorithmus unverständlich. Darüber hinaus können zu umfassende Informationen auch dazu führen, dass der Bürger diese aus zeitökonomischen Gründen gar nicht beachtet und sie deshalb keine echte Aufklärungswirkung entfalten können.⁸² Eine auf die wesentlichen tatsächlichen und rechtlichen Gründe reduzierte Begründung ist insgesamt vorzugswürdig.

Die Begründungsanforderung dürfte so lange erfüllbar sein, wie der Algorithmus nur zur bloßen Entscheidungsunterstützung eingesetzt wird. Insofern kann der oder die für die abschließende Entscheidung zuständige Amtswalter:in die wesentlichen tatsächlichen und rechtlichen Gründe i. S. d. § 39 Abs. 1 VwVfG angeben, ohne die Wirkweise des Algorithmus offenbaren zu müssen. Trifft der Algorithmus allerdings selbst die Entscheidung, dürfte auch eine solche reduzierte Begründung eine besondere technische Herausforderung darstellen. In einfachen Konstellationen wird man eine ausreichende Begründung noch durch den Einsatz von Textbausteinen erzeugen können. Darüber hinaus dürfte es hingegen kurz- bis mittelfristig regelmäßig nicht gelingen, eine den Anforderungen des § 39 Abs. 1 VwVfG genügende Begründung in natürlicher Sprache automatisiert zu erstellen.⁸³ Aber auch dieses Problem wird sich zumindest teilweise lösen lassen.⁸⁴

bb) Absehen von einer Begründung im Einzelfall möglich

Unter bestimmten Voraussetzungen kann auf eine Begründung verzichtet werden. Trotz ihrer verfassungsrechtlichen Verankerung gilt die Begründungspflicht nicht absolut und auch nicht für jede Entscheidung in glei-

80 In diese Richtung auch *Wischmeyer* (Fn. 37), 54 ff.

81 *Martini/Nink*, NVwZ-Extra 2017, 1 (11 Fn. 108); *Wischmeyer* (Fn. 37), 52 f.

82 *Golla* (Fn. 8), 674.

83 So beschränken sich die bisherigen Methoden im Bereich der „Explainable AI“ häufig auf die Visualisierung bestimmter Bereiche eines neuronalen Netzes oder die Identifizierung besonders relevanter Eingabedaten, vgl. *Samek/Müller* (Fn. 15), S. 16 f.

84 *Guckelberger*, Verwaltung (Fn. 1), Rn. 523.

cher Weise.⁸⁵ Vielmehr ist sie in ihrer verfassungsrechtlichen Dimension als abwägungsbedürftiges Prinzip und nicht als unbedingte Regel zu verstehen. Das veranschaulichen die in § 39 Abs. 2 VwVfG normierten, verfassungskonformen Ausnahmen.⁸⁶ Insofern ist im Hinblick auf die jeweilige Situation zu prüfen, ob möglicherweise überwiegende Gründe das einzelfallbezogene oder bereichsspezifische Absehen von einer Begründung rechtfertigen können.⁸⁷ Zu denken ist beispielsweise an den Einsatz von Algorithmen im Rahmen eines Risikomanagementsystems im Besteuerungsverfahren (vgl. § 88 Abs. 5 AO). Die Geheimhaltung dürfte hier zwingend erforderlich sein, damit das Risikomanagementsystem seine Aufgabe, verdächtige Fälle herauszufiltern, überhaupt erfolgreich ausüben kann. Es dürfte deshalb auch das Informationsinteresse der Bürger:innen überwiegen. Auch beim Einsatz von Algorithmen zur intelligenten Steuerung von Ampeln dürfte die Angabe einer Begründung ohne größere Probleme entbehrlich sein. Das entspricht auch der bisherigen Praxis. Allerdings reicht allein der Hinweis auf die möglichen Effizienzvorteile des Algorithmen-einsatzes nicht aus, sondern es müssen auch die möglichen Schutzinteressen der Bürger:innen einzelfallbezogen berücksichtigt werden.⁸⁸ Der möglicherweise etwas weit formulierte § 39 Abs. 2 Nr. 3 VwVfG ist nach diesen Maßgaben auszulegen.⁸⁹

c) Zwischenergebnis

Im Hinblick auf die Begründungspflicht⁹⁰ ergibt sich, dass diese nicht die technische Transparenz der eingesetzten Algorithmen einfordert, sondern dass die Entscheidung rechtfertigenden rechtlichen und tatsächlichen Gründe auch bei einer durch oder mit Hilfe von Algorithmen getroffenen

85 Dazu *Kischel*, Begründung (Fn. 71), S. 63 f. u. 160 ff. mit näherer Erläuterung; *Schmidt-Aßmann* in: Dürig/Herzog/Scholz (Hrsg.), GG, 95. Ed. (Stand: Juli 2021), Art. 19 Abs. 4 Rn. 253 f.

86 *Stelkens* (Fn. 25), § 39 Rn. 4.

87 Vgl. *Kischel*, Begründung (Fn. 71), S. 176 ff. mit vielfältigen Beispielen für mögliche Ausnahmen.

88 Für eine sorgfältige Prüfung auch *Guckelberger*, Verwaltung (Fn. 1), Rn. 525; *Hufen/Siegel*, Fehler im Verwaltungsverfahren, 7. Aufl., 2021, Rn. 480 f.

89 Dazu *Guckelberger*, Verwaltung (Fn. 1), Rn. 524 ff.

90 Eine weitere Frage ist es hingegen, inwiefern möglicherweise Informationszugangsrechte eine Offenlegung verlangen, vgl. beispielsweise *Guckelberger*, Verwaltung (Fn. 1), Rn. 494 ff.

Entscheidung grundsätzlich angegeben werden müssen. Lässt sich eine solche Begründung aus technischen Gründen nicht verwirklichen, kann geprüft werden, ob überwiegende Gründe eine Ausnahme rechtfertigen. Andernfalls ist der Algorithmeneinsatz ausgeschlossen.

IV. Fazit

Im Ergebnis ist festzuhalten, dass Defizite im Bereich der Transparenz dem staatlichen Einsatz von Algorithmen nicht per se entgegenstehen. So wird man einzelfallbezogen auf die Interpretier- und Erklärbarkeit eines Algorithmenmodells verzichten können. Auch in einem demokratischen Rechtsstaat kommt der Transparenz kein absoluter Wert zu.⁹¹ Letztlich kommt es bei der Entscheidung über den Einsatz von Algorithmen auf eine Gesamtabwägung an, bei der die Transparenz nur als ein mögliches Kriterium zu berücksichtigen ist. Vielfach dürfte es technisch erfolgsversprechender sein, auf alternative Maßnahmen abzustellen, um den verlässlichen, sicheren und rechtskonformen Einsatz von Algorithmen zu gewährleisten. In dieser Hinsicht bieten sich insbesondere das „Blackbox-Testing“ in unterschiedlichen Varianten⁹² sowie das Prinzip der Redundanz an.

Die im Rahmen des Beitrags entwickelte Argumentation war primär am Vergleich mit menschlichen Amtswalter:innen und an technischen Lösungsmöglichkeiten jenseits der Transparenz orientiert. Als wichtige Ergänzung dieses Gedankengangs ist darauf hinzuweisen, dass es genauso möglich und sinnvoll ist, das bestehende Rechtssystem fortzuentwickeln⁹³ und an den Einsatz von intransparenten Algorithmen anzupassen. Denk-

91 In diese Richtung aber das VerfGH Saarland, Urt. v. 5.7.2019 – Lv 7/17, NJW 2019, 2456 (Rn. 1 ff.). Allerdings ist als wesentlicher Unterschied zu berücksichtigen, dass in der streitgegenständlichen Konstellation die Messdaten technisch hätten gespeichert werden können (Rn. 53, 72).

92 Vielversprechend dürfte beispielsweise der Ansatz *Hermstrüwers* (Fn. 1), 505 ff. sein, maschinell implementierbare Fairnessprinzipien herauszuarbeiten, die zur Kontrolle von Algorithmen eingesetzt werden können, ohne auf deren Interpretierbarkeit angewiesen zu sein.

93 In diese Richtung der Ansatz *Martinis* (Fn. 57), Rn. 89 ff. und *Rademachers*, Künstliche Intelligenz und neue Verantwortungsarchitektur, in: Eifert (Hrsg.), Digitale Disruption und Recht, 2020, S. 45, dessen Ausführungen sich aber nicht auf den staatlichen Einsatz beschränken. Grundzüge einer neuen Digitalordnung arbeitet *Barczak* (Fn. 7), 1004 heraus.

bar wäre es beispielsweise, eine neue niedrigschwellige Kontrollmöglichkeit für von Algorithmen getroffene Entscheidungen einzurichten.⁹⁴ Das würde es ermöglichen, Algorithmen tendenziell großzügiger einzusetzen, solange die Entscheidung unproblematisch rückgängig gemacht werden kann. Letztlich sind Lösungen für den gewinnbringenden hoheitlichen Einsatz von Algorithmen sowohl aus der rechtlichen als auch der technischen Perspektive zu entwickeln.

⁹⁴ In diese Richtung beispielsweise *Martini/Nink*, NVwZ-Extra 10/2017, 1 (4); *Guckelberger*, Verwaltung (Fn. 1), Rn. 616.