

Verantwortungsvolle Maschinen ohne Verantwortlichkeit?

Datenintensive Algorithmen in Organisationen

Cristina Besio, Cornelia Fedtke, Michael Grothe-Hammer, Athanasios Karafillidis, Andrea Pronzini

Author accepted manuscript

Autoreninformation:

Prof. Dr. Cristina Besio

Helmut-Schmidt-Universität Hamburg, Institut für Gesellschaftswissenschaften, Holstenhofweg 85, 22043 Hamburg, Deutschland. E-Mail: cristina.besio@hsu-hh.de

Cornelia Fedtke, M.A.

Helmut-Schmidt-Universität Hamburg, Institut für Geisteswissenschaften, Holstenhofweg 85, 22043 Hamburg, Deutschland. E-Mail: fedtkec@hsu-hh.de

Assoc. Prof. Dr. Michael Grothe-Hammer

Norwegian University of Science and Technology, Department of Sociology and Political Science, Edvard Bulls veg 1, 7491 Trondheim, Norwegen. E-Mail: michael.grothe-hammer@ntnu.no

PD Dr. Athanasios Karafillidis

CVJM Hochschule, Evangelische Bank Institut für Ethisches Management, Hugo-Preuß-Str. 40, 34131 Kassel. E-Mail: atha@karafillidis.com

Andrea Pronzini, Dipl.-Soz.

Centro formazione di polizia, Via Ferriere 5, 6512 Giubiasco. E-mail: apronzini@sunrise.ch

Korrespondenzautorin:

Prof. Dr. Cristina Besio

Helmut-Schmidt-Universität Hamburg, Institut für Gesellschaftswissenschaften, Holstenhofweg 85, 22043 Hamburg, Deutschland. E-Mail: cristina.besio@hsu-hh.de

Zusammenfassung:

Die Frage der Verschiebung von Verantwortung, die durch den vermehrten Einsatz von datenintensiven Algorithmen verursacht wird, beschäftigt seit einigen Jahren die sozialwissenschaftliche Forschung. Dabei geht es vor allem darum, welche Personen oder Instanzen dafür verantwortlich sein sollen, wenn sich Entscheidungen als diskriminierend, sachlich falsch oder ungerecht herausstellen. Aus soziologischer Perspektive fällt auf, dass in diesen Debatten nicht trennscharf zwischen Verantwortung und Verantwortlichkeit unterschieden wird. In unserem Beitrag greifen wir diese Unterscheidung so wie sie von Niklas Luhmann (1964) formuliert wurde auf, um Veränderungen und Kontinuitäten des Organisierens zu analysieren, die mit dem Einsatz datenintensiver Algorithmen verbunden sind. Wir zeigen, dass Algorithmen in organisationalen Entscheidungsprozessen Unsicherheit absorbieren und somit durchaus Verantwortung übernehmen, aber nicht für Fehler verantwortlich gemacht werden können. Durch den Einsatz von Algorithmen wird Verantwortung in Assemblagen von Personen und Technik zerlegt, während die Zuschreibung von Verantwortlichkeit hochgradig kontrovers wird. Daraus entstehen neue Diskrepanzen zwischen Verantwortung und Verantwortlichkeit, die insbesondere für das organisationsinterne Vertrauen und die Innovationsfähigkeit von Organisationen folgenreich sein können.

Abstract:

The question of the shift of responsibility caused by the increased use of data-intensive algorithms has occupied social science research for several years. Existing research has been mainly concerned with discussing which persons or entities should be accountable when decisions turn out to be discriminatory, factually incorrect or unjust. From a sociological perspective, it is striking that in these debates no clear distinction is made between responsibility and accountability. In our paper, we take up this distinction as formulated by Niklas Luhmann (1964) to analyze changes and continuities of organizing associated with the use of data-intensive algorithms. We show that algorithms absorb uncertainty in organizational decision-making processes and thus can indeed assume responsibility but cannot be held accountable for mistakes. Through the use of algorithms, responsibility is decomposed into assemblages of people and technology, while the attribution of accountability becomes highly controversial. This gives rise to new discrepancies between responsibility and accountability, which can be particularly consequential for the internal trust and innovative capacity of organizations.

1 Einleitung

Die Verschiebung von Verantwortung, die durch den vermehrten Einsatz von datenintensiven Algorithmen verursacht wird, beschäftigt seit einigen Jahren die sozialwissenschaftliche Forschung zum Thema Digitalisierung (Mittelstadt et al. 2016). In den Debatten geht es häufig darum, welche Personen oder Instanzen dafür verantwortlich sein sollen, wenn sich sogenannte algorithmische Entscheidungen beispielsweise als diskriminierend, ungerecht oder sachlich falsch herausstellen. Sind immer noch Menschen Urheber aller Handlungen oder sind inzwischen auch „intelligente“ Maschinen fähig, eigenständig in die Kommunikation einzugreifen? Werden klassische Vorstellungen einer Zurechnung von Verantwortung für Entscheidungen obsolet, wenn die Urheberschaft für Entscheidungen nicht mehr eindeutig bei Menschen verortet werden kann (Matthias 2004)? Benötigen wir Modelle der verteilten Verantwortung (Loh 2019), um die Verantwortlichkeit von Entscheidungen als Produkte soziotechnischer „assemblages“ erfassen zu können (Annany/Crawford 2018)? Solche Fragen sind in den letzten Jahren über verschiedene Disziplinen hinweg diskutiert worden. Sie befassen sich unter anderem mit autonomen Fahrzeugen und Fahrstuhlsystemen, automatischer Verschlagwortung von Bildern oder „Killer Robots“ (Bonneton et al. 2016; ANONYMISIERT; Matthias 2004; Sparrow 2007).

Diese Diskussion findet auch in der Organisationsforschung¹ statt. Zahlreiche Arbeiten haben darauf hingewiesen, dass der zunehmende Gebrauch datenintensiver Algorithmen in Organisationen zu weitreichenden und teils fundamentalen Veränderungen führt (z. B. Büchner 2018; Constantiou/Kallinikos 2015; Faraj et al. 2018; Büchner/Dosdall 2021). Das betrifft Prozesse der Entscheidungsfindung, etablierte Machtgefüge und gewohnte Legitimationsmechanismen (z. B. Besio et al. 2022). In der Konsequenz stellt sich auch die klassische Frage nach der Verantwortung und Verantwortlichkeit für algorithmisierte Entscheidungen auf eine neue und brisante Weise.

Dabei fällt aus soziologischer Perspektive auf, dass zumeist nicht trennscharf zwischen Verantwortung und Verantwortlichkeit unterschieden wird. Zwar tauchen die englischsprachigen Äquivalente *responsibility* und *accountability* in der Literatur auf; eine Unterscheidung der Begriffe wird jedoch in entsprechenden Studien weder analytisch noch empirisch vorgenommen (siehe exemplarisch Mittelstadt et al. 2016). Wenn beide Bereiche doch unterschieden werden, wird Verantwortung auf individuelle moralische Verpflichtung reduziert

¹ Einige Konzepte für diesen Artikel wurden im Forschungsprojekt LEAD.Bw „Führungskulturen im digitalen Zeitalter – Der Fall der Bundeswehr“ im Forschungszentrum dtec.bw®, Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr, entwickelt.

(Meijer/Grimmelikhuijsen 2020). Aus unserer Sicht verschenkt diese Unschärfe allerdings das explanatorische Potenzial der Unterscheidung von Verantwortung und Verantwortlichkeit, das Niklas Luhmann (1964, 172-190) bereits in den 1960er Jahren hervorgehoben hat. Insbesondere führt eine Klärung beider Begriffe und ihrer Relation zu der Einsicht, dass Algorithmen in organisationalen Informationsverarbeitungsprozessen durchaus Verantwortung übernehmen, aber eben nicht verantwortlich gemacht werden können. Es handelt sich um verantwortungsvolle² Maschinen ohne Verantwortlichkeit.

Im Anschluss an die Vorstellung des Forschungsstandes zum Einsatz von Algorithmen in Organisationen und zur Verantwortung von Algorithmen (2) greifen wir die Unterscheidung von Verantwortung und Verantwortlichkeit auf, um Veränderungen und Kontinuitäten des Organisierens zu analysieren, die mit dem Einsatz datenintensiver Algorithmen verbunden sind. Wir verstehen Verantwortung nicht als moralische Verpflichtung, sondern als „sozialen Prozeß der Informationsverarbeitung [...], der zugleich der Absorption von Unsicherheit und der Bewußtseinsentlastung dient“ (Luhmann 1964, 174), und verwenden das Konzept der Verantwortlichkeit für die Rechenschaftspflicht bei Entscheidungen (3). Wir gehen davon aus, dass maschinelle Informationsverarbeitung tatsächlich *Verantwortung* übernehmen kann, sofern sie in alltäglichen Organisationsprozessen Unsicherheit absorbiert und Orientierung schafft. Dabei wird Verantwortung im Kontext von Assemblagen, bestehend aus Personen und Objekten, zerlegt (4.1). Algorithmen verunsichern zugleich die Organisation in Bezug auf Fragen der *Verantwortlichkeit*: Herkömmliche Formen der Rechenschaftspflicht werden infrage gestellt und neue entstehen. Verantwortlichkeit wird hochgradig kontrovers (4.2). Dies hat auf verschiedenen Ebenen Folgen. Wir weisen insbesondere auf Veränderungen im organisationalen Vertrauen und in der Innovationsfähigkeit von Organisationen hin (5). Unsere Ausführungen bestätigen die Auffassung, dass die Effekte der Digitalisierung nicht linear oder deterministisch verlaufen und deshalb jeweils die spezifischen Mensch-Technik-Konstellationen untersucht werden müssen, die die Praxis sowie die organisationalen Folgen von Algorithmen gestalten (Christin et al. 2015; Ananny/Crawford 2018; Zarsky 2016). Die Unterscheidung zwischen Verantwortung und Verantwortlichkeit bietet neue Gesichtspunkte, um diese Effekte zu reflektieren (6).

2 Algorithmen in Organisationen und die Frage der Verantwortung

Algorithmen sind Entscheidungsregeln, die schematisch angewandt werden, um zielgerichtet eine bestimmte, bekannte Problematik zu lösen. Grundsätzlich kann zwischen lernenden und

² Der Begriff verantwortungsvoll wird hier in der Bedeutung „mit Verantwortung verbunden“ benutzt und nicht im Sinne von gewissenhaft oder pflichtbewusst.

nichtlernenden Algorithmen unterschieden werden. Nichtlernende Algorithmen sind einfache, vorgegebene Regeln und derart programmiert, dass aus einem bestimmten Input immer der gleiche Output hervorgeht. Solche automatisierten Prozesse der Informationsverarbeitung sind typisch für die soziale Form der Organisation. Demnach sind Wenn-dann-Algorithmen weder ein neues noch ein rein technisches Phänomen.³ Die Bedeutung dieser einfachen Algorithmen hat in den letzten Jahren allerdings stark zugenommen, weil neue technische Systeme ihren Einsatz deutlich erleichtert haben und teils erfordern (Heiland 2018). Der zunehmende Einsatz von Algorithmen bringt sogar neue kommunikative Phänomene hervor. Ein prominentes Beispiel ist der Hochfrequenzhandel in der Finanzbranche, der weitgehend auf sehr einfachen Algorithmen basiert – deren Einsatz aber im heute bekannten Ausmaß erst durch schnelle Datenübertragungs- und Verarbeitungsraten möglich geworden ist und zuvor ungeahnte Möglichkeiten hervorbringt (Schwartz 2015; siehe Jöstingmeier in diesem Band).

Lernende Algorithmen hingegen sind komplexer. Anhand empirischen Datenmaterials entwickeln sie eigene Regeln, das heißt, sie können die Ausgangsregeln, mit denen sie gestartet sind, verändern und an eine neue, veränderte Datenlage anpassen. In einem Modell werden dazu Merkmale aus Beobachtungs- oder Trainingsdaten ausgewertet und mit einer Zielvariable verknüpft (z. B. kann in Onlineshops Klickverhalten mit Verkaufserfolg verknüpft werden). Dabei modifiziert der lernende Algorithmus selbstständig, aber deterministisch⁴ die ihm zur Entscheidungsfindung vorgegebenen Regeln und entwickelt eigene Regeln (vgl. Mittelstadt et al. 2016; Matthias 2004). Ein Beispiel sind neuronale Netze, bei denen Gewichtsparameter und Netzwerke von Verbindungen über die Zeit mittels der Menge an Daten, mit denen sie trainiert werden, evolvieren. Die daraus resultierenden Veränderungen der algorithmischen Regeln bleiben trotzdem das Ergebnis deterministischer Berechnungen. Wie diese Veränderung zustande kommt, ist jedoch aufgrund der Komplexität des gelernten Modells für die Nutzer:innen intransparent und nicht mehr nachvollziehbar (Burrell 2016).

Lernende Algorithmen sind typischerweise datenintensiv, weil sie mit großen Datenmengen trainiert werden müssen. Auch nichtlernende Algorithmen sind heute fähig, viele Daten zu verarbeiten und dementsprechend in der Lage, schnell überraschende Informationen zu liefern. Wir wollen in diesem Beitrag nicht den Begriff Big Data benutzen, weil dieses Konzept häufig

³ Auch wenn die Bedeutung von Algorithmus über technisch programmierte Algorithmen hinausgeht, beziehen wir uns mit diesem Begriff ausdrücklich auf computertechnisch umgesetzte Formen.

⁴ Grundsätzlich hat beim Einsatz von neuronalen Netzen der Zufall ein gewisses Maß an Einfluss auf die Ergebnisse, bedingt durch die Inputs und die Inputreihenfolge. Die Parameterinitialisierung wird zufällig getroffen. Je mehr Daten verarbeitet werden, desto mehr ist der Zufall jedoch zu vernachlässigen, weil die Parameter während des Trainings an die Daten angepasst werden (Reimers/Gurevych 2018).

mit unstrukturierten Daten vor allem aus den sozialen Medien assoziiert wird. Die Algorithmen, die in Organisationen verwendet werden, nutzen hingegen häufig umfassende und teilweise gut strukturierte, organisationsinterne Datenbestände, etwa zu Personal, Produkten oder Investitionen.

Nichtlernende und lernende Algorithmen kommen auf zwei Weisen zum Einsatz: entweder um Entscheidungen zu unterstützen oder um Entscheidungen zu ersetzen (ANONYMISIERT; Lee 2018; Karafillidis/Weidner 2015, 80ff.). Als algorithmengestützte Entscheidungen wollen wir Entscheidungen bezeichnen, bei denen Algorithmen zum Einsatz kommen, um Entscheider:innen bei ihren Entscheidungen zu informieren. Erörterungen zu dieser Art der Entscheidungsunterstützung findet man seit den 1970er Jahren im Bereich der decision support systems und heute zumeist unter dem Schlagwort Big Data. Beispiele sind unternehmerische Entscheidungen, die durch Big Data Analytics gestützt werden (McAfee/Brynjolfsson 2012), Predictive Policing (Brayne 2017; siehe Esposito et al. sowie Büchner/Dosdall in diesem Band) sowie Gerichtsurteile, bei denen die Richter:innen durch algorithmisierte Prognoserechnungen informiert werden (Cevolini/Esposito 2020; Vincent/Viljoen 2020; siehe Hedler in diesem Band). Davon zu unterscheiden sind Fälle, in denen Entscheider:innen durch Algorithmen ersetzt werden (Mittelstadt et al. 2016), beispielsweise bei der Verschlagwortung von Bildern in Online-datenbanken oder der automatischen Generierung von Kategorien für Werbezzielgruppen.

Beide Einsatzarten werden in Zusammenhang mit Verantwortung sowohl aus sozial- und geisteswissenschaftlichen als auch informatisch-ingenieurwissenschaftlichen Perspektiven diskutiert. Da die Begriffe Verantwortlichkeit (accountability) und Verantwortung (responsibility) nicht scharf konzeptionell unterschieden werden, vertreten verschiedene Studien teils sehr unterschiedliche Begriffsauffassungen, die von Rechenschaftspflichtigkeit (Reddy et al. 2019) bis hin zur verantwortungsbewussten Programmierung und Verwendung reichen (siehe beispielsweise Neyland 2016). Auch Meijer und Grimmelikhuijsen (2020), die Verantwortung und Verantwortlichkeit explizit trennen, liefern keine überzeugende Definition dieser Begriffe, denn für sie ist Verantwortung eine individuelle moralische Verpflichtung und Verantwortlichkeit stellt die damit verbundene Rechenschaftspflicht dar. Das unscharfe Verständnis beschränkt jedoch die Auseinandersetzung mit den unterschiedlichen strukturellen Konsequenzen der dahinterstehenden Konzepte (bspw. Ranerup/Hinriksen 2019; Shackelford/Raymond 2014).

Die zentrale Problematik in der Literatur betrifft die Verantwortlichkeit (Accountability) und fragt, welche Instanz bei Fehlern oder Problemen zur Rechenschaft gezogen werden kann (siehe Beckers/Teubner sowie Kette in diesem Band). Dieser Aspekt wird häufig in engem

Zusammenhang mit moralisch-ethischen Prinzipien diskutiert. In Arbeiten zu Assemblagen von Personen und Objekten – womit kontingente Konstellationen menschlicher und technischer Elemente bezeichnet werden, die an Entscheidungen beteiligt sind – ist man sich meist einig, dass Algorithmen (Prinzing 2019), Prognosesoftware wie COMPAS (Gualdi/Cordella 2021) oder Roboter (Loh 2019) zwar anfällig für Fehler sind, aber nur bedingt für diese verantwortlich gemacht werden können. Die Argumentation unterscheidet sich jedoch mit Blick auf Algorithmen als technische Elemente oder auf Assemblagen: Prinzing (2019, 280) etwa erläutert, dass Algorithmen keinen „Personen-Status“ haben und somit nicht in der Lage sind, aus kontingenten Möglichkeiten zu wählen und Entscheidungen zu treffen, weshalb sie letztlich nicht verantwortlich gemacht werden können. Gualdi und Cordella (2021) argumentieren hingegen, dass die gesamte Assemblage von Technologie und Mensch in den Blick genommen werden muss, um Verantwortlichkeit verorten zu können (siehe auch Annany/Crawford 2018). Sie konstatieren, dass in den soziotechnischen Systemen komplexe Gefüge entstehen, in denen technische und menschliche Akteure auf bestimmte Weise zusammenwirken, sodass sich die Verantwortlichkeiten verschieben. Diese Verschiebung findet jedoch nicht nur unter den Menschen statt, sondern überträgt sich auch auf den Algorithmus, was die Zuschreibung von Verantwortlichkeit verwischt. Sie fordern daher einen transparenten Umgang mit Algorithmencodes, verbleiben aber dabei, dass Menschen die Verantwortung übernehmen müssten.

Die hohe Komplexität im Entstehungsprozess maschineller Entscheidungen sowie die Intransparenz algorithmischer Kalkulationen macht eine eindeutige Identifikation von Fehlerquellen nahezu unmöglich und lässt Forderungen nach Transparenz laut werden. Mehr Transparenz und eine Offenlegung des Programmcodes soll eine bessere Zurechenbarkeit auf die Urheber:innen hinter den Algorithmen erlauben, so die Hoffnung vieler (kritisch hierzu Reddy et al. 2019). Damit verbunden sind Forderungen nach Komitees (Liao/Liao 2020), Ethikkodizes (Prinzing 2019) und sogar einer Social Responsibility of AI Algorithms (Cheng et al. 2021).⁵ Sie sollen Verantwortlichkeiten in algorithmisch gestützten Abläufen festlegen, damit diese nicht mehr „nach Belieben zwischen den Individuen und zwischen Individuen und Organisation hin- und hergeschoben“ werden (Debatin 2016, 68).

⁵ Mittlerweile gibt es mehrere Versuche, Transparenz zu schaffen. So hat beispielsweise die NGO Algorithm-Watch das Projekt Unding gegründet. Es soll Nutzer:innen unterstützen, „automatische Entscheidungen an[zu]fechten“, wie es auf der Website selbst heißt (unding.de; 19.08.2021) und listet Beispiele für algorithmische Verfehlungen auf. Darunter fallen die Nichterkennung durch Fotoautomaten von Menschen mit schwarzer Hautfarbe oder eine vermeintlich ungerechtfertigte negative Schufa-Auskunft. Das Projekt bietet an, die Kommunikation mit der Firma für den Beschwerdefall zu übernehmen. Unding setzt sich also für die Betroffenen automatisierter Fehlschlüsse ein und versucht, Verantwortliche für ihr Problem zu finden, indem es auf die Anbieter:innen und Betreiber:innen der Algorithmen Druck ausübt, Verantwortlichkeit zu übernehmen.

Transparenz kann die komplexe Problematik der Zuordnung von Verantwortlichkeiten in Assemblagen aber nicht lösen (Ananny/Crawford 2018). Einen Algorithmus offenzulegen, ist für die meisten wirtschaftlich orientierten Unternehmen ohnehin keine Option, da sie zugleich ihr Firmengeheimnis und Alleinstellungsmerkmal preisgeben würden. Mehr noch bringt eine Offenlegung des Codes bei lernenden Algorithmen häufig wenig, weil der vom Algorithmus selbst fortgeschriebene Code sogar für Expert:innen zu komplex ist, um verstanden werden zu können (Burrell 2016). Von einer eindeutigen Zuschreibung von Verantwortlichkeiten in soziotechnischen Assemblagen sehen Ananny und Crawford (2018, 984f.) unter anderem aus diesen Gründen ab. Sie argumentieren, dass eine Offenlegung keine verantwortlichen Systeme schaffe, weil in komplexen Assemblagen vollständige Transparenz unmöglich sei. Vielmehr gehe es darum, die spezifischen Limitationen der Transparenz in spezifischen Assemblagen offenzulegen. Daher verorten sie Verantwortlichkeit in der soziotechnischen Assemblage als Einheit.

Wir wollen zeigen, dass die Konsequenzen von Algorithmen im Hinblick auf Verantwortung und Verantwortlichkeit deutlich weitreichender sind als in der Literatur diskutiert: Assemblagen weisen eine zerlegte Verantwortung auf, die Verantwortlichkeiten so verunsichert, dass neue Möglichkeiten der Schuldzuschreibung entstehen, während herkömmliche Modi an Relevanz verlieren. Das verkompliziert das Verhältnis beider Bereiche, was zahlreiche Folgen hat. Wir weisen in diesem Beitrag auf die Umstrukturierung der Bedingungen von organisationalem Vertrauen und die Innovationsfähigkeit von Organisationen hin.

3 Der Unterschied zwischen Verantwortung und Verantwortlichkeit

Wie erwähnt, wird selten zwischen Verantwortung und Verantwortlichkeit differenziert. Für eine fundierte Diskussion ist eine Abgrenzung jedoch unerlässlich, denn die Begriffe adressieren unterschiedliche Dimensionen von Entscheidungsprozessen. *Verantwortung* bezeichnet den „sozialen Prozeß der Informationsverarbeitung [...], der zugleich der Absorption von Unsicherheit und der Bewußtseinsentlastung dient“ (Luhmann 1964, 174). Verantwortung wird im Prozess des Entscheidens denen zugeschrieben, die Information verarbeiten. Hierdurch wird Unsicherheit absorbiert und fehlende Gewissheit aufgefangen. Mit *Verantwortlichkeit* ist hingegen die Rechenschaftspflicht für Fehler gemeint. In jeder Entscheidung werden beide Dimensionen aktualisiert. Das Besondere an Organisationen ist nun, dass sie sowohl Verantwortung als auch Verantwortlichkeit formal zuteilen.

3.1 Unsicherheitsabsorption durch Verantwortung

Verantwortung ist „der ungedeckte Informationswert einer Entscheidung, der Überschuss an Information, die jemand gibt, im Vergleich zu der, die er erhalten hat“ (Luhmann 1964, 175).

Sie füllt gewissermaßen die Lücke, die dadurch entsteht, dass jede Entscheidung nur trotz fehlender Information getroffen werden kann. Die Bedeutung der Verantwortung wird erst klar, wenn man sich vergegenwärtigt, dass Entscheidungen immer nur unter Bedingungen der Unsicherheit getroffen werden können (March/Simon 1958; Luhmann 2000). Nach Luhmann sind Entscheidungen stets kontingent, das heißt, jede Entscheidung wird als Wahl zwischen Alternativen realisiert und ist nicht zwangsläufiges Ergebnis eines Kalküls. Würden dagegen Berechnungen eindeutig aufeinander folgen, vollständige Informationen zur Verfügung stehen und über jede Verknüpfung rational reflektiert werden können, könnte von einer Entscheidung keine Rede sein. In diesem Sinne formuliert Heinz von Foerster: „Nur *die* Fragen, die im Prinzip unentscheidbar sind, können *wir* entscheiden“ (von Foerster 1993, 73; Herv. i. Orig.).

Mit dem Begriff der Verantwortung spricht Luhmann aber nicht nur die sachliche Dimension der Informationsverarbeitung an, sondern betont zugleich, dass dadurch eine Entlastung in der Sozialdimension stattfindet. Verantwortung zu übernehmen, heißt also, in einer spezifischen Situation das Risiko der anderen zu übernehmen, unter Bedingungen der Ungewissheit zu handeln. Wenn die Leitung etwa ein Budget festlegt, kann das Team planen, ohne sich mit den Unsicherheiten zu beschäftigen, die mit der Finanzierung verbunden sind. So werden Orientierungsgrundlagen für weitere Kommunikationsprozesse und Kommunikationsteilnehmer:innen geschaffen und nachgelagerte Entscheidungen von einigen Dimensionen der Unsicherheit entlastet. Dieser Prozess der Unsicherheitsabsorption (March/Simon 1958, 165; Luhmann 2000, 183-192) findet immer dann statt, wenn die Informationsverarbeitung einer Stelle von anderen Mitgliedern einfach hingenommen wird und die Informationsgrundlage der vorherigen Entscheidung nicht eigens kontrolliert werden muss, um weiterhin arbeiten und entscheiden zu können. In dieser soziologischen Begriffsfassung meint Verantwortung also keine moralisch gebotene oder ethisch notwendige individuelle Haltung, wie von anderen Autor:innen vorgeschlagen (etwa Meijer/Grimmelikhuijsen 2020), sondern bezeichnet eine Relation kollektiver Informationsverarbeitung und Risikoübernahme.

Dies erklärt aber noch nicht, wieso eine Entscheidung, die unter Unsicherheit getroffen wird, überhaupt akzeptiert wird und Orientierungsgrundlagen für andere schafft. In Organisationen hängt dies letztlich von strukturellen Absicherungen im sozialen Setting ab. Die Besonderheit von Organisationen als soziale Systeme impliziert, dass sie Stellen mit Entscheidungszuständigkeiten festlegen; sie definieren somit Verantwortungsbereiche besonders klar. Sie etablieren auf diese Weise die Erwartung, dass die Informationsverarbeitung dem Organigramm entsprechend von bestimmten Personen auf bestimmten Stellen (Personal, Einkauf, PR usw.) mit Bezug auf unterschiedliche Aufgaben geleistet wird (Luhmann 1964, 177). Um Unsicherheit zu

absorbieren, verlassen sich Organisationen nicht (nur) auf Automatismen, Gewohnheiten und persönliche Kontakte, sondern ebenfalls auf die formale Erwartung, dass Entscheidungen an bestimmten Stellen und zu bestimmten Zeiten auch dann getroffen werden, wenn die Informationssituation es eigentlich nicht zulässt (ebd.).

Hier liegt ein wesentlicher Grund dafür, dass Stellen mit bestimmten Kompetenzen ausgestattet werden. Kompetenzen lassen einen Informationsvorsprung im Hinblick auf bestimmte Fragen erwarten, so dass andere von der Suche nach richtigen Informationen, Informationsquellen und -zusammenhängen entlastet werden (ebd., 179). Unsicherheitsabsorption gelingt, weil die organisationale Strukturvorgabe nicht nur festlegt, wer und wann jemand legitimiert ist, Entscheidungen zu treffen, sondern auch, wer gegenüber welchen Stellen berichtspflichtig ist, für wen also gewisse Entscheidungen als Orientierungsgrundlage dienen sollen oder müssen. So wird die Wahrscheinlichkeit erhöht, bestimmte Entscheidungen annehmen zu müssen (ebd., 161; 177f.). Da aber die Unsicherheit der Entscheidungssituation allen bekannt ist und durch die Stellenstruktur nicht eliminiert werden kann, können mehrere Faktoren die Annahme einer Entscheidung zusätzlich erleichtern, seien es Ansehen und Autorität bestimmter Entscheider:innen oder eine überzeugende Darstellung. Dabei fällt die Bewertung, ob eine Entscheidung relevant ist, unterschiedlich aus. Wird der Entscheidungsspielraum als groß eingeschätzt, so wird der Beitrag zur Unsicherheitsabsorption entsprechend höher bewertet (Luhmann 2000, 198f.).

3.2 Verantwortlichkeit als Rechenschaftspflicht

Die Stellenstruktur schafft auch eine Grundlage, um in Organisationen Verantwortung mit Rechenschaftspflicht zu verknüpfen, wobei Verantwortung und Verantwortlichkeit nicht immer koinzidieren, wie wir im Abschnitt 3.3 ausführen werden. Verantwortlichkeit bedeutet, dass Personen, Teams oder ganze Organisationen als Urheber:innen und somit Verantwortliche für Erfolge und vor allem für Misserfolge einer Entscheidung sowie für (vermeintliche) Fehler zur Rechenschaft gezogen werden können. In Organisationen kann man also erwarten, dass die Personen, die formal für bestimmte Entscheidungen zuständig sind, von Vorgesetzten bei Problemen zur Rechenschaft gezogen werden können und im Erfolgsfall eventuell Anerkennung finden. Verantwortlichkeit ermöglicht es der Organisation, die komplexen Umstände, Personen oder Prämissen, die zu einer Entscheidung geführt haben, auszublenden, sodass eindeutig ein:e Urheber:in oder Schuldige:r gefunden werden kann (Luhmann 2000, 197f.).

Organisationen können unterschiedliche Verantwortlichkeiten festlegen. Grundlegend geht es um den Umgang mit und die Kontrolle von Fehlern oder Normverstößen. Verantwortlichkeit ist in Organisationen daher eine wesentliche Bedingung dafür, dass Verantwortung überhaupt

funktionieren kann, das heißt Bedingung dafür, dass man sich auf Entscheidungen anderer verlassen kann, weil man ihnen formal vertraut (Luhmann 1964, 180). Insofern durch Verantwortlichkeit fehlerhaftes Verhalten reflektiert werden kann, bietet sie Lerngelegenheiten. Denn wenn die Urheber:innen von Fehlern identifiziert werden, hat man feste Ansatzpunkte, um Probleme mit dem Ziel zu erörtern, sich bei der nächsten Gelegenheit anders zu verhalten (Luhmann 2000, 198). Dabei ist anzumerken, dass durch Verantwortlichkeit nicht der gesamte Unsicherheitsbereich von Entscheidungen abgedeckt werden kann. Manche Fehler sind auch bei bestem Wissen und Gewissen nicht zu vermeiden, weil sie der Komplexität der Entscheidungslage geschuldet sind. Die Wirkungskraft der Verantwortlichkeit kann weiterhin reduziert werden, indem eine Zuschreibung von Verantwortung auf Personen möglicherweise gar nicht in Anspruch genommen wird, etwa um zu verhindern, dass die Unzulänglichkeiten der Organisationsspitze oder Vorgesetzter sichtbar werden (Luhmann 1964, 183). Umgekehrt kann eine personalisierte Zuschreibung dafür verwendet werden, die Komplexität von Entscheidungslagen zu leugnen. Bei Skandalen lässt sich zum Beispiel beobachten, dass moralische Verantwortlichkeit ins Spiel gebracht wird, um trotz der Unübersichtlichkeit und Vielschichtigkeit des Problems Sündenböcke zu finden und bestrafen zu können. Mit dieser Praxis werden letzten Endes weitreichende organisationale Veränderungen erfolgreich blockiert (Besio 2014).

3.3 Die Diskrepanz zwischen Verantwortung und Verantwortlichkeit

Organisationssoziologisch ist besonders interessant, dass Verantwortung und Verantwortlichkeit nicht zwangsläufig zusammenfallen (dazu auch Brunsson 1989, 179-193). Das kann aufgrund formaler Festlegungen der Fall sein oder informale Gründe haben. Formal fallen Verantwortung und Verantwortlichkeit auseinander, wenn eine Stelle Entscheidungen treffen darf, eine andere aber dafür rechenschaftspflichtig ist. Das ist typischerweise dann zu beobachten, wenn Vorgesetzte für die problematischen Entscheidungen ihrer Mitarbeitenden von der Organisationsleitung belangt werden oder sogar rechtlich dafür haften müssen. Häufig sind Organisationen formal so strukturiert, dass grundlegende Unsicherheitsabsorption (Verantwortung) unten stattfindet, weil dort Informationen aus verschiedenen Quellen bearbeitet werden und in Entscheidungsprozesse einfließen, während der Spitze Erfolge und Fehler zugeschrieben werden (Luhmann 1964, 184).

Informal werden Verantwortung und Verantwortlichkeit getrennt, wenn jemand für eine Entscheidung verantwortlich ist, also das Risiko der fehlenden Information anderen abgenommen hat, aber jemand anderes (oft hierarchisch höherstehend) sich im Erfolgsfall damit brüstet – also die eigene Verantwortlichkeit in den Vordergrund stellt. Es lässt sich auch beobachten, dass im Problemfall Verantwortlichkeit in einer Art Blame Game hin- und hergeschoben wird.

Sobald Fehler sichtbar werden, wird schnell versucht, Entscheidungen durch die Zurechnung auf scheinbar zwingende äußere Auslösefaktoren (etwa krisenhafte Ereignisse oder die Gesetzeslage) als alternativlos zu stilisieren, sodass niemand verantwortlich gemacht werden muss (Brunsson 1989, 182-184). Strategien wie die Verschiebung von Verantwortung, die Verwässerung der Verantwortlichkeit durch Beteiligung anderer Stellen, die Bezugnahme auf etablierte Routinen oder auf angeblich objektive Fakten, sind in Organisationen keine Seltenheit (Luhmann 1964, 185).

Luhmann beschreibt nicht nur pointiert das Verhältnis zwischen Verantwortung und Verantwortlichkeit, sondern macht gleichsam auf positive wie negative Folgen der Diskrepanz zwischen diesen aufmerksam. Ein starker Kontrast zwischen Verantwortung und Verantwortlichkeit führt dazu, dass die Kontrollfunktion der Verantwortlichkeit grundlegend außer Kraft gesetzt wird und das Vertrauen, das sie erzeugen soll, unterminiert wird (Luhmann 1964, 180f.). In der Regel vertraut man Entscheider:innen, wenn man weiß, dass man sie für ihre Entscheidungen gegebenenfalls zur Rechenschaft ziehen kann. Dies wird allerdings durch eine zu große Kluft zwischen Verantwortung und Verantwortlichkeit verunmöglicht. Wenn man aber andererseits nicht immer zur Rechenschaft gezogen werden kann oder wenn es gelingt, die eigene Verantwortung zu vertuschen oder zu verstecken, dann entstehen für einzelne Entscheidungen Freiheitsgrade mitsamt der Möglichkeit, riskantere und innovative Entscheidungen zu treffen (Luhmann 1964, 187).

Dies alles gehört zum Alltag der Organisation. Der vermehrte Einsatz von datenintensiven Algorithmen verändert jedoch die Konstruktionen von Verantwortung und Verantwortlichkeit in Organisationen sowie ihr Verhältnis zueinander, weil Algorithmen zunehmend als Entscheidungsteilnehmer oder sogar Entscheider behandelt werden (Diakopoulos 2015; Esposito 2017). Informationsverarbeitung und entsprechende Zuständigkeiten für Entscheidungen (also Verantwortung) verschieben sich teils auf Algorithmen selbst, teils auf diffuse Assemblagen aus Personen und Objekten (Annany/Crawford 2018), und Verantwortlichkeiten bleiben weitgehend ungeklärt und problematisch (Matthias 2004; Mittelstadt et al. 2016).

4 Implikationen datenintensiver Algorithmen für Verantwortung und Verantwortlichkeit

Die Implikationen des Einsatzes von datenintensiven Algorithmen diskutieren wir im Folgenden mit Bezug auf die dadurch ausgelöste Zerlegung von Verantwortung, die damit verbundene kontroverse Verantwortlichkeit sowie die Verkomplizierung ihrer Verhältnisse.

4.1 Zerlegung von Verantwortung

Durch den Einsatz von datenintensiven Algorithmen wird Verantwortung in Organisationen teils neu verteilt und teils verschoben. Zunächst werden Entscheidungen mit maschinellen Kalkulationen vorbereitet oder Entscheider:innen sogar durch algorithmische Berechnungen ersetzt (ANONYMISIERT; Lee 2018; Karafillidis/Weidner 2015, 80ff.). Folgen können Verschiebungen im Rollensystem und somit im Kompetenz- und Machtgefüge der Organisation sein, die auch davon abhängen, wem die Kontrolle über den jeweiligen Prozess zugerechnet wird (Karafillidis/Weidner 2015).

Wir behaupten, dass Maschinen heutzutage Verantwortung übernehmen können oder ihnen Verantwortung zugerechnet werden kann – sofern man anerkennt, dass nichtmenschliche Objekte wie Algorithmen in Kommunikations- und Entscheidungsprozessen mitwirken (hierzu ausführlich Esposito 2017). Dazu ist es nötig, das Konzept der Verantwortung als Form der Unsicherheitsabsorption vom Menschen zu lösen (siehe hierzu auch allgemeine, breit rezipierte Annahmen aus den Science and Technology Studies: Cerulo 2009; Bauer et al. 2017). Bereits Luhmann hatte dies grundsätzlich so festgestellt, als er schrieb:

„Verantwortung steckt in jedem Beitrag zur Unsicherheitsabsorption, auch in dem, den die Maschine leistet. Grundsätzlich ist Verantwortung im Sinne von Unsicherheitsabsorption eine Funktion kommunikativer Prozesse und keine ‚Eigenschaft‘ von Menschen. [...] Wenn man Verantwortung in unserem Sinne als Funktion kommunikativer Prozesse definiert, hat es keinen Sinn mehr, sie individuell zuzurechnen“ (Luhmann 1966, 105)

Nach Luhmann (1964, 176) kann die maschinelle Verarbeitung von Informationen ein funktionales Äquivalent für Verantwortung sein: Neben Institutionen, Gewohnheiten und anderen automatisierten Verhaltensweisen leisten Maschinen Unsicherheitsabsorption, bieten Orientierung und machen Entscheidungen samt Verantwortungsübernahme an einigen Stellen sogar überflüssig.

Algorithmisch unterstützte Entscheidungen

Trotz dieser Betonung der Funktion von maschinellen Kalkulationen stellt Luhmann für die Situation der 1960er Jahre fest: „Die heiklen Phasen des Prozesses der Unsicherheitsabsorption müssen immer noch von Menschen getragen werden; den Computer kann man allenfalls verantworten lassen, was nicht angegriffen werden kann“ (Luhmann 1966, 60). Dies hat sich geändert, da Algorithmen mit unterstützender Funktion nicht nur in der Privatwirtschaft, sondern zunehmen in der öffentlichen Verwaltung (u.a. Meijer/Grimmelikhuijsen 2020), bei

richterlichen und polizeilichen Entscheidungen (Brayne 2017; Christin et al. 2015) oder im Militär verwendet werden. In all diesen Bereichen bilden sie eine Informationsbasis, um heikle Situationen einzuschätzen, etwa das Risiko der Rückfälligkeit von Straftäter:innen oder die automatische Detektion von Menschen auf Bildern in der Seenotrettung oder beim Militär.

Algorithmen können dazu führen, dass sich in diesen Bereichen die persönliche Verantwortung einzelner Entscheider:innen verkleinert. „Verantwortung ist der ungedeckte Informationswert einer Entscheidung“ (Luhmann 1964, 175), das heißt, es gibt einen Überschuss an Information, den die zuständige Person weitergibt im Vergleich zur Information, die sie als Input bekommt. Im Fall von Big Data Analytics liegt das Ausmaß der verfügbaren Information vergleichsweise hoch. In dem Moment, in dem sich die Entscheider:innen auf algorithmische Informationen stützen und dies kommunizieren, reduziert sich die Verantwortung der entscheidenden Person in dem Ausmaß, in dem sich der Bereich der ungedeckten Information durch den Einsatz von Algorithmen verkleinert. Denn es ist nicht mehr Verdienst der entscheidenden Person, etwa Risiken und Chancen ungewisser Situationen evaluiert zu haben, sondern es handelt sich um eine Vorleistung der Maschine.

Jedoch lässt sich feststellen, dass auch heutzutage relevante Aspekte der Entscheidungsverantwortung bei konkreten Personen liegen. Bei algorithmengestützten Entscheidungen sind es weiterhin die Richter:innen, die die Urteile fällen und die Polizist:innen, die entscheiden, welcher Analyse und Prognose sie folgen und welcher nicht. In solchen Fällen gibt es nicht zwangsläufig formale Strukturen, die zu einer Verwendung von Algorithmen verpflichten und es bleibt im Ermessensspielraum der Entscheider:innen, den algorithmischen Output zu nutzen oder ihn zu vernachlässigen. So zeigen empirische Studien, dass Richter:innen und Polizist:innen algorithmisch verarbeitete Informationen unter Umständen außer Acht lassen (Christin et al. 2015; Brayne 2017).

Das hohe Maß an Verantwortung, das die Entscheider:innen in dieser Situation übernehmen, ergibt sich aus der Tatsache, dass die Entscheidung, den Algorithmus zu berücksichtigen oder zu vernachlässigen, in einer intransparenten Situation stattfindet. Angesichts der Komplexität moderner Algorithmen können Entscheider:innen algorithmische Verarbeitungsprozesse nicht im Detail nachvollziehen. Vielmehr können sie nur darauf vertrauen, dass qualitativ hochwertige Daten in den Algorithmus eingespeist und sinnvoll ausgewertet wurden (Lee 2018). Dass dies aber nicht immer der Fall ist, zeigen etliche Forschungsbeiträge. Daten, die in die algorithmischen Analysen einfließen, sind oft durch Vorurteile vorselektiert und Algorithmen häufig (unintendiert) diskriminierend programmiert (u.a. Eubanks 2018; Turner Lee 2018).

Dass die Verantwortung einzelner Entscheider:innen ihre Relevanz nicht verliert, zeigt sich zudem besonders eindeutig, wenn mehrere Algorithmen gleichzeitig eingesetzt werden. Diese sind für unterschiedliche Probleme entworfen und geben nicht immer kohärente Outputs, sodass es Aufgabe der verantwortlichen Personen ist, aus den verschiedenen Informationen, aus der Komplexität der Relation verschiedener algorithmischer Kalkulationen, eine Entscheidung zu treffen, die organisationsintern durchsetzbar ist. Wie schwer und risikoreich das wird, zeigen mehrere Studien (Berthod/Müller-Seitz 2018; Daipha 2015). So können im Notfall beispielsweise verschiedene Signale im Cockpit auf unterschiedliche technische Störungen hindeuten und es bleibt Aufgabe der Pilot:innen, aus dieser Gemengelage eine sinnvolle (und unter Umständen rettende) Interpretation zu entwickeln (Berthod/Müller-Seitz 2018).

Algorithmen als Ersatz für Entscheidungen

In Fällen, in denen Algorithmen Entscheider:innen ersetzen und nicht nur unterstützen, lassen sich noch stärkere Verschiebungen von Verantwortungen feststellen. Zunächst treffen Algorithmen auch hier keine Entscheidungen im eigentlichen Sinn, sondern produzieren auf Basis bestimmter Inputs bestimmte Outputs. Ihre Outputs sind daher immer alternativlos. Aber die Entscheider:innen fehlen, die die Outputs der Algorithmen noch einmal abwägen. Jedoch bleiben Teile der Verantwortung beim Menschen; dies gilt in erster Linie für die Programmierer:innen. Bei nichtlernenden Algorithmen, in denen bestimmte Wenn-dann-Regeln automatisiert ausgeführt werden, sind die Ergebnisse vorab bereits durch Programmierer:innen und sonstige am Programmierprozess beteiligte Stellen festgelegt. Dies hat sich seit den 1960er Jahren nicht geändert. Aus diesem Grund kann von einer Verlagerung von Verantwortung auf Algorithmen auch weiterhin nur begrenzt gesprochen werden.

Wenn man den Blick nun von nichtlernenden auf selbstlernende Algorithmen richtet, stellt sich das Ausmaß der Verantwortungsverlagerung auf die Maschine erheblich größer dar. Zwar schreibt auch ein selbstlernender Algorithmus lediglich „bestimmte Regeln auf Basis bestimmter Regeln im Zusammenspiel mit diversen Inputs“ fort (ANONYMISIERT) und bleibt daher in seiner Funktionsweise deterministisch. Man mag hier argumentieren, dass Verantwortung demnach in weiten Teilen bei den Programmierer:innen, den teils unbewussten Inputgeber:innen sowie den Nutzer:innen liege, da sie die Grundlagen für die tatsächlichen Regeln liefern, die der Algorithmus ausführt. Allerdings wissen die Programmierer:innen in diesen Fällen nicht, wie das gelernte Modell sich entwickeln wird. Sie können es gar nicht wissen, denn die Grundidee hinter einem lernenden Algorithmus ist, dass dieser trainiert wird und somit lernt, nachdem er programmiert wurde. Demnach wird von den Programmierer:innen nicht die

Unsicherheit ungedeckter Informationswerte absorbiert, sondern im Gegenteil bewusst in den Algorithmus überführt. Bei den Inputgeber:innen eines solchen Algorithmus ist es ähnlich. Nutzer:innen von technischen Systemen fungieren unter Umständen als Inputgeber:innen und trainieren manchmal wissentlich, manchmal unwissentlich Algorithmen – so wie es häufig der Fall ist, wenn Personen schlicht ihren E-Mail-Service (z. B. Spam-Erkennung) oder ein Bilder-Management-Tool (z. B. Tagging von Personen) nutzen. An diesen Stellen übernehmen die Nutzer:innen unbewusst Verantwortung, insofern sie Festlegungen auf Basis der ihnen zur Verfügung stehenden Informationen treffen, auch wenn es sich nicht um eine direkte Programmierung der Algorithmen handelt. Die Algorithmen lernen hingegen, auf Basis solcher Inputs ähnliche Festlegungen in anderen Informationssettings eigenständig zu treffen. Entsprechend lässt sich Verantwortung bei selbstlernenden Algorithmen, die Entscheidungen ersetzen, in verteilter Form sowohl bei den Programmierer:innen, den Nutzer:innen und den sonstigen Inputgeber:innen als auch bei den Algorithmen selbst verorten.

Algorithmen, Verantwortung und Stellenstruktur

In den neu entstandenen sozialen Konstellationen werden typischerweise Aufgaben des mittleren Managements von Algorithmen unterstützt oder übernommen, um große Datenmengen bearbeiten zu können und Prozesse zu beschleunigen. Dabei können die Unsicherheitsabsorption erleichtert und Entscheider:innen entlastet werden. Dies zeigen die hohen Erwartungen an die Effizienzsteigerung durch Digitalisierung. Da der Entscheidungsspielraum des mittleren Managements und somit seine Verantwortung reduziert werden kann, geht die Einführung von Algorithmen häufig damit einher, dass dieses Personal an Macht und Anerkennung verliert (Beverungen 2017; Constantiou/Kallinikos 2015; Faraj et al. 2018). Weitere Aspekte, die eine große Rolle bei der Durchsetzung von Entscheidungen auf dieser Ebene spielen, wie die Darstellung und Inszenierung von Entscheidungen oder das persönliche Charisma, verlieren ebenfalls an Relevanz.

Gleichzeitig gewinnen manche Entscheidungen in Organisationen deutlich an Reichweite, etwa bezüglich der Einführung und der Eigenschaften spezifischer Software, die tendenziell an der Organisationsspitze getroffen werden. Entscheidungen dieser Art sind komplex, weil sie einerseits die Kooperation mit externen Firmen betreffen, welche die Algorithmen programmieren, während sie andererseits gleichzeitig die interne Entwicklung neuer Expertisen erfordern und alte Arbeitsabläufe neu modellieren. Es zählt heutzutage zu den Aufgaben digitaler Führung, die passenden Programme für die Organisation auszusuchen (Thiemann 2021, 147-149). Insbesondere die Entscheidung, bestimmte Entscheidungen an Algorithmen abzugeben, ist

folgenreich, denn Entscheidungsunsicherheit wird dabei bereits weitgehend absorbiert (Luhmann 1966, 71). Dies gilt vor allem, wenn lernende Algorithmen als Ersatz für Entscheidungen verwendet werden. In solchen Fällen ist eine deutliche Verlagerung von Verantwortung nach vorn im organisationalen Entscheidungsprozess evident. Verantwortung verschiebt sich also vermehrt hin zu den Positionen, die vorab entscheiden, welche Aufgaben durch Algorithmen ausgeführt werden. Der Hochfrequenzhandel ist dafür ein gutes Beispiel. In diesem Bereich wird eine extrem hohe Zahl von Handelsentscheidungen in Sekundenbruchteilen von Algorithmen automatisiert ausgeführt, sodass vorab getroffene Entscheidungen über den Einsatz von Algorithmen bis dato nicht gekannte Auswirkungen haben (Gresser 2018).

Insgesamt erweitert der Einsatz von Algorithmen den Verantwortungsbereich der Führungskräfte, denn während auf oberen Hierarchieebenen klassisch die Schwerpunkte in der Verantwortlichkeit liegen (Luhmann 1964, 183), verschiebt sich nun auch die Verantwortung nach oben. Große Teile der Informationsverarbeitung und entsprechender Unsicherheitsabsorption können von Maschinen erledigt werden, die herkömmlich vom mittleren Management oder von Mitarbeiter:innen mit spezifischer Expertise geleistet wurden. Gewiss trägt die Spitze etwa durch Budgetierung oder Planungsentscheidungen seit jeher wesentlich zur Unsicherheitsabsorption im Entscheidungsprozess bei. Doch vor allem in den Fällen, in denen Algorithmen Entscheidungen ersetzen, erweitert sich ihr Einflussbereich auf inhaltliche Aspekte der Entscheidungsfindung wesentlich.

Im herkömmlichen Modell der Organisation erledigen verschiedene organisationale Stellen die Arbeit in geteilter Verantwortung und treffen Entscheidungen im Rahmen ihrer Kompetenz. Mit dem gesteigerten Einsatz von Algorithmen kommt es aber zu einer Veränderung organisationsinterner Relationen, die eine verschärfte Verteilung und somit Zerlegung von Verantwortung beinhaltet. An Entscheidungsprozessen sind nun komplexe Konstellationen von Akteuren und technischen Objekten (Assemblagen) beteiligt: Führungspersonen, die über die Einführung von Algorithmen entscheiden, Programmierer:innen und Trainer:innen, die für die Umsetzung sorgen, Expert:innen und Datenanalyst:innen, die die Outputs der Algorithmen lesen und darstellen, Stellen, die algorithmische Schlussfolgerungen in ihre Entscheidungen einfließen lassen und nicht zuletzt die Algorithmen selbst, die von Akteuren zwar programmiert werden, aber mit ihren Möglichkeiten und Limitationen wiederum das Handeln und Erleben der Akteure mitbestimmen. Deterministische Kalkulationen spielen hier eine Rolle, aber Erfahrung, Intuition und Risiko verschwinden in diesen Konstellationen nicht einfach, sondern werden, wie beschrieben, in einem komplexen Prozess neu verteilt und rekombiniert. Dies hat weitgehende organisationale Konsequenzen, etwa die Veränderung von Legitimationsmechanismen (Besio

et al. 2022) oder die Neuordnung der Machtverteilung betreffend (Beverungen 2017; Constantiou/Kallinikos 2015; Faraj et al. 2018). Vor allem aber hat dies Folgen für die Zuschreibung von Leistungen und Fehlern, also für die Verantwortlichkeit.

4.2 Kontroverse Verantwortlichkeit

Die Neuverteilung von Verantwortung muss nicht mit einer formalen Neuverteilung von Verantwortlichkeit einhergehen. Es kann durchaus vorkommen, dass die Verarbeitung von Informationen und die damit verbundene Reduktion von Komplexität maschinell realisiert wird, die Zurechnung von Verantwortlichkeit jedoch wie gewohnt verläuft. Es wird jedoch immer schwieriger, nur eine bestimmte Instanz für Fehler verantwortlich zu machen (Anany/Crawford 2018; Gualdi/Cordella 2021; Matthias 2004). Während die Einbeziehung verschiedener Instanzen die Informationsverarbeitung und die Unsicherheitsabsorption erleichtern kann, verkompliziert sie gleichzeitig die Zuschreibung von Leistungen und Fehlern. Denn Verantwortlichkeit verstreut sich in Assemblagen und lässt sich nicht mehr eindeutig verorten. Wenn Mensch-Maschine-Assemblagen beim Zustandekommen von Entscheidungen beteiligt sind, fällt eine wichtige Voraussetzung für die Funktionsweise der Verantwortlichkeit weg: die eindeutige Zuschreibung von Handlungen auf eine Person (Luhmann 1964, 180f.). Dies führt zu veränderten Strategien der Schuldzuschreibung und dazu, dass Verantwortlichkeit verwischt.

Der Verweis auf die Maschine

Brisant sind Situationen, in denen die involvierte Maschine (etwa ein Such- oder Klassifikationsalgorithmus) diskriminierend wirkt oder fehlerhaft ist. Häufig kann in solchen Fällen die Verantwortung auf den Algorithmus in einer Assemblage abgeschoben werden, weil das Gefüge von Mitwirkenden an einer Entscheidung undurchsichtig ist und konkrete Zuordnungen verschleiert werden können. Dann wird der Algorithmus, die nichtmenschliche Entität, als Sündenbock vorgeschoben, sodass die rechenschaftspflichtigen Personen im Verborgenen bleiben. Wenn die Bezugnahme auf die Maschine oder den Algorithmus gelingt und dieser als zwingender Auslösefaktor inszeniert werden kann, entlastet das je nach Kontext die Organisation und erspart den beteiligten Entscheider:innen Diskussionen um die konkrete Verantwortlichkeit für Fehler und unangemessene Entscheidungen.

Diese durchaus gängige organisationale Praxis kann anhand einer Studie von Wynants et al. (2020) zu lernenden Algorithmen illustriert werden. Eine Untersuchung von Klassifikationsalgorithmen in der medizinischen Diagnostik, welche auf Basis von Lungenröntgenaufnahmen vorhersagen sollten, ob Patient:innen an Covid-19 erkrankt sind oder nicht, identifiziert als eine

Fehlerquelle Kreuzkorrelationen mit diagnostisch irrelevantem Text auf den Röntgenaufnahmen. Beispielsweise gab es einen Trainingsdatensatz mit Aufnahmen aus einem Provinzkrankenhaus mit sehr wenigen Covid-Kranken und einem Schwerpunktkrankenhaus, in dem sehr viele Covid-19-Fälle behandelt wurden. Auf den Röntgenbildern der Patient:innen befanden sich ebenfalls Metadaten in Textform, die je nach Krankenhaus in verschiedenen Schriftarten gehalten waren. Die mit diesem Material trainierten Algorithmen lernten die Unterscheidung in Covid/Nicht-Covid nicht nur anhand der Lungenaufnahme festzulegen, sondern berücksichtigten auch die Schriftart auf den Röntgenbildern, um ihre Genauigkeit zu erhöhen. Optimiert man die Vorhersage auf Grundlage solcher für die Diagnosefunktion irrelevanter Daten, wird das Modell für den praktischen Einsatz im Klinikbetrieb wertlos. Im schlimmsten Fall werden diagnostische Fehlentscheidungen mit weitreichenden Konsequenzen für die Behandlung von Patient:innen getroffen, die dann, scheinbar naheliegend, dem Algorithmus zugerechnet werden, ohne die verzweigten Entscheidungslagen, die zu dieser Situation geführt haben, in das Verantwortlichkeitskalkül aufzunehmen.

Meist ist der Bezug auf die Maschine aber strittig. Stattdessen müssen Entscheider:innen Verantwortung zeigen, obwohl ihre Entscheidungen auf algorithmisch verarbeiteten Informationen basieren, die sie selbst nicht zurückverfolgen können. In solchen Situationen kann die Verfügbarkeit von zahlreichen maschinell verarbeiteten Daten eine weitere Strategie der Entlastung von Verantwortung eher erschweren: den Verweis auf die Komplexität der Lage. Da man in Organisationen häufig mit der Illusion lebt, dass Software zuverlässige, objektive und umfassende Informationen liefert (u.a. Turner Lee 2018.), wird es immer schwieriger, auf die Komplexität, die Undurchschaubarkeit und die Ungewissheit der Entscheidungslage zu verweisen, um Entscheidungen zu rechtfertigen und sich dadurch der Verantwortung zu entziehen.

Zu berücksichtigen ist zudem, dass der Verweis auf Maschinen nicht zufriedenstellend ist, weil sich diese weder rechtfertigen und entschuldigen können, noch sind sie anfällig für organisationale Sanktionen wie Lohnminderungen oder Entlassungen. So können Fehler oder Verstöße gegen Normen und Regeln, die etwa im Fall von algorithmisch basierten Diskriminierungen zutage treten, nicht mit üblichen organisationalen Mitteln bearbeitet werden. Folglich bleibt die Suche nach Schuldigen eine zentrale Praxis von Organisationen.

Blame Game

Das komplexe Gefüge verschiedener Akteure und technischer Objekte in der Mitwirkung an Entscheidungsprozessen eröffnet weitere Möglichkeiten der Schuldzuweisung bei Fehlern

sowie der Verschleierung von Verantwortlichkeit und des Einsatzes von Sündenbockstrategien. So können Personen oder Teams, ganze Abteilungen, Firmen oder Behörden zur Rechenschaft gezogen werden, sofern sie sich für die Nutzung der Maschine eingesetzt haben oder sich unkritisch auf das maschinelle Ergebnis verlassen. Doch auch diese Strategien führen zu keiner eindeutigen und stabilen Zuschreibung von Leistungen und Fehlern.

Entscheidungsprozessen, die Algorithmen einbeziehen, sind Programmierer:innen und Datenwissenschaftler:innen maßgeblich beteiligt. Sie erstellen die algorithmischen Modelle, trainieren sie, sammeln die dafür notwendigen Daten und kontrollieren ihre Qualität (vgl. Feier et al. 2021). Daher wird häufig dafür plädiert, die Verantwortlichkeit bei Personen, insbesondere den Programmierer:innen und Inputgeber:innen zu verorten (Elish 2016; Mittelstadt et al. 2016). Allerdings ist auch diese Zuschreibung im Fall lernender Algorithmen nicht einfach. Denn die Programmierer:innen können, wie erwähnt, nicht wissen, wie der Algorithmus sich entwickeln wird. Und die Inputgeber:innen wissen oftmals nicht einmal, dass sie Inputs für einen lernenden Algorithmus produzieren und falls doch, wissen sie nicht, für welche algorithmischen Kalkulationen ihre Inputs verwendet werden.

Das Beispiel automatisierter Bildverschlagnwortung von Google verdeutlicht das Problem. Vor einigen Jahren wurde bekannt, dass diese Bildverschlagnwortung rassistische Ergebnisse produziert und Menschen mit dunkler Hautfarbe oftmals als „Gorillas“ kategorisierte (Grush 2015). Es liegt berechtigterweise nahe zu argumentieren, dass diejenigen verantwortlich sind, die den Algorithmus zuvor hauptsächlich mit Bildern weißer Menschen trainiert haben. Allerdings spielen hier auch andere Faktoren eine Rolle – etwa, dass Fotokameras seit jeher einer rassistischen Praxis folgen, weil sie auf das Ablichten weißer Hautfarbe optimiert wurden (Lewis 2019). Für solche durchaus relevanten Aspekte kann man den Inputgeber:innen rund um den Algorithmus zur Verschlagnwortung nur schlecht Verantwortlichkeit zurechnen. Die bereits aufgezeigte Zerlegung der Verantwortung tritt an dieser Stelle brisant zutage und macht es schwierig, Verantwortlichkeiten plausibel zuzurechnen. Verantwortlichkeit löst sich auch hier vielmehr in einem Netzwerk von Kommunikationshistorien oder in Assemblagen auf.

Von der Assemblage zur Organisation

In der Literatur wird dafür plädiert, dass bei fehlerhaften Entscheidungen ganze Assemblagen zur Rechenschaft gezogen werden sollen (Ananny/Crawford 2018). Doch wie lassen sich die Grenzen solcher Assemblagen eindeutig festlegen? Ein Ausweg in der Praxis ist, die gesamte Organisation statt einzelner Personen, Stellen oder Assemblagen verantwortlich zu machen (Martin 2019).

Solche Zuschreibungsprozesse sind nicht nur für komplexe, lernende Algorithmen, sondern auch für lineare Algorithmen, die in komplexe Systeme eingebettet sind, zu beobachten. Die Tragweite dieser Entwicklung zeigt sich zum Beispiel beim Umgang mit den Unfällen, die der Flugzeugtyp Boeing 737 verursacht hat. Die Abstürze der beiden Maschinen der neuesten Generation 737 MAX können, simplifiziert ausgedrückt, als Folge der Digitalisierung betrachtet werden, weil es ohne den Einbau eines digitalen Assistenzsystems keine Abstürze gegeben hätte (McDermid 2020). Die Frage, wer oder was für diese Abstürze verantwortlich war, begleitete lange die mediale Berichterstattung. Allerdings konnte man für die Abstürze kaum die Techniker:innen, die die Software programmiert hatten, verantwortlich machen. Was die Pilot:innen angeht, stellte sich sogar heraus, dass sie nicht einmal um die Existenz dieses Systems wussten. Andere Verantwortliche ließen sich ebenfalls nicht benennen – auch beispielsweise nicht Manager:innen, die über den Einsatz dieser Software entschieden hatten. Als verantwortlich für diese Katastrophen wurden am Ende das Unternehmen Boeing und die zuständige Aufsichtsbehörde, die Federal Aviation Administration (FAA), identifiziert, also zwei Organisationen.

Organisationen werden in dem Maße als soziale Systeme verantwortlich gemacht, wie Algorithmen Entscheidungen nicht nur unterstützen, sondern Entscheider:innen durch solche ersetzt werden. Im Fall der Verschlagwortung von Bildern wurde Verantwortlichkeit schlussendlich nur auf Ebene der Organisation verortet: Google wird als Unternehmen für die diskriminierende Praxis verantwortlich gemacht und soll für Besserung sorgen (Vincent 2018).

Angesichts der Komplexität und Intransparenz werden rechtlich zwei herkömmliche Mechanismen für vielversprechend erachtet, um Verantwortlichkeit den beteiligten Organisationen zuzuschreiben. Ein erster Mechanismus ist die Produkthaftung, die besagt, dass Hersteller:innen oder Betreiber:innen technischer Systeme für Fehler eigener Produkte haften müssen. Diese Form der Haftung unterstreicht die Neutralität der Technik und verweist eindeutig auf diejenigen, die entschieden haben, Algorithmen einzusetzen. In Bereichen, in denen besonders hohe Schäden möglich sind, kann zudem die Gefährdungshaftung hinzutreten. Diese greift, wenn kein direktes Fehlverhalten festgestellt werden kann. Jedoch können beim Einsatz von Algorithmen nicht nur das Design und die Produktion Fehler verursachen, sondern auch Datenqualität und Analyseverfahren. Insbesondere bei lernenden Algorithmen ist die Distanz zwischen Einführungs- bzw. Entwicklungsentscheidung und algorithmischem Output groß, und so sind mögliche Schäden schwer vorherzusehen und nicht kontrollierbar. Dementsprechend scheint eine Zuschreibung, die einzig auf die Hersteller:innen oder Betreiber:innen abzielt, unverhältnismäßig. Diese Probleme und mögliche Lösungen werden in der Rechtswissenschaft intensiv

diskutiert (Ballestrem et al. 2020; Beckers/Teubner 2021, siehe auch Teubner/Beckers in diesem Band).

Aus der Perspektive der Organisationen, die Algorithmen betreiben, und/oder deren Versicherungen implizieren beide Haftungsformen hohe und kaum kalkulierbare Risiken. Externe Schuldzuschreibungen können unter den genannten intransparenten Bedingungen zudem intern schwerwiegende Folgen mit sich bringen.

Der Verweis auf die Organisation als Verantwortliche bedeutet konkret, dass die hierarchische Spitze der Organisation, also Personen, die qua formal-hierarchischem Entscheidungsbaum nach außen für die gesamte Organisation stehen, für das Verhalten und die Entscheidungen ihrer Mitarbeiter:innen zur Rechenschaft gezogen werden. Ihre hierarchisch legitimierte Gesamtverantwortlichkeit befugt sie aber auch, intern Schuldige unter Führungskräften zu suchen, die wiederum die Schuldfrage nach unten weiterreichen. Das kommt einer Einführung von Unsicherheit in die Mechanismen der Unsicherheitsabsorption auf unteren Hierarchieebenen gleich. Dies kann zudem den Wunsch nach neuen Überwachungsmöglichkeiten und weitergehenden hierarchischen Befugnissen mit sich bringen. Digitale Technologien gelten auch in dieser Hinsicht als vielversprechend, um beispielsweise Ressourcenverbrauch oder Zeiteinsatz von Mitarbeiter:innen für Vorgesetzte digital transparent zu machen. Wer an der Spitze steht und zur Rechenschaft gezogen werden kann, verfügt ohnehin über Macht, die Arbeit der anderen zu kontrollieren (Luhmann 1964, 188). Mit neuen Überwachungstechniken kann aber die Machtquelle, die von der hierarchischen Position ausgeht, gestärkt werden. Auch wenn in der Außenperspektive die Spitze verantwortlich bleibt, löst dies intern eine weitere Suche nach Verantwortlichkeiten aus. So kann diese Praxis der organisationalen Gesamtverantwortlichkeit intern Kaskadeneffekte auslösen, die Mitarbeiter:innen auf unteren hierarchischen Ebenen möglicherweise verunsichern. Dies kann die Verschleierung der Verantwortung sowie das Blame Game noch einmal ankurbeln.

5 Verkomplizierung des Verhältnisses von Verantwortung und Verantwortlichkeit

Während Verantwortung im Kontext von Assemblagen weiterhin wirkt und die dadurch geleistete Unsicherheitsabsorption sich sogar mit gesteigerter Wirksamkeit entfaltet, scheinen herkömmliche Mechanismen der Verantwortlichkeit infrage gestellt zu werden. Dies verkompliziert das Verhältnis zwischen Verantwortung und Verantwortlichkeit. Wir präsentieren abschließend einige Annahmen, die unserer Ansicht nach wichtige Aspekte dieser Problematik hervorheben.

Generell fällt auf, dass mit dem gesteigerten Einsatz von datenintensiven Algorithmen die Kontrollfunktion der Verantwortlichkeit geschwächt ist. Diejenigen, die zur Rechenschaft gezogen werden, seien es externe oder interne Programmierer:innen, Führungskräfte oder die Organisationspitze, sind an den relevanten Prozessen der Informationsverarbeitung ihrer Organisation gar nicht mehr oder nur teilweise beteiligt. Sofern sie daran teilhaben, können sie aufgrund der Intransparenz (Ananny/Crawford 2018; Burrell 2016; Zarsky 2016) von Algorithmen und der Komplexität der Zusammenhänge nicht eindeutig zurückverfolgen, wie es unter algorithmischer Mitwirkung zu bestimmten Entscheidungen gekommen ist. Somit können diese Personen nicht mehr die Enttäuschung normativer Erwartungen bearbeiten. Das heißt aber, dass die Kontrolle, die durch die Thematisierung von Verantwortlichkeit gegebenenfalls geleistet ist, nicht mehr greift und eine Fehlerdiskussion verunmöglicht wird. Die mit Verantwortlichkeit verbundene Kontrollfunktion setzt voraus, dass jene, die zur Rechenschaft gezogen werden, die fehlerhafte Entscheidung nachvollziehen können, so dass sie ihre Beteiligung daran erkennen: Sie können ihre Fehler besprechen, sich rechtfertigen und sich vornehmen, ihr Verhalten bei nächster Gelegenheit zu ändern. Bei der kontroversen Verantwortlichkeit ist dies unter Umständen aber nicht mehr möglich.

Aus dieser Schwächung der Kontrollfunktion kann folgen, dass etablierte Formen des Vertrauensaufbaus in Führungskräfte oder andere Mitarbeiter:innen in spezifischen Positionen erodieren. Wie erwähnt, ist formales Vertrauen unmittelbar mit Verantwortlichkeit verknüpft: Man kann Entscheider:innen vertrauen, dass sie Informationen korrekt verarbeiten und gewissenhaft Entscheidungen treffen, weil man sie für Fehler zur Rechenschaft ziehen kann. Dieser Mechanismus war zwar auch vor dem Einsatz von Algorithmen problematisch (Luhmann 1964, 180-182), weil schon immer die Möglichkeit bestand, die eigene Verantwortlichkeit zu verwischen oder diskursiv zu reduzieren. Die gesteigerte Intransparenz in komplexen Assemblagen verkompliziert jedoch die Lage und kann die Zuschreibung von Leistungen und Fehlern dermaßen erschweren, dass ein Vertrauensverlust in organisationales Personal und somit in die Organisation selbst ausgelöst werden kann. Auch dies kann eine Verunsicherung der Organisation bedeuten. Es ist zweifelhaft, ob Vertrauen in die Rechenmaschinen diesen Verlust tatsächlich kompensieren können. Kann in anderen Worten der häufig beschworene Glaube in die Objektivität, Geschwindigkeit und Vollkommenheit algorithmischer Kalkulationen das Vertrauen in das Personal ersetzen oder gar überflüssig machen? Empirische Arbeiten deuten darauf hin, dass dies nicht immer der Fall ist (Lee 2018). Eher lassen sich Versuche beobachten, auf IT-Dienstleister:innen und spezialisierte Rechtsberater:innen zu rekurrieren, die die Vertrauenswürdigkeit der Maschine affirmieren könnten.

Die Diskrepanz zwischen Verantwortung und Verantwortlichkeit kann aber auch positive Folgen haben. Insbesondere kann sie Raum für Innovation schaffen. Auf den ersten Blick scheint es, dass die Beteiligung von Algorithmen an Entscheidungsprozessen Kreativitätsspielräume dort schließen kann, wo statt Entscheidungen Kalkulationen eingesetzt werden. Algorithmische Analysen basieren auf Daten aus vergangenen Erfahrungen und auf der Basis der Vergangenheit bieten sie Analysen oder gar Prognosen für die Zukunft. Dies kann die Innovationsfähigkeit von Organisationen reduzieren und vor allem dann problematisch werden, wenn auf die Beteiligung von Personal bei schwierigen Entscheidungen verzichtet wird. Das Personal kann sich nicht mehr in der Auseinandersetzung mit praktischen Fällen weiterentwickeln und neues Erfahrungswissen sammeln (Faraj et al. 2018). Jedoch eröffnet die Automatisierung von einfachen Prozessen die Möglichkeit, kreative Potenziale des Personals freizusetzen, insofern Mitarbeiter:innen von langwierigen Routineaufgaben befreit werden. Diesen bekannten Argumenten fügt unsere Analyse der Diskrepanz von Verantwortung und Verantwortlichkeit hinzu, dass gerade die vieldiskutierten Schwierigkeiten der Zuschreibung von Verantwortung und die Intransparenz von Assemblagen im Hinblick auf Innovationen von Vorteil sein können. Denn gerade diese Bedingungen eröffnen neue Möglichkeiten informeller strategischer Spiele, die nicht nur organisationsinterne, sondern auch organisationsexterne Akteure einbeziehen können. Wenn man nicht unmittelbar die Folgen für die eigenen Fehler tragen muss, neigt man dazu, mehr zu wagen (Luhmann 1964, 187). Die positive Kehrseite einer problematischen Zuschreibung von Verantwortlichkeit kann somit als Schutz für riskante Innovationsprozesse dienen.

Ob und wie diese Spielräume für Innovation genutzt werden, hängt von den organisationalen Kontexten ab. Die Aura von Objektivität (Turner Lee 2018), die Algorithmen umgibt, der Glaube also, dass sie nicht nur präzise, sondern auch objektive Analyseinstrumente seien, kann auch zu Konformität führen und Innovation blockieren. Wenn Algorithmen Entscheidungen nicht ersetzen, sondern nur unterstützen, bleibt die Verantwortlichkeit an bestimmten, mit Personen besetzten Stellen verortet (bspw. Richter:innen und Polizist:innen). Entscheider:innen können Verantwortlichkeit kaum auf den Algorithmus abwälzen, der sie lediglich unterstützt, weil deren Stellen – mit ihren Aufgaben, ihrer organisationalen Einbettung und ihrer persönlichen Besetzung – eindeutig mit einem gewissen Ermessenspielraum ausgestattet sind. Setzen sich Entscheider:innen allerdings über algorithmisch prognostizierte Ereignisse hinweg, treffen sie eine explizite Entscheidung *gegen* die errechneten Resultate, steht die Verantwortlichkeit der Person im Vordergrund. Solche Entscheidungen erfordern wesentlich mehr Rechtfertigungsaufwand (Büchner/Dosdall 2021, 350). Dies erzeugt einen gewissen Druck,

algorithmenkonforme Entscheidungen zu treffen – eine Folge, die bereits im Fall von Sozialarbeiter:innen beobachtet wurde (Ranerup/Hinriksen 2019).

6 Schluss

Wir haben gezeigt, inwiefern datenintensive Algorithmen Organisationen betreffen und einige ihrer Folgen skizziert. Die Unterscheidung zwischen Verantwortung und Verantwortlichkeit hat uns ermöglicht, aktuelle Studien in ein anderes Licht zu rücken. Einerseits konnten wir schwerwiegende Verschiebungen und sogar eine Zerlegung von Verantwortung feststellen, andererseits veränderte Strategien der Rechenschaftspflicht sowie des Verwischens und Versteckens von Verantwortung herausarbeiten, die Verantwortlichkeit in höchstem Maße kontrovers werden lassen. Diese Veränderungen verkomplizieren das Verhältnis von Verantwortung und Verantwortlichkeit. Vor allem in dieser Hinsicht soll unsere Perspektive als Anregung für empirische Studien dienen, die unseren Vermutungen zur Minderung der Kontrollfunktion der Verantwortlichkeit, zum Risiko eines Vertrauensverlusts sowie zu den Chancen und Grenzen der Innovationsfähigkeit von algorithmisch unterstützten Organisationen nachgehen.

Unsere Perspektive betont die soziale Konstruktion von Verantwortung und Verantwortlichkeit. Dies bedeutet, dass Organisationen im Vergleich zu anderen sozialen Kontexten Verantwortung und Verantwortlichkeit spezifisch konstruieren. Weiterführende empirische Studien könnten untersuchen, inwiefern Organisationen Chancen nutzen oder Risiken erleiden, zum Beispiel, inwieweit die Illusion der Objektivität von Algorithmen in spezifischen Kontexten wirkt oder wo sich Vertrauensverluste feststellen lassen. Dabei sollte berücksichtigt werden, dass Verantwortung und Verantwortlichkeit nicht für alle Organisationstypen denselben Stellenwert haben. So müssen, um auf ein Beispiel von Luhmann zurückzugreifen, Kampftruppen eine verantwortliche Informationsverarbeitung liefern, was für Sportvereine weniger gilt (Luhmann 1964, 176). Erstere haben aber angesichts der Komplexität ihrer Aufgaben zugleich mehr Bedarf an Algorithmen.

Unsere Analyse legt auch Folgen für das Management offen. In der Literatur werden angesichts der Intransparenz von Algorithmen Forderungen nach moralischen und ethischen Verpflichtungen laut, die Algorithmen, aber auch die dahinterstehenden Organisationen erfüllen sollten. Andere haben schon darauf verwiesen, dass Intransparenz nicht immer aufzulösen ist und dass mehr Transparenz nicht immer zu mehr Verantwortlichkeit in Assemblagen beiträgt (Ananny/Crawford 2018). Wir zeigen zudem, dass die Verkomplizierung des Verhältnisses zwischen Verantwortung und Verantwortlichkeit zahlreiche Folgen haben kann. Wir plädieren daher dafür, dass auf der Suche nach neuen Modellen organisationaler Verantwortung und

Verantwortlichkeit nicht nur die Frage der eindeutigen Zuschreibung von Schuld und Leistung gestellt wird, sondern auch andere Aspekte berücksichtigt werden. Insbesondere müsste die kontroverse Verantwortlichkeit mit der Zerlegung von Verantwortung und den einhergehenden Veränderungen der Stellenstruktur verknüpft werden. Darüber hinaus sollten den Folgen dieser Verkomplizierung nachgegangen werden: Es muss gefragt werden, inwiefern unter diesen Umständen Vertrauen und Innovationsfähigkeit erodieren können und wie diese (wieder) hergestellt werden können.

ZWEI QUELLEN ENTFERNT ZWECKS ANONYMISIERUNG

Literatur

- Ananny, Mike/Crawford, Kate (2018): Seeing without knowing: Limitations of the Transparency Ideal and its Application to Algorithmic Accountability. *New Media & Society* 20 (3), 973-989. DOI: 10.1177/1461444816676645.
- Ballestrem, Johannes Graf/Bär, Ulrike/Gausling, Tina/Hack, Sebastian/Oelffen, Sabine von (2020): *Künstliche Intelligenz. Rechtsgrundlagen und Strategien in der Praxis*. Wiesbaden: Springer Gabler.
- Bauer, Susanne/Heinemann, Torsten/Lemke, Thomas (Hrsg.) (2017): *Science and Technology Studies. Klassische Positionen und aktuelle Perspektiven*. Berlin: Suhrkamp.
- Beckers, Anna/Teubner, Gunther (2021): *Three Liability Regimes for Artificial Intelligence. Algorithmic Actants, Hybrids, Crowds*. Oxford: Hart.
- Berthod, Olivier/Müller-Seitz, Gordon (2018): Making Sense in Pitch Darkness: An Exploration of the Sociomateriality of Sensemaking in Crises. *Journal of Management Inquiry* 27 (1), 52-68. DOI: 10.1177/1056492616686425.
- Besio, Cristina (2014): Transforming Risks into Moral Issues in Organizations, in: Christoph Luetge/Johanna Jauernig (Hrsg.), *Business Ethics and Risk Management* 43. Dordrecht: Springer Netherlands, 71-84.
- Besio, Cristina/Fedtke, Cornelia/Pronzini, Andrea (2022): Veränderungen organisationaler Legitimationsmuster durch Digitalisierung, in: Corinna Onnen/Rita Stein-Redent/Birgit Blättel-Mink/Torsten Noack/Michael Opielka/Katrin Späte (Hrsg.): *Organisationen in Zeiten der Digitalisierung*. Wiesbaden: Springer VS. [im Erscheinem]
- Beverungen, Armin (2017): Algorithmisches Management, in: Timon Beyes/Jörg Metelmann/Claus Pias (Hrsg.), *Nach der Revolution. Ein Brevier digitaler Kulturen*. Berlin: Tempus Corporate, 51-64.

- Bonnefon, Jean-François/Shariff, Azim/Rahwan, Iyad (2016): The Social Dilemma of Autonomous Vehicles. *Science* 352 (6293), 1573-1576. DOI: 10.1126/science.aaf2654.
- Brayne, Sarah (2017): Big Data Surveillance: The Case of Policing. *American Sociological Review* 82 (5), 977-1008. DOI: 10.1177/0003122417725865.
- Brunsson, Nils (1989): *The Organization of Hypocrisy. Talk, Decisions and Actions in Organizations*. Chichester: Wiley.
- Büchner, Stefanie (2018): Zum Verhältnis von Digitalisierung und Organisation. *Zeitschrift für Soziologie* 47 (5), 332-348. DOI: 10.1515/zfsoz-2018-0121.
- Büchner, Stefanie/Dosdall, Henrik (2021): Organisation und Algorithmus. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 73, 333-357. DOI: 10.1007/s11577-021-00752-0.
- Burrell, Jenna (2016): How the Machine ‚Thinks‘: Understanding Opacity in Machine Learning Algorithms. *Big Data & Society* 3 (1) [online first]. DOI: 10.1177/2053951715622512.
- Cerulo, Karen A. (2009): Nonhumans in Social Interaction. *Annual Review of Sociology* 35 (1), 531-552. DOI: 10.1146/annurev-soc-070308-120008.
- Cheng, Lu/Varshney, Kush R./Liu, Huan (2021): Socially Responsible AI Algorithms: Issues, Purposes, and Challenges. Online unter arxiv.org/pdf/2101.02032.
- Christin, Angèle/Rosenblat, Alex/Boyd, Danah (2015): Courts and Predictive Algorithms, in: Danah Boyd/Corrine Yu/Harlan Yu (Hrsg.), *Data & Civil Rights: A New Era of Policing and Justice*. Washington, D.C. Data & Society Research Institute. Online unter law.nyu.edu/sites/default/files/upload_documents/Angele%20Christin.pdf.
- Constantiou, Ioanna D./Kallinikos, Jannis (2015): New Games, New Rules: Big Data and the Changing Context of Strategy. *Journal of Information Technology* 30 (1), 44-57. DOI: 10.1057/jit.2014.17.
- Daipha, Phaedra (2015): From Bricolage to Collage: The Making of Decisions at a Weather Forecast Office. *Sociological Forum* 30 (3), 787-808. DOI: 10.1111/socf.12192.
- Debatin, Bernhard (2016): Verantwortung. Grundbegriffe der Kommunikations- und Medienethik (Teil 3). *Communicatio Socialis* 49 (1), 68-73. Online unter ejournal.communicatio-socialis.de/index.php/cc/article/view/1170/1168, zuletzt geprüft am 08.10.2021.
- Diakopoulos, Nicholas (2015): Algorithmic Accountability. *Digital Journalism* 3 (3), 398-415. DOI: 10.1080/21670811.2014.976411.
- Elish, Madeleine Claire (2016): Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction (Pre-Print). *Engaging Science, Technology, and Society*. SSRN Journal. DOI: 10.2139/ssrn.2757236.

- Esposito, Elena (2017): Artificial Communication? The Production of Contingency by Algorithms. *Zeitschrift für Soziologie* 46 (4), 249-265. DOI: 10.1515/zfsoz-2017-1014.
- Eubanks, Virginia (2018): Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press.
- Faraj, Samer/Pachidi, Stella/Sayegh, Karla (2018): Working and Organizing in the Age of the Learning Algorithm. *Information and Organization* 28 (1), 62-70. DOI: 10.1016/j.infoandorg.2018.02.005.
- Feier, Till/Gogoll, Jan/Uhl, Matthias (2021): Hiding Behind Machines: When Blame Is Shifted to Artificial Agents. Online unter arxiv.org/pdf/2101.11465.
- Foerster, Heinz von (1993): *KybernEthik*. Berlin: Merve.
- Gresser, Uwe (2018): *Hochfrequenzhandel*. Wiesbaden: Springer Fachmedien.
- Grush, Loren (2015): Google Engineer Apologizes after Photos App Tags Two Black People as Gorillas. *The Verge*, 02.07.2015. Online unter theverge.com/2015/7/1/8880363/google-apologizes-photos-app-tags-two-black-people-gorillas, zuletzt geprüft am 08.10.2021.
- Galdi, Francesco/Cordella, Antonio (2021): Artificial Intelligence and Decision-Making: the question of Accountability. *Proceedings of the 54th Hawaii International Conference on System Sciences*. DOI: 10.24251/HICSS.2021.281.
- Heiland, Heiner (2018): Algorithmus = Logik + Kontrolle. *Algorithmisches Management und die Kontrolle der einfachen Arbeit*, in: Daniel Houben/Bianca Prietl (Hrsg.), *Datengesellschaft*. Bielefeld: transcript, 233-252.
- Karafilidis, Athanasios/Weidner, Robert (2015): *Technische Unterstützungssysteme*. Berlin: Springer. Online unter openhsu.ub.hsu-hh.de/handle/10.24405/10881.
- Lee, Min Kyung (2018): Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management. *Big Data & Society* 5 (1), 1-16. DOI: 10.1177/2053951718756684.
- Lewis, Sarah (2019): The Racial Bias Built Into Photography. *The New York Times*, 25.04.2019. Online unter nytimes.com/2019/04/25/lens/sarah-lewis-racial-bias-photography.html, zuletzt geprüft am 08.10.2021.
- Liao, Wenguo/Liao, Guangping (2020): Research on Accountability Mechanism of Big Data Algorithm Decision, in: 2020 International Conference on Computer Science and Engineering Technology (CSET2020), 23.-24.06.2020. Online unter proceedings-online.com/proceedings_series/SH-INFORMATION/CSET2020/emsse07501.pdf, zuletzt geprüft am 23.05.2021.

- Loh, Janina (2019): Verantwortung und Roboterethik, in: Matthias Rath/Friedrich Krotz/Matthias Karmasin (Hrsg.), *Maschinenethik. Normative Grenzen autonomer Systeme*. Wiesbaden: Springer VS, 91-105.
- Luhmann, Niklas (1964): *Funktionen und Folgen formaler Organisation*. Berlin: Duncker & Humblot.
- Luhmann, Niklas (1966): *Recht und Automation in der öffentlichen Verwaltung. Eine verwaltungswissenschaftliche Untersuchung*. Berlin: Duncker & Humblot.
- Luhmann, Niklas (2000): *Organisation und Entscheidung*. Opladen: Westdeutscher Verlag.
- March, James G./Simon, Herbert A. (1958): *Organizations*. New York: Wiley.
- Martin, Kirsten (2019): Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics* 160 (4), 835-850. DOI: 10.1007/s10551-018-3921-3.
- Matthias, Andreas (2004): The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology* 6 (3), 175-183. DOI: 10.1007/s10676-004-3422-1.
- McAfee, Andrew/Brynjolfsson, Erik (2012): *Big Data: The Management Revolution*. Hg. v. Harvard Business Review. Online unter hbr.org/2012/10/big-data-the-management-revolution, zuletzt aktualisiert am 08.10.2014, zuletzt geprüft am 20.05.2021.
- McDermid, John (2020): Boeing 737 Max: why was it grounded, what has been fixed and is it enough? *The Conversation*, 27.11.2020. Online unter theconversation.com/boeing-737-max-why-was-it-grounded-what-has-been-fixed-and-is-it-enough-150688, zuletzt geprüft am 26.03.2022.
- Meijer, Albert/Grimmelikhuijsen, Stephan G. (2020): Responsible and Accountable Algorithmization. How to Generate Citizen Trust in Governmental Usage of Algorithms, in: Marc Schuilenburg/Rik Peeters (Hrsg.), *The Algorithmic Society. Technology, Power, and Knowledge*. London, New York: Routledge, 53-66.
- Mittelstadt, Brent Daniel/Allo, Patrick/Taddeo, Mariarosaria/Wachter, Sandra/Floridi, Luciano (2016): The Ethics of Algorithms: Mapping the Debate. *Big Data & Society* 3 (2), 1-21. DOI: 10.1177/2053951716679679.
- Neyland, Daniel (2016): Bearing Account-Able Witness to the Ethical Algorithmic System. *Science, Technology, & Human Values* 41 (1), 50-76. DOI: 10.1177/0162243915598056.
- Prinzing, Marlis (2019): Auf dem Weg in eine Diktatur der Wahrscheinlichkeit? Fragen nach der Verantwortung beim Einsatz von Prognosesoftware, in: Carsten Ochs/Michael Friedewald/Thomas Hess/Jörn Lamla (Hrsg.), *Die Zukunft der Datenökonomie. Zwischen Geschäftsmodell, Kollektivgut und Verbraucherschutz*. Wiesbaden: Springer VS, 269-284.

- Ranerup, Agneta/Henriksen, Helle Zinner (2019): Value Positions Viewed Through the Lens of Automated Decision-Making: The Case of Social Services. *Government Information Quarterly* 36 (4), 1-13. DOI: 10.1016/j.giq.2019.05.004.
- Reddy, Elizabeth/Cakici, Baki/Ballestero, Andrea (2019): Beyond Mystery: Putting Algorithmic Accountability in Context. *Big Data & Society* 6 (1) [online first]. DOI: 10.1177/2053951719826856.
- Reimers, Nils/Gurevych, Iryna (2018): Why Comparing Single Performance Scores Does Not Allow to Draw Conclusions About Machine Learning Approaches. Online unter arxiv.org/pdf/1803.09578.pdf, zuletzt geprüft am 15.03.2021.
- Schwarting, Rena (2015): Hochfrequenzhandel zwischen Entscheidungsautomation und Entscheidungsautonomie, in: Maja Apelt/Konstanze Senge (Hrsg.), *Organisation und Unsicherheit*. Wiesbaden: Springer VS, 159-174.
- Shackelford, Scott/Raymond, Anjanette (2014): Building the Virtual Courthouse: Ethical Considerations for Design, Implementation, and Regulation in the World of ODR. *SSRN Journal*. Kelley School of Business Research Paper No. 2014-10. DOI: 10.2139/ssrn.2387912.
- Sparrow, Robert (2007): Killer Robots. *Journal of Applied Philosophy* 24 (1), 62-77. DOI: 10.1111/j.1468-5930.2007.00346.x.
- Thiemann, Daniel (2021): Führen und Entscheiden in modernen Arbeitswelten – Was verändert die Digitalisierung?, in: Stephan Kaiser/Arjan Kozica/Frank Böhringer/Josef Wissinger (Hrsg.), *Digitale Arbeitswelt*. Wiesbaden: Springer Fachmedien, 143-162.
- Turner Lee, Nicol (2018): Detecting Racial Bias in Algorithms and Machine Learning. *Journal of Information, Communication and Ethics in Society* 16 (3), 252-260. DOI: 10.1108/JICES-06-2018-0056.
- Vincent, Gina M./Viljoen, Jodi L. (2020): Racist Algorithms or Systemic Problems? Risk Assessments and Racial Disparities. *Criminal Justice and Behavior* 47 (12), 1576-1584. DOI: 10.1177/0093854820954501.
- Vincent, James (2018): Google ‚fixed‘ its Racist Algorithm by Removing Gorillas from its Image-Labeling Tech. *The Verge*, 12.01.2018. Online unter theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai, zuletzt geprüft am 08.10.2021.
- Wynants, Laure/Calster, Ben van/Collins, Gary S./Riley, Richard D./Heinze, Georg/Schuit, Ewoud et al. (2020): Prediction Models for Diagnosis and Prognosis of Covid-19: Systematic Review and Critical Appraisal. *BMJ* 369 [online first]. DOI: 10.1136/bmj.m1328.

Zarsky, Tal (2016): The Trouble with Algorithmic Decisions. *Science, Technology, & Human Values* 41 (1), 118-132. DOI: [10.1177/0162243915605575](https://doi.org/10.1177/0162243915605575).