

Algorithmische „Entscheidungen“ in der Medizin? Eine Reflexion zu einem handlungsbezogenen Ausdruck

Susanne Hahn

I. Vorklärung: Handlungsbezogene Ausdrücke in unterschiedlichen Kontexten

„Entscheidung“, „entscheiden“ und benachbarte Vokabeln gehören zu einem begrifflichen Feld, das mit dem *Handeln* zu tun hat. Wenn davon die Rede ist, dass „Algorithmen entscheiden“ oder „künstliche Intelligenz bzw. Maschinen entscheiden“, dann werden handlungsbezogene Ausdrücke, die üblicherweise in Kontexten menschlichen Handelns Verwendung finden, auf einen anderen Bereich, nämlich Algorithmen oder Maschinen, übertragen. Eine solche Übertragung findet derzeit nicht nur mit der Entscheidungsrede statt. Die Rede von „autonomen Akteuren“ in Bezug auf künstliche Intelligenz, vom „Vertrauen in künstliche Intelligenz“ oder Zusammenfügungen wie „Algorithmenethik“ oder „Mensch-Maschine-Interaktion“ illustrieren diesen Umstand.

Diese und weitere Fälle geben Anlass zu Fragen: Welche Präsuppositionen, Konnotationen und Bedeutungsnetze werden übertragen, wenn ein Ausdruck, der in menschlichen Handlungszusammenhängen verwendet wird, auf den Bereich der künstlichen Intelligenz¹ übertragen wird? Welche Irreführungen ergeben sich möglicherweise daraus?

Mit dem vorliegenden Beitrag soll – eher beispielhaft und als Anfang einer noch anzustellenden umfassenderen Betrachtung – erprobt werden, wie eine Analyse der Begrifflichkeiten im üblichen menschlichen Handlungskontext ausfällt, welche Präsuppositionen ein Gebrauch im Bereich von KI enthält und welche Verwendungen im KI-Bereich plausibel sind. Dies soll am Beispiel der Entscheidungsrede dargelegt werden. Eine solche begriffliche Analyse ist auch als Wegbereitung und Grundlage für eine konstruktiv ausgerichtete Position zur künstlichen Intelligenz zu sehen:

1 Mit „künstlicher Intelligenz“ ist hier und im Folgenden stets das derzeit realisierbare und auch realisierte maschinelle Lernen, d.h. die Mustererkennung durch Algorithmen und ihre Nutzung als Voraussagemodelle gemeint.

Statt einzelne Anwendungen in ihren potenziellen Auswirkungen zum Gegenstand von Kritik zu machen, ließe sich positiv gewendet fragen: Welche Ziele lassen sich sinnvoll mit dem Einsatz maschinellen Lernens verfolgen und welche eher nicht? Welche Beschränkungen sind zu beachten?

Kandidaten für algorithmische „Entscheidungen“, die Anlass zur Analyse bieten, könnten beispielsweise der Einsatz von Modellen maschinellen Lernens bei der Kreditvergabe sein, aber auch bei der Einladung zu Folgeuntersuchungen beim Brustkrebsscreening oder bei der Aufforderung zur Durchführung einer Darmspiegelung nach Einreichen eines Stuhltests. In diesen Fällen würde „automatisch“ eine Maßnahme an ein algorithmisch ermitteltes Ergebnis geknüpft.² Die (Nicht-)Vergabe eines Kredits kann für Menschen durchaus schwerwiegende Konsequenzen haben. In medizinischen Kontexten ist die (Nicht-)Einladung zu weiteren Untersuchungen folgenreich. Insbesondere Falsch-Negativ-Ergebnisse – jemand wird nicht als untersuchungsbedürftig eingeschätzt, nicht eingeladen und entwickelt im weiteren Verlauf z.B. eine Krebserkrankung –, aber in geringerem Maße auch Falsch-Positiv-Ergebnisse – jemand wird fälschlicherweise zu weiteren Untersuchungen eingeladen, gerät aufgrund langer Wartezeiten für Untersuchungstermine unter psychischen Druck und wird womöglich durch weitere Diagnostik geschädigt –, haben für die Betroffenen erhebliche Konsequenzen.³

Die (potentiellen) Einsätze haben entsprechende Kritik auf den Plan gerufen. Die Beschreibungen des Wirkens von Algorithmen enthalten eben jene erwähnten handlungsbezogenen Ausdrücke, darunter prominent das Vokabular des Entscheidens. Ein Zitat aus dem für Laien sehr hilfreichen Sachbuch von Hannah Fry zur Künstlichen Intelligenz illustriert dieses Vorgehen:

-
- 2 Hier und an weiteren Stellen (s. Abschnitt II) lassen sich bereits im Fall menschlichen Handelns Fragen zur Abgrenzung von Überlegung, Entscheidung und Handlung aufwerfen. Vgl. zu diesem Aspekt und weiteren Entscheidungen S. McCall, *Decision*, *Canadian Journal of Philosophy*, 1987, 17, 261 (261ff.).
 - 3 Weitere, prominent gewordene Beispiele sind der Einsatz des COMPAS-Algorithmus in einigen Bundesstaaten der USA bei der Frage, ob potenzielle Straftäter bis zur Hauptverhandlung auf freien Fuß gesetzt werden oder – deutlich weniger brisant, aber vermutlich häufiger vorkommend – der Einsatz von Algorithmen bei der gezielten Werbung. Vgl. H. Fry, *Hello World. How to Be Human in the Age of the Machine*, London 2018, C. Rudin, *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead*, *Nature Machine Intelligence* 1(5), 2019, 206.

„Besides, all the hype over AI is a distraction from much more pressing concerns and – I think – much more interesting stories. Forget about omnipotent artificially intelligent machines for a moment and turn your thought from the far distant future to the here and now – because there are already **algorithms** with free rein to act as **autonomous decision-makers**. To **decide** prison terms, treatments for cancer patients and what to do in a car crash. They're already making life-changing **choices** on our behalf at every turn.

The question is, if we're handing over all that **power** – are they deserving of our **trust**?“⁴

Algorithmen werden als **autonome Entscheider** angesprochen, die **entscheiden**, ob jemand inhaftiert wird, ob Personen bestimmte Krebsbehandlungen bekommen etc., die somit über existentielle Zustände entscheiden. Ihnen wird **Macht** und **Handlungsbefugnis** zugeschrieben und es wird die Frage aufgeworfen, ob sie unser **Vertrauen** verdienen.

Im Folgenden sollen zunächst Verständnisse von Entscheidung dargestellt werden, um darauf aufbauend zu fragen, ob und wenn ja in welchem Sinn, Algorithmen entscheiden.

II. Was ist eine Entscheidung?

Um diese Frage umfassend zu beantworten, wäre eine umfangreichere Bestandsaufnahme angezeigt. Für die hier verfolgten Zwecke – Anregung zum Nachdenken darüber, wozu sich Algorithmen (nicht) einsetzen lassen und Sensibilisierung für eine Übertragung handlungsbezogenen Vokabulars auf informationstechnische Kontexte – wird eine sehr reduzierte Skizze erstellt. Dazu erfolgt zunächst ein Rückgriff auf ein klassisches begriffshistorisches Nachschlagewerk der Philosophie. Eine dort gelegte begriffliche „Spur“ wird weiterverfolgt, um das vorgeschlagene Verständnis von Entscheidung in eine Standardeinteilung der Entscheidungstheorie einzuordnen. In diesem Rahmen wird wiederum der Einsatz von Algorithmen betrachtet.

Der Artikel „Entscheidung“ des „Historische[n] Wörterbuch[s] der Philosophie“ liefert die folgende allgemeine Charakterisierung einer Entscheidung: „Der Begriff <E.> wird am häufigsten im Zusammenhang von rechtlichen und ethisch-politischen Fragen gebraucht und meint dort insgesamt,

4 H. Fry, Hello World (Fn. 3), S. 15. – Hervorhebungen S.H.

daß *etwas Ungewisses, Zweifelhafte zur Klärung kommt* [...].⁵ Dieses Verständnis, das die Nicht-Eindeutigkeit des Entscheidungsausgangs und die Determinierung des weiteren Verlaufs eines Geschehens durch die Entscheidung als spezifische Merkmale hervorhebt,⁶ findet weitere Illustration in einem Aufsatz von Hermann Lübbe: „Die Entscheidung überspringt einen Mangel an rationalen Bestimmungsgründen des Handelns. Sie ist deswegen nicht irrational. Die Vernunft der Entscheidungssituation besteht gerade darin, sich zum Handeln zu bestimmen, obwohl ausreichende Gründe, so und nicht anders zu handeln, fehlen.“⁷ Als Beispiel erwähnt Lübbe die Entscheidung über eine Invasion angesichts einer nicht gesicherten Wetterlage: „Die Antwort der Meteorologen hat keinen anderen Gewißheitsgrad, als ihre Wissenschaft ihn nach aller Erfahrung erreichen kann. Dennoch muß die Entscheidung heute fallen, und sie fällt als Entscheidung, sofern sie einen Abgrund von Ungewißheit überspringt.“⁸

Insgesamt lässt sich damit ein Verständnis von Entscheidung skizzieren, das die folgenden Merkmale aufweist: ein Akteur hat mehrere Handlungsoptionen, die Entscheidungssituation zeichnet sich durch einen gewissen Druck aus, es gibt einen Zwang zum Handeln, es ist nicht eindeutig, welche Konsequenzen jede Handlungsoption hat und es gibt keine zwingenden oder ausreichenden Gründe für eine Handlungsoption. – Diese Bedeutung von Entscheidung lässt sich durchaus als eine „emphatische“ bezeichnen, die wenig mit Entscheidungssituationen wie sie alltäglich an der Theke in der Bäckerei, im Restaurant oder auch bei der Wahl des Abendprogramms anfallen, zu tun hat. Zum emphatischen Verständnis von Entscheidung passen Vokabeln wie Entscheidungsbefugnis, Ermächtigung, autonome Akteure, diskretionäre Spielräume, Abwägung, Ziele, situative Notwendigkeiten, etc.

5 C. von Bormann, «Entscheidung», in: J. Ritter (Hrsg.): *Historisches Wörterbuch der Philosophie*, Basel 1972ff, DOI: 10.24894/HWPh.876 – Hervorhebung S.H.

6 Die erwähnte Ungewissheit und Zweifelhaftheit wird auf den Handlungsverlauf bezogen. Durch die Entscheidung wird die in ihrem Verlauf offene Handlungssituation geschlossen. Auf welche Faktoren sich die Offenheit bezieht – hier kommen die epistemische (Un-)Gewissheit bezüglich der Handlungskonsequenzen sowie die Bewertung derselben ins Spiel – wird, anders als in der weiter unten behandelten Entscheidungstheorie nicht thematisiert.

7 H. Lübbe, *Zur Theorie der Entscheidung*, in: *Collegium Philosophicum. Studien*. Joachim Ritter zum 60. Geburtstag, Basel/Stuttgart 1965, 118, (131f.)

8 H. Lübbe, *Entscheidung* (Fn. 7), 132.

Beispiele für passende Handlungsszenarien aus dem Bereich der Medizin sind: Jemand trifft für sich selbst eine weitreichende Entscheidung, z.B. eine Chemotherapie nicht durchführen zu lassen. Eine Ärztin entscheidet, einem bestimmten Patienten eine konkrete Therapie nicht zukommen zu lassen.⁹ Das oben angeführte Beispiel einer Kreditentscheidung lässt sich ebenfalls als Beispiel für eine solches Entscheidungsverständnis anführen.

In diesen Fällen sind *ex post* Fragen an die entscheidende Person denkbar wie: Warst Du befugt zur Entscheidung? Welchen Wissensstand hattest Du zum Zeitpunkt der Entscheidung? Welche Verantwortlichkeiten bzw. Pflichten hattest Du zum Zeitpunkt der Entscheidung? Welche Zwänge gab es zum Zeitpunkt der Entscheidung? War die Entscheidung wohlüberlegt?

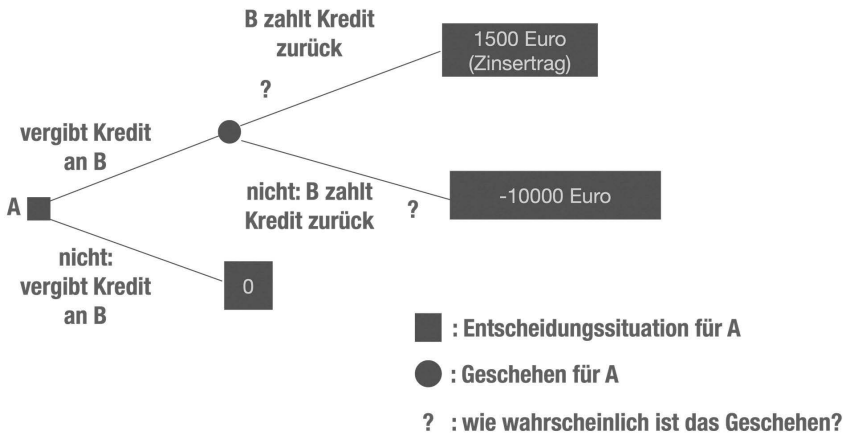
Die Theorie der Entscheidungen ist ein Gebiet, das sowohl in der Ökonomik als auch in einem Zweig der formal orientierten analytischen Philosophie verortet ist und die Rationalität von Entscheidungen in das Zentrum der deskriptiven und normativen Überlegungen stellt. Aus der Entscheidungstheorie ist eine Klassifikation von Entscheidungen geläufig, die auf den Wissensstand zum Zeitpunkt der Entscheidung abhebt. Es werden unterschieden: Entscheidungen unter Sicherheit, Entscheidungen unter Risiko und Entscheidungen unter Unsicherheit.¹⁰ Bei Entscheidungen unter Sicherheit ist bekannt, dass eine Handlung sicher zu einem bestimmten Resultat führt. Bei Entscheidungen unter Risiko führt jede Handlung zu einer Menge von Konsequenzen, bei denen jeweils die Wahrscheinlichkeiten, mit denen sie eintreten, bekannt sind. Dies unterscheidet sie von Entscheidungen unter Unsicherheit, bei denen sich zwar angeben lässt, welche Konsequenzen sich einstellen können, denen man jedoch keine Wahrscheinlichkeiten zuordnen kann.

9 Hier – wie an vielen anderen Stellen – ließen sich weitere Unterscheidungen anführen und Fragen aufwerfen. Genannt seien hier die Unterscheidung von Entscheidungen, die nur die entscheidende Person betreffen und solchen, die (auch mit entsprechender Befugnis) andere Personen betreffen. Weiter ließen sich Einzelentscheidungen und Entscheidungen über eine Klasse von Handlungssituationen unterscheiden. – Vgl. zu diesen und weiteren Aspekten aus philosophischer Sicht die immer noch sehr instruktive Analyse von *Nicholas Rescher*, *Risk. A Philosophical Introduction to the Theory of Risk Evaluation and Management*, Washington 1983. Zur Einordnung des Verhältnisses zwischen Entscheidung, Überlegung und Handlung vgl. *C. Budnik*, *Überlegen und Entscheiden*, in: *M. Kühler/M. Rütger* (Hrsg.), *Handbuch Handlungstheorie*, DOI 10.1007/978-3-476-05359-6_19.

10 In klarer Darstellung: *R.D. Luce/H. Raiffa*, *Games and Decisions. Introduction and Critical Survey*, New York 1957, chap. 2.

Das oben geschilderte „emphatische“ Verständnis von Entscheidung zeichnet sich insbesondere dadurch aus, dass „sie einen Abgrund von Ungewissheit überspringt“. Bezogen auf die Klassifikation von Entscheidungssituationen könnte man dieses Verständnis dem Typ „Entscheidung unter Unsicherheit“ zuordnen. Viele Bemühungen in den Wissenschaften lassen sich als Bestreben deuten, Entscheidungen unter Unsicherheit zu Entscheidungen unter Risiko zu transformieren. So versucht beispielsweise die klinische Forschung, Zusammenhänge zwischen einer medizinischen Maßnahme und dem Eintreten von Konsequenzen statistisch oder anhand klinischer Studien zu etablieren. Die Gewinnung von Wahrscheinlichkeiten ist ein zentrales Anliegen empirischer Forschung.

Anhand der folgenden Darstellung eines Standardbeispiels lassen sich die Unterscheidungen erläutern:



Die handelnde Person A kann sich zwischen zwei Handlungsoptionen entscheiden, nämlich den Kredit an B zu vergeben oder dies nicht zu tun. Im letzteren Fall ist die Konsequenz für die handelnde Person klar, sie hat weder einen Zinsertrag noch einen Verlust. Im anderen Fall ist das Handeln von B für A ein Geschehnis, d.h. ein Verlauf der Welt, der nicht mehr der eigenen Kontrolle unterliegt. Zwei mögliche Verläufe werden angenommen, die komplette Rückzahlung, einschließlich der Zinsen, und die vollständige Unterlassung der Rückzahlung. Die Einschätzung der Konsequenzen für A hängt von zwei Faktoren ab, der Einschätzung bzw. Angabe, wie wahrscheinlich der Kreditausfall ist und der Bewertung der möglichen Konse-

quenzen. Je nachdem als wie gravierend der Kreditausfall gewertet wird und wie stark gewünscht der Zinsertrag ist, wird die Bewertung ausfallen. Im Falle einer Risikoentscheidung, d.h. dann, wenn Wahrscheinlichkeiten zum Kreditausfall bekannt sind, berechnen Theorien rationalen Handelns einen Erwartungswert der Handlungsoptionen. Dazu werden die bewerteten möglichen Konsequenzen mit den Wahrscheinlichkeiten gewichtet und aufsummiert. Rational handeln Personen, wenn sie ihren Erwartungswert maximieren.¹¹

Dieses Vorgehen lässt sich als Ausweg aus der Situation deuten, die oben geschildert wurde: Man möchte Entscheidungssituationen vermeiden, in denen es keine hinreichenden Gründe für die eine oder andere Handlungsoption gibt. Daher versucht man, die epistemischen Unsicherheiten durch statistische und/oder experimentelle Forschung zu überwinden, um Wahrscheinlichkeiten für das Eintreten von Konsequenzen zu gewinnen. – Daneben kann ein Kern der Rationalwahltheorien in dem Bemühen gesehen werden, die Bewertungskomponente ebenfalls methodisch zu behandeln, indem den Präferenzen der Akteure bestimmte Konsistenzbedingungen auferlegt werden. – Wenn man das Kreditszenario durch eines aus dem medizinischen Kontext ersetzt, wird noch deutlicher, dass auch der Umgang mit dem Bewertungselement der Entscheidung mit gravierenden spezifischen Schwierigkeiten einhergeht. In diesen Fällen sind nicht lediglich monetäre Konsequenzen zu betrachten, sondern Zustände, die das Wohlbefinden von Menschen möglicherweise tiefgreifend beeinflussen und die zudem von unterschiedlichen Individuen unterschiedlich bewertet werden.

Insgesamt spricht einiges dafür, Entscheidungen im obigen „emphatischen“ Sinn als Entscheidungen unter Unsicherheit in der Klassifikation der Entscheidungstheorie anzusprechen. Für Entscheidungen unter Unsicherheit sehen die Rationalwahltheorien unterschiedliche, durchaus miteinander konfligierende Empfehlungen vor: Die Maximin-Strategie empfiehlt die Handlungsoptionen jeweils hinsichtlich ihres schlechtestmöglichen Verlaufs zu betrachten und dann diejenige zu wählen, die im Vergleich der schlechtest möglichen Konsequenzen am besten abschneidet. Die optimistische Maximax-Strategie zieht auf das bestmögliche Ergebnis ab. Die Theorie der Maximierung des subjektiven Erwartungsnutzens, als dritte

11 C. Bicchieri, Rationality and Game Theory, in: A.R. Mele/P. Rawling, The Oxford Handbook of Rationality, DOI: 10.1093/oxfordhb/9780195145397.003.0010, 183f.

Variante, entwickelt ein Konzept, um die fehlenden objektiven Wahrscheinlichkeiten durch subjektive Wahrscheinlichkeiten zu kompensieren.¹²

Ein Verständnis von Entscheidung, das den Entscheidungsakt als eine unvermeidliche Überwindung von Unsicherheit betrachtet, wird dieses Vorgehen anhand verschiedener Strategien vermutlich als pseudo-objektiv ablehnen und die Betrachtung des Einzelfalls und die Nutzung diskretionärer Spielräume und somit das individuelle Ermessen als notwendig betrachten. Dazu passt die Auffassung, dass es nicht die eine richtige Entscheidung gibt, sondern dass man in der jeweiligen Entscheidungssituation zu unterschiedlichen plausiblen Entscheidungen kommen kann, da die Bedingungen gerade nicht zu eindeutigen Schlüssen führen. So können Entscheidungen, von denen man ex post weiß, dass sie zu schlechten Konsequenzen geführt haben, im Nachhinein gerechtfertigt werden durch die Überlegungen in der jeweiligen Situation.

III. Entscheiden Algorithmen?

Algorithmen maschinellen Lernens – dasjenige, was derzeit als Künstliche Intelligenz verfügbar ist –, sind im Kern Verfahren der Mustererkennung. Aus vorhandenen Daten werden Muster extrahiert, wie z.B., dass Kreditnehmer, die jung, männlich und ohne Immobilienbesitz sind, ihren Kredit mit der-und-der Häufigkeit nicht zurückzahlen, und auf neue Fälle angewendet. Es handelt sich um eine statistische Nachbildung des Lernens aus Erfahrung.¹³ Der Prozess der Entwicklung und Anwendung von Algorithmen, die als „Entscheidungsalgorithmen“ angesprochen werden, lässt sich in fünf Schritten erfassen:

- Generierung des Daten-Inputs,
- Datenaufbereitung,
- Erstellung des Voraussagemo Modells,

12 C.G. Hempel, Rationales Handeln, in: G. Meggle, Analytische Handlungstheorie, Frankfurt a. M. 1985, 388 (392ff.).

13 Hier und an anderen Stellen bestehen Anknüpfungspunkte für wissenschaftsphilosophische Überlegungen zur Generierung von allgemeinen Zusammenhängen. Vgl. dazu allgemein z.B. G. Schurz, Einführung in die Wissenschaftstheorie, Darmstadt 2006, Kap. 3.6 – 3.9 und 4.

- Erstellung von Entscheidungsregeln und
- Generierung des Outputs.¹⁴

Die ersten beiden Schritte sind nicht spezifisch für die Algorithmen maschinellen Lernens, sondern gehören standardmäßig zu statistischen Verfahren.¹⁵ Im dritten Schritt wird die Erstellung eines Voraussagemodells vollzogen, z.B. durch künstliche neuronale Netze oder – eher klassisch – durch logistische Regression.¹⁶ Mit einem solchen Voraussagemodell können neue Fälle, z.B. eine neue Anfrage für einen Kredit oder eine Patientin, die Kandidatin für eine Untersuchung oder Therapie ist, eingeschätzt werden. Resultat der Einschätzung ist ein Wahrscheinlichkeitswert für eine bestimmte Konsequenz, z.B. die vollständige Rückzahlung eines Kredits oder die Zugehörigkeit zu einer gefährdeten Patientengruppe. Bis zu diesem Schritt ist „noch nichts passiert“, d.h. es wurden lediglich Prognosen erstellt oder Klassifikationsaussagen getroffen.

Erst mit der Aufstellung besonderer Wenn-Dann-Verknüpfungen werden Maßnahmen eingeleitet (in der obigen Aufstellung von Finlay wird dieser Schritt als „Erstellung von Entscheidungsregeln“ bezeichnet): *Wenn* eine Person einen Score von so-und-so-viel hat, d.h. wenn die Wahrscheinlichkeit für die Kreditrückzahlung im Beispiel der Kreditvergabe oder die Wahrscheinlichkeit, von einer Untersuchung oder einer Therapie zu profitieren, wenigstens ein bestimmtes Maß hat, *dann* wird eine Maßnahme durchgeführt. In den Beispielen wäre der Output eine Finanzierungszusage bzw. eine Einladung zu einer Untersuchung oder ein Schreiben für die Vergabe einer Therapie.

Die Wenn-Dann-Verknüpfungen sind keine Regeln im Sinne handlungsanleitender Regeln für menschliche Akteure. Handlungsregeln unterscheiden sich von Wenn-dann-Verknüpfungen der Algorithmen dadurch, dass sie deontische Anteile enthalten, die sich an die Regeladressaten richten:

-
- 14 S. Finlay, *Artificial Intelligence and Machine Learning for Business: A No-Nonsense Guide to Data Driven Technologies*, Preston 2017, chap. 2.
- 15 B. Baesens, *Analytics in a Big Data World: The Essential Guide to Data Science and its Applications*. Minneapolis 2014, chapt. 2.
- 16 Nicht alle Einsätze erfordern künstliche neuronale Netze oder ähnliche Verfahren, die wegen ihrer Nicht-Nachvollziehbarkeit besonderer Kritik ausgesetzt sind. Für die Einschätzung der Kreditwürdigkeit ist beispielsweise eine klassische logistische Regression, umgesetzt in ein Verfahren maschinellen Lernens, ausreichend. Vgl. Finlay, *Artificial Intelligence* (Fußn. 14), chapt. 7.

Akteuren sind Handlungen *geboten*, *verboten* oder *erlaubt*.¹⁷ Klassische Beispiele für soziale Handlungsregeln sind rechtliche oder moralische allgemeine Normen. Für diese ist charakteristisch, dass im Fall des Verstoßes gegen eine Verbots- oder Gebotsregel Sanktionen vorgesehen sind. Wer gegen das Diebstahlverbot agiert und ertappt wird, muss mit einer Sanktion rechnen. Wer einer Lüge überführt wird, muss mit Missbilligung oder Ächtung rechnen. Wenn es um soziale Regeln geht, ist die Möglichkeit der Zuwiderhandlung von vornherein mitbedacht. Menschen können sich gegen Handlungsverbote oder -gebote entscheiden. Sie sind zudem in der Lage, Handlungsermächtigungen (eine bestimmte Form von Erlaubnissen) zu nutzen, um in einer Situation entsprechend ihrer Beurteilung zu agieren.

All dies ist bei den Wenn-dann-Verknüpfungen von Algorithmen nicht der Fall. Hier ist die Maßnahme an das Erfüllen einer Voraussetzung, letztlich an einen bestimmten Wahrscheinlichkeitswert, geknüpft. Auch wenn ein Bereich von Wahrscheinlichkeiten bestimmt würde, innerhalb dessen die Durchführung der Maßnahme an einen Zufallsgenerator gebunden würde, unterscheidet sich dieses Vorgehen von einer Erwägung wie sie beispielsweise bei ermächtigten Personen vorkommt. Bei der Einräumung und Nutzung diskretionärer Spielräume wird man von den Personen erwarten, dass sie auch *ex post* Gründe für ihre Entscheidung vorbringen können, und sich nicht auf Zufall oder Beliebigkeit berufen.

Was bedeutet das für die Frage, ob Algorithmen entscheiden? Wenn man ein emphatisches Verständnis von Entscheidung unterstellt, wie es oben dargestellt wurde, dann passen dazu Konzepte wie „Entscheidungsmacht“, „autonome Akteure“, „Handlungsbefugnis“, „diskretionäre Spielräume“, „Ziele“ usw. Die einzelnen Sortierungsfälle, die durch Algorithmen vorgenommen werden, sind *keine Entscheidungen in diesem Sinne*. „Der Algorithmus“ ist kein autonomer Akteur, der unter Ausnutzung diskretionärer Spielräume Entscheidungen trifft und dem man auch prima facie „unsinnige“ Entscheidungen nachsehen würde, wenn er sie situativ erklärt. Übertragen in eine „mechanische“ Maschine: Wenn man nur noch Äpfel einer bestimmten Größe verkaufen will, dann kann man zur Sortierung größerer Apfelmengen ein Förderband konstruieren, bei dem nur Äpfel

17 Zu Regeln (hier synonym zu allgemeinen Normen) vgl. N. Hoerster, *Ethik und Interesse*, Stuttgart 2003, Kap.2; G. Siegart, *Regel*, in: P. Kolmer/G. Wildfeuer (Hrsg.), *Neues Handbuch philosophischer Grundbegriffe*, Freiburg 2011, 1864; S. Hahn, *Regeln und Normen*, in: L. Heidbrink/A. Lorch/V. Rauen (Hrsg.), *Handbuch Wirtschaftsphilosophie III, Praktische Wirtschaftsphilosophie*, https://doi.org/10.1007/978-3-658-22107-2_9.

ab einer bestimmten Größe zur Verpackungsstation weitergeleitet werden. Äpfel, die zu klein sind, fallen durch eine Öffnung hindurch. Die Maschine sortiert kleine Äpfel heraus, im geschilderten emphatischen Sinn entscheidet sie nicht. Analog kann man sagen, dass der Algorithmus Fälle sortiert.

Ausgangspunkt für die vorgebrachten Überlegungen war ein gegenteiliges Verständnis von Algorithmen, das diese als Träger von Autonomie und Macht sieht, somit ein emphatisches Verständnis von Entscheidung unterstellt und dementsprechend die Frage aufwirft: „The question is, if we're handing over all that **power** - are they deserving of our **trust**?¹⁸ – Dieses Verständnis – diese These soll hier vertreten werden – verstellt den Blick auf den Umstand, dass mit der *Herstellung* einer Wenn-Dann-Verknüpfung eine Entscheidung gefällt wird, die die Sortierung der zukünftigen Fälle festlegt. *Ex ante* wird über die Annahme eines Voraussagemodells und – insofern ein output automatisiert verknüpft wird – über diese Verknüpfung entschieden.¹⁹ Hier handelt es sich um eine Entscheidung im emphatischen Sinn: Wenn man einen Wahrscheinlichkeitswert festlegt, unterhalb dessen keine Kreditvergabe erfolgt oder keine Einladung zu einer Untersuchung, dann weiß man nicht, welche Einzelfälle, die man bei einer Einzelfallbetrachtung vielleicht noch gerade zu den positiv beschiedenen Fällen gezählt hätte, der Grenzziehung zum Opfer fallen; bzw. wer noch zu den positiv beschiedenen Fällen zählt, aber besser in die andere Gruppe gefallen wäre. Die Situation ähnelt derjenigen, in der über die Empfindlichkeit eines Testverfahrens in der Medizin entschieden wird. Hier sind stets die in Kauf genommenen falsch-negativen Ergebnisse gegen die falsch-positiven Ergebnisse abzuwägen. Typischerweise lassen sich hier für verschiedene Entscheidungen, d.h. verschiedene Verhältnisse dieser Ergebnisse, Gründe angeben, das Verhältnis ist nicht vorgegeben.

Der Einsatz eines Algorithmus mit einer Maßnahmenverknüpfung stellt somit eine Entscheidung dar, für die die entscheidende(n) Person(en) auch verantwortlich ist (sind). Eine Sicht auf Algorithmen als Sortierungsverfahren, über deren Einsatz *ex ante* entschieden wird, ermöglicht somit eine differenzierte Sicht auf die Zuschreibung von Verantwortung. Die Vorentscheidung über den Einsatz eines Sortierungsalgorithmus bedeutet

18 Fry, Hello World (Fn. 3), 15.

19 Hier ließe sich einwenden, dass sich die Sachlage durch sich verändernde Algorithmen ändert, wie dies bei nicht-überwachtem maschinellen Lernen der Fall ist. Allerdings sind in diesen Fällen Zielvorgaben zu machen, und der Ort der Entscheidung könnte wiederum in dieser Festlegung gesehen werden.

nicht, dass damit auch keine Verantwortungszuschreibung mehr möglich ist. Allerdings liegt diese nicht mehr in der jeweiligen Einsortierung des Einzelfalls, sondern in der Entscheidung über ganze Klassen von Fällen. Werden Algorithmen als Sortierungsmechanismen eingesetzt, werden Einzelfallentscheidungen durch ex ante-Entscheidungen über Klassen von Fällen ersetzt. – Zwei Vergleichsbeispiele sollen die Besonderheit dieser Entscheidung über künftige Sortierungen verdeutlichen: In der sogenannten evidenzbasierten Medizin sollen Leitlinien der Fachgesellschaften den besten medizinischen Kenntnisstand in die vielen einzelnen Therapieentscheidungen bringen. Anders als bei den Sortierungsalgorithmen mit ihren Wenn-Dann-Verknüpfungen handelt es sich jedoch in diesen Fällen wiederum um allgemeine Handlungsnormen, denen (s.o.) zuwidergehandelt werden kann und die insgesamt unter dem Vorbehalt der individuellen Prüfung der behandelnden Ärztinnen stehen.²⁰ Bei der Kreditvergabe wird ebenfalls versucht, die Parameter, die sich in der Vergangenheit als relevant für die Rückzahlung ergeben haben, in die Formulierung von Richtlinien zur künftigen Vergabe fließen zu lassen. Die Personen, die diese Richtlinien umsetzen, könnten als analoge Algorithmen zur Sortierung von Kreditanträgen betrachtet werden. Insofern jedoch Ermessensspielräume bestehen und Sachbearbeiterinnen von den Vorgaben abweichen können, handelt es sich nicht um bloße, vorgegebene Sortierungen.

Die eingangs zum Ausdruck gebrachte Anregung, in konstruktiver Weise über den Einsatz von Algorithmen nachzudenken, kann auf diesem Hintergrund mit einigen Fragen vorangebracht werden: Welche Entscheidungen eignen sich dazu, von Einzelfallentscheidungen mit Ermessensspielraum zu Sortierfällen aufgrund vorgängiger Entscheidungen über Muster-Maßnahme-Verknüpfungen transformiert zu werden? Welche Qualitätsanforderungen sind an die Entscheidung über diese Einsetzung von Voraussagemodellen und die Wenn-dann-Verknüpfung zu stellen?²¹ Wie ist das Problem der Bewertung von Konsequenzen zu behandeln? Welche Entscheidungen sollen möglicherweise von vornherein als Kandidaten für Sortierungsverfahren

20 „Leitlinien sind als „Handlungs- und Entscheidungskorridore“ zu verstehen, von denen in begründeten Fällen abgewichen werden kann oder sogar muss.“ AWFM-Regelwerk Leitlinien, Arbeitsgemeinschaft der wissenschaftlichen medizinischen Fachgesellschaften, <https://www.awmf.org/regelwerk/>; Abrufdatum 28.09.2023.

21 Die Möglichkeit, Korridore von Wahrscheinlichkeitswerten einzurichten, in denen Fälle von der Sortierung ausgenommen, um einer menschlichen Einzelfallbeurteilung zugeführt zu werden, könnte das Spektrum der in Frage kommenden Entscheidungen erweitern.

ren ausgeschlossen werden? Welche institutionellen Arrangements könnten einen fruchtbaren Einsatz von Algorithmen begleiten?

Die Analyseskizze zur Verwendung der Entscheidungsrede sollte zeigen, dass eine unreflektierte Übertragung auf Algorithmen ein emphatisches Verständnis von Entscheidung präsupponiert, das nicht der tatsächlichen Arbeitsweise von Algorithmen als nachgeordneten Sortierungsmechanismen entspricht. Damit wird zudem der Blick auf den Umstand erschwert, dass die Entscheidung über den Einsatz eines Algorithmus eine folgenreiche Entscheidung darstellt, die mit Verantwortlichkeiten einhergeht. Die – bestenfalls von informatischer Expertise begleitete – Sicht auf die Arbeitsweise von Algorithmen und ihre Leistungsfähigkeit und Limitationen einerseits und die Analyse unterschiedlicher Entscheidungstypen andererseits eröffnen den Raum für die Frage nach konstruktiven Einsätzen für algorithmische Sortierungen.²²

22 Insofern solche Überlegungen unter der Perspektive *de lege ferenda* erfolgen (angesichts bestehender Regulierung hieße das konkret unter einer *Änderungsperspektive*) ließe sich auch fragen, für welche Einsätze sich das Verbot komplett „automatisierter Entscheidungen“ durch eine Öffnung für stark routinisierte Vorgänge, flankiert durch ein institutionelles Arrangement leicht zu erhebender Einsprüche modifizieren ließe.

