

Dear reader,

This is an author-produced version of an article published in *Theologie und Glaube* 114 (2024). It agrees with the manuscript submitted by the author for publication but does not include the final publisher's layout or pagination.

Original publication:

Klöcker, Katharina

Expandierende Verantwortungslücken? Algorithmisierte Entscheidungsprozesse in theologisch-ethischer Reflexion

in: *Theologie und Glaube* 114 (2024), pp. 400–409

Münster: Aschendorff Verlag 2024

Access to the published version may require subscription.

Published in accordance with the policy of Aschendorff Verlag: <https://www.aschendorff-buchverlag.de/openaccesstop>

Your IxTheo team

Liebe*r Leser*in,

dies ist eine von dem/der Autor*in zur Verfügung gestellte Manuskriptversion eines Aufsatzes, der in *Freiburger Zeitschrift für Theologie und Glaube* 114 (2024) erschienen ist. Der Text stimmt mit dem Manuskript überein, das der/die Autor*in zur Veröffentlichung eingereicht hat, enthält jedoch *nicht* das Layout des Verlags oder die endgültige Seitenzählung.

Originalpublikation:

Klöcker, Katharina

Expandierende Verantwortungslücken? Algorithmisierte Entscheidungsprozesse in theologisch-ethischer Reflexion

in: *Theologie und Glaube* 114 (2024), S. 400–409

Münster: Aschendorff Verlag 2024

Die Verlagsversion ist möglicherweise nur gegen Bezahlung zugänglich.

Diese Manuskriptversion wird im Einklang mit der Policy des Aschendorff Verlags publiziert: <https://www.aschendorff-buchverlag.de/openaccesstop>

Ihr IxTheo-Team

Expandierende Verantwortungslücken? Algorithmisierte Entscheidungsprozesse in theologisch-ethischer Reflexion

Katharina Klöcker

Abstract

Der Text nimmt unterschiedliche Formen algorithmisierter Entscheidungsprozesse in den Blick. Während bei algorithmenbasierten Entscheidungen die Verantwortungszuschreibungen lediglich komplizierter werden, entstehen in der Mensch-Maschine-Interaktion durch selbstlernende Algorithmen immer mehr Verantwortungslücken. Diese werden in diesem Text als mögliche Gefährdungen der Demokratie in theologisch-ethischer Perspektive reflektiert.

The text looks at different forms of algorithmic-decision-making-processes. While the attribution of responsibility merely becomes more complicated in algorithm-based decisions, self-learning algorithms lead to a growing number of vacuums in terms of who is responsible in human-machine interaction. This text reflects on these challenges as potential threats to democracy from a theological-ethical perspective.

1. Wider eine wirklichkeitsferne Ethik

So wirklichkeitsfern vielen Menschen die kirchliche Morallehre hinsichtlich bestimmter Themen bis heute auch erscheinen mag - die theologische Ethik hat in den vergangenen Jahrzehnten viel dafür getan, sich diesen Vorwurf der Wirklichkeitsferne nicht einzuhandeln. Die postvatikanische Abkehr von einem die katholische Moraltheologie lange Zeit prägenden naturrechtlich-essentialistischen Paradigma ging einher mit einer expliziten Hinwendung zu den empirischen Wissenschaften, die mit einer inter- und transdisziplinären Perspektivierung korrespondierte und zu einer wissenschaftstheoretischen Neuausrichtung des Fachs führte. Das Ringen um eine adäquate Verhältnisbestimmung der theologischen Ethik zu den Erkenntnissen der Human- und Sozialwissenschaften ist bis heute zentraler Bestandteil des im Fach andauernden Selbstvergewisserungsprozesses.

Das Nachdenken darüber, welche Rolle empirische Einsichten in der theologischen Ethik jenseits eines Seins-Sollen-Fehlschluss spielen können, ist deshalb so wichtig, weil nur eine auf die *Wirklichkeit* bezogene, eine sich ihren Kontexten zuwendende Ethik prinzipiell dazu in der Lage ist, Orientierung zu stiften. Es gilt dabei, dass etwas, das empirischen Erkenntnissen widerspricht, nicht Gegenstand normativ-ethischer Aussagen sein kann. Jede Ethik, die dies außer Acht lässt, setzt ihre Plausibilität aufs Spiel und wird in den gesellschaftlichen Debatten ihrer Zeit Zaungast bleiben.

Doch nicht nur der Wirklichkeitsbezug in Form empirischer Einsichten ist Bezugspunkt normativer Anstrengungen. Auch die Ethik übernimmt hinsichtlich empirischer Erkenntnisse eine wichtige Funktion. Sie thematisiert die normativ wirksamen, aber oftmals implizit bleibenden epistemologischen Vorannahmen empirischer Wissenschaften und verhilft dadurch der Einsicht zur Geltung, dass auch die vermeintlich objektiven empirischen Einsichten ihrerseits abhängig von bestimmten Deutungshorizonten sind.

Dieses hier skizzierte Verhältnis einer gegenseitigen Verwiesenheit von Ethik und Empirie¹ lässt sich noch um den Aspekt ergänzen, dass die normative Geltung grundlegender ethischer Prinzipien zwar

¹ Ausführlicher dazu z. B. W. Schaupp (Hg.), Ethik und Empirie. Gegenwärtige Herausforderungen für Moraltheologie und Sozialethik, Freiburg u.a. 2014 (Studien zur theologischen Ethik 142); E. Kos, Ethik und Empirie. Die Suche nach einer angemessenen Zuordnung in der römisch-katholischen Moraltheologie, in: Una

nicht von empirischen Erkenntnissen abhängig gemacht werden kann, diese Prinzipien sich aber – so hat es der niederländische Medizinethiker Bert Musschenga als generalistische Form Empirischer Ethik beschrieben² – in der Wirklichkeit zu bewähren haben und möglicherweise aufgrund empirischer Einsichten auch modifiziert werden müssen.³ Das heißt, Prinzipien und moralische Urteile müssen ihre Praktikabilität in konkreten Situationen unter Beweis stellen und sich bei einem Nicht-Bestehen dieses Realitäts-Checks einem Veränderungs- und Korrekturprozess unterziehen.

Ausgehend von dieser methodischen Vorbemerkung wenden sich die folgenden Überlegungen einer gegenwärtig mit der Digitalisierung der Gesellschaft einhergehenden Problematik zu, die auch theologische Ethik aus oben genannten Gründen herausfordert: Im Fokus steht die Frage, welche Auswirkungen es hat, wenn zunehmend Algorithmen auch bei moralisch relevanten Entscheidungen mit im Spiel sind. Wer übernimmt die Verantwortung, wenn KI-Systeme in Entscheidungen involviert sind oder diese vielleicht sogar vollkommen automatisiert treffen?

Eines steht fest: Antworten darauf, wie in noch näher zu spezifizierenden Mensch-Maschine-Interaktionen Verantwortung zugeschrieben wird, haben weitreichende Folgen für das Zusammenleben in digitalen Gesellschaften. Nicht zuletzt stehen Fragen nach Verantwortlichkeiten immer auch in einem engen Zusammenhang mit Entwicklungen, die die Demokratie betreffen, weil diese maßgeblich von der Bereitschaft zur Übernahme von Verantwortung lebt. „[K]eine Gesellschaftsform ist so auf Verantwortungsübernahme durch die Individuen, die gleichberechtigten und gleichverpflichteten Bürger der Zivilgesellschaft angewiesen wie die demokratische Gesellschaft. Demokratie verpflichtet die Individuen zu verantwortlichem Handeln, und verantwortliches Handeln muss gelernt werden.“⁴ Entwicklungen, die Verantwortungszuschreibungen erschweren oder verunmöglichen, drohen damit zugleich auch demokratische Strukturen auszuhöhlen.

Solche Alarmzeichen für Demokratiegefährdungen fordern insbesondere auch theologische Ethik, die sich explizit als Verantwortungsethik versteht, heraus. Will sie sich an den Debatten über Digitalität und Demokratie beteiligen, wird auch sie, wie die folgenden Überlegungen ein Stück weit entfalten wollen, im Angesicht der Algorithmisierung unserer Lebenswelt über Verantwortung als zentrale ethische Kategorie neu nachzudenken haben und daraus konzeptionelle Konsequenzen ziehen müssen.

2. Algorithmisierte Entscheidungsprozesse

Die Digitalisierung verändert menschliches Wahrnehmen, Denken und Handeln tiefgreifend. In den Transformationsprozessen digitaler Gesellschaften gibt der Algorithmus⁵ den Takt an. Er ist das entscheidende Bindeglied zwischen Daten und Digitalisierung. In Form von Softwareprogrammen ermöglicht der Algorithmus durch Selektion, Sortierung und Bewertung von Daten die effiziente Automatisierung von Prozessen aller Art. Vieles spricht dafür, dass der Algorithmus auch die „DNA der digitalen Zukunft“⁶ sein wird. Bereits jetzt ist Algorithmisierung ubiquitär und aus vielen Bereichen

Sancta 69 (2014) 195-206; T. Moos, Ethik und Empirie in der Theologie, in: Zeitschrift für Evangelische Ethik 67 (2023) 185-199.

² Vgl. B. Musschenga, Was ist empirische Ethik?, in: Ethik in der Medizin 21 (2009) 187-199.

³ Vgl. Moos, Ethik und Empirie (s. Anm. 1), 188.

⁴ W. Edelstein, Erziehung zu Demokratie und Verantwortung. Einführungsvortrag zum Workshop „Demokratische Schule – Schule in der Demokratie“, Heinrich-Böll-Stiftung, Berlin, 5. Dezember 2007, online verfügbar unter: www.boell.de (Zugriff 23.05.2024).

⁵ Algorithmus bezeichnet eine präzise, formalisierte Abfolge von Anweisungen, die ein mathematisches Problem lösen sollen. „Mit Computern geht das Maschinenhafte von Algorithmen apparativ in Erfüllung.“ T. A. Heilmann, Algorithmus, in: K. Liggieri/O. Müller (Hg.), Mensch-Maschine-Interaktion: Handbuch zu Geschichte - Kultur - Ethik, Stuttgart 2019, 229-231, 229.

⁶ M. Martini, Blackbox Algorithmus – Grundfragen einer Regulierung Künstlicher Intelligenz, Berlin 2019, 1.

des Lebens nicht mehr wegzudenken. Während diese Entwicklung in der Informatik und Mathematik längst Gegenstand intensiver Forschung ist, wenden sich zeitverzögert zunehmend auch die Geistes-, Kultur- und Sozialwissenschaften den durch Algorithmen und KI-Systemen virulent werdenden Fragen zu. Hier ist mittlerweile sogar von einem „algorithmic turn“⁷ die Rede.

Ein Aspekt, der aus ethischer Perspektive besonders virulent ist, betrifft die Entwicklung, dass in digitalen Gesellschaften in immer mehr Bereichen des Lebens immer häufiger Entscheidungen mit Hilfe von Algorithmen getroffen werden. Je nach Grad der Involviertheit von Algorithmen, wird von teilautomatisierten oder automatisierten Entscheidungen gesprochen. So zeigen KI-Systeme auf der Grundlage bisheriger Suchverläufe personalisierte Nachrichten oder Werbung auf dem Smartphone an, Algorithmen treffen auf Online-Partnerbörsen eine Auswahl potenzieller Partner(innen), berechnen für einen Stadtteil die Wahrscheinlichkeit von Einbrüchen, prognostizieren Wahlergebnisse oder nehmen Einfluss darauf, wer einen Kredit, eine künstliche Hüfte oder einen Kita-Platz bekommt. Das so genannte *Algorithmic-decision-making* (ADM) greift tief in das Leben ein und bestimmt nicht selten über Wohl und Wehe von Menschen.⁸ Die Frage, welche Macht den Algorithmen hinsichtlich welcher Entscheidungen zukommen sollte, und wie Chancen und Risiken automatisierter Prozesse abgewogen werden können, drängt sich auf. Letztlich geht es hier, ebenso wie in vielen anderen Digitalisierungskontexten auch, im Kern um die Frage nach dem Verhältnis von Mensch und Maschine.

Die Vorteile der auf Algorithmen basierenden Softwareanwendungen sind immens: (Teil)automatisierte Entscheidungsprozesse steigern Effizienz, ermöglichen Prognosen, nehmen Kosten-Nutzen-Bewertungen vor oder eröffnen neue Analyse- und Diagnosemöglichkeiten. Von den Vorteilen des Einsatzes von KI-Systemen profitieren sehr viele Menschen. Ein Verzicht auf Algorithmen scheint daher eigentlich nicht mehr vorstellbar. Hochtechnologisierte Gesellschaften würden in kürzester Zeit zumindest in manchen Bereichen schlicht kollabieren. Wie hoch der Grad der Abhängigkeit ist, ist für viele Menschen allerdings kaum erkennbar, da algorithmisierte Prozesse an unzähligen Stellen unbemerkt im Hintergrund ablaufen. „Algorithmen arbeiten verborgen im Maschinenraum des Cyberspace und nehmen von dort aus immer zielsicherer Navigationsfunktionen wahr, die bislang dem Menschen als Leiter der Kommandobrücke vorbehalten waren. Die Auswertung vollzieht sich jenseits des eigenen Wahrnehmungshorizonts – auf der Grundlage der IP-Adresse, angeklickter Symbole oder der Reaktionsschnelligkeit – als Abfallprodukt eines eigenen Nutzungsvorteils.“⁹

Nur langsam rückt hingegen ins öffentliche Bewusstsein, dass die Algorithmisierung auch mit hohen Risiken verbunden ist, sowohl für den Einzelnen als auch für die Gesellschaft insgesamt. Bei einer repräsentativen Umfrage der Bertelsmann-Stiftung¹⁰ aus dem Jahr 2022 gaben vier von fünf Befragten an, weder bei sich noch bei anderen Menschen Benachteiligungen durch algorithmenbasierte Entscheidungen bemerkt zu haben. Tatsächlich erzeugt der Einsatz von Algorithmen jedoch durchaus immer wieder Gerechtigkeitsprobleme.¹¹ Die zugrundeliegenden Software-Programmierungen bleiben intransparent hinsichtlich der ihnen inhärenten Rationalitäten

⁷ W. Uricchio, The algorithmic turn: photosynth, augmented reality and the changing implications of the image, in: *Visual Studies* 26 (2011) 25-35.

⁸ Beispielhaft zeigt dies für den Gesundheitsbereich etwa die Stellungnahme der Zentralen Kommission zur Wahrung ethischer Grundsätze in der Medizin und ihren Grenzgebieten (Zentrale Ethikkommission) bei der Bundesärztekammer „Entscheidungsunterstützung ärztlicher Tätigkeit durch Künstliche Intelligenz“, in: *Deutsches Ärzteblatt* 118 (2021) Heft 33-34, DOI: 10.3238/arztebl.zeko_sn_cdss_2021 [Zugriff: 24.05.2024].

⁹ Martini, *Blackbox Algorithmus* (s. Anm. 5), 7.

¹⁰ Vgl. Bertelsmann-Stiftung, Was Deutschland über Algorithmen und Künstliche Intelligenz weiß und denkt, Gütersloh 2022, online unter www.bertelsmann-stiftung.de [Zugriff: 24.05.2024].

¹¹ Vgl. z.B. K. Klöcker, *Autoritäre Algorithmen*, in: *Herder Korrespondenz* 75 (2021) Heft 7, 37-39.

und Wertvorstellungen. Bei Entscheidungen, die mit Hilfe von Algorithmen getroffen werden, kommt es mittlerweile nicht selten dazu, dass Menschen aufgrund bestimmter Merkmale, die mit Vorurteilen verbunden sind, diskriminiert werden. Die Allgegenwärtigkeit solcher *algorithm bias* ist mittlerweile wissenschaftlich vielfach belegt.¹²

3. Frage nach der Zuschreibung von Verantwortung

Spätestens an diesem Punkt rückt die Frage ins Zentrum, wer für Benachteiligungen und Ungerechtigkeiten, die durch ADM entstehen, verantwortlich zu machen ist. Wer ist also in (teil)automatisierten Entscheidungsprozessen Akteur und kann bzw. muss damit auch für die Folgen seiner Handlungen bezüglich geltender Normen einstehen? In rechtlicher Perspektive ist zu fragen, wer prinzipiell haftungsfähig ist.¹³ Geht man von einem normativen Verantwortungskonzept aus, das zwischen Verantwortungssubjekt (wer), Verantwortungsobjekt (wofür), Verantwortungsinstanz (gegenüber wem) und der Begründung (warum) differenziert¹⁴, dann zeigt sich, dass in Mensch-Maschine-Interaktionen die Frage, wer wofür verantwortlich zu machen ist, nicht mehr ohne weiteres klar zu beantworten ist. Bereits die Tatsache, dass von „intelligenten“ Systemen gesprochen wird, legt nahe, dass die konventionelle Rollenverteilung mit dem Menschen als Verantwortungssubjekt und der Maschine als Verantwortungsobjekt zumindest begründungsbedürftig wird.

Wenn KI-Systeme, also „hybride sozio-technische Konstellationen“¹⁵ etwas *entscheiden*, bedeutet das dann, dass Maschinen im Zuge der Digitalisierung in einen Bereich vordringen und dort in Aktion treten, auf den bislang der Mensch einen Exklusivanspruch zu haben schien, dass Maschinen also auch moralisch entscheiden, was implizieren würde, dass sie sich in Freiheit für oder gegen eine Handlung entscheiden könnten? Lässt sich mit Blick auf KI-Systeme zudem tatsächlich davon sprechen, dass diese nicht nur aufgrund eines vorgegebenen Algorithmus teils auch ethisch relevante Entscheidungen treffen, sondern dass diesen Vorgängen auch eine moralische Urteilskraft¹⁶ zugrunde liegt? Diese Fragen werden kontrovers diskutiert.¹⁷ Als Hauptargument gegen die Annahme, Maschinen könnten eigenständig moralische Prinzipien auf konkrete Situationen anwenden, wird darauf verwiesen, dass KI-Systeme weder über Rationalität noch über Autonomie oder ein Gewissen verfügten.¹⁸ Ebenfalls hervorgehoben wird, dass allem Handeln eine Intentionalität anhafte, und dass diese mit der Fähigkeit einhergehe, das Handeln und die ihm zugrundeliegenden Überzeugungen und Emotionen begründen zu können: „Dieses *Geben und Nehmen von Gründen* ist konstitutiv für Interaktion, Kommunikation und menschliche Handlung im Allgemeinen. Es kann nicht in ein auf Algorithmen basiertes Verhaltenssystem übersetzt werden.“¹⁹ Auch wenn Julian Nida-Rümelin in diesem Zitat betont, dass diese Fähigkeit zur Angabe von Gründen für eigenes Handeln prinzipiell auch nicht-menschlichen Lebewesen zukommen könne, so bleibe im Gegensatz zur Maschine doch der Mensch derjenige, der entscheide und der auch für die Folgen seines Handelns einzustehen habe.

¹² Vgl. etwa die Veröffentlichung der Agentur der Europäischen Union für Grundrechte (FRA) aus dem Jahr 2022: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf [Zugriff: 23.05.2024].

¹³ Vgl. L. Heidbrink, Definitionen und Voraussetzungen der Verantwortung, in: Ders./C. Langbehn/J. Loh (Hg.), Handbuch Verantwortung, Wiesbaden 2017, 3-33, 5.

¹⁴ Vgl. M. H. Werner, Verantwortung, in: A. Grunwald (Hg.), Handbuch Technikethik, Stuttgart 2013, 44-48, 45f.

¹⁵ F. Saurwein, Automatisierung, Algorithmen, Accountability. Eine Governance Perspektive, in: M. Rath/F. Krotz/M. Karmasin (Hg.), Maschinenethik. Normative Grenzen autonomer Systeme, Wiesbaden 2019, 35-56, 50.

¹⁶ Unter moralischer Urteilskraft wird die Fähigkeit verstanden, moralische Prinzipien und Regeln auf konkrete Situationen anzuwenden, unter Berücksichtigung der dabei für die konkrete moralische Entscheidungsfindung relevanten Aspekte.

¹⁷ Vgl. dazu J. Nida-Rümelin: Handlung, Technologie und Verantwortung, in: W. Pietsch/J. Wernecke/M. Ott (Hg.), Berechenbarkeit der Welt? Philosophie und Wissenschaft im Zeitalter von Big Data, Wiesbaden 2017, 497-513.

¹⁸ Vgl. etwa P. Kirchschräger, Ethisches Entscheiden, Baden Baden 2023, 41.

¹⁹ J. Nida-Rümelin, Handlung, Technologie und Verantwortung (s. Anm. 14), 500 (Hervorh. i. Orig.).

Der Mensch könne zwar Maschinen befähigen, innerhalb eines vorgegeben Rahmens von Regeln Entscheidungen in Form von Handlungsanweisungen zu treffen, doch bleibe er für die Programmierung entsprechender Software zuständig und somit auch Verantwortungsträger.

Bereits seit den 1970er Jahren wird unter dem Stichwort der Technikfolgenabschätzung breit diskutiert, wie individuelle und institutionelle Dimensionen von Verantwortung ineinandergreifen.²⁰ Unter den Vorzeichen der Digitalisierung wird daher ebenfalls zu eruieren sein, wie Verantwortung in algorithmisierten Entscheidungen zu verteilen ist. Dieses auch unter Berücksichtigung der Unterscheidung von Handlungsverantwortung und Regelverantwortung zu bestimmen, wird allerdings desto komplizierter, je länger und unübersichtlicher die „lange Kette der Verantwortlichkeiten“²¹ ist und je komplexer die soziotechnischen Systeme des digitalen Zeitalters sich entwickeln. Trotzdem aber gilt, dass Verantwortung auch in den hier fokussierten algorithmisierten Entscheidungsprozessen weiterhin menschlichen Akteuren zuschreibbar ist. Zwar verschärfen sich die Fragen. Doch sie stellen sich nicht radikal neu.

4. Regelbasierte oder selbstlernende KI-Systeme

Mit den bisherigen Ausführungen zu komplexer werdenden Verantwortungszuschreibungen ist die mit der Algorithmisierung sich verbindende Verantwortungsproblematik allerdings noch nicht ausreichend beschrieben. Eine Differenzierung im Hinblick auf unterschiedliche Arten von KI-Systemen eröffnet den Blick auf einen weiteren Problemhorizont, der sich hinsichtlich der Zuschreibung von Verantwortung auftut. Bislang wurden KI-Systeme in diesem Text als nicht nur grundsätzlich, sondern auch als ausschließlich regelbasiert arbeitend in dem Sinne behandelt, dass sie Daten auf der Grundlage vorgegebener Algorithmen verarbeiten. Und in der Tat liegen solche Systeme dem Großteil von algorithmisierten Entscheidungsprozessen zugrunde. Die von der Bundesregierung im Jahr 2018 eingesetzte Datenethikkommission etwa bezeichnet in ihrem Gutachten Entscheidungen, die auf solchen Systeme beruhen, als „algorithmienbasiert“ - in Abgrenzung von „algorithmiengetriebenen“ und „algorithmien determinierten“²² Entscheidungen. Mit Blick auf regelbasierte Prozesse wird in diesbezüglichen Debatten vielfach gefragt, welchen Anteil an moralisch möglicherweise nicht zu rechtfertigenden Entscheidungen Programmierer(innen), Nutzer(innen), Auftraggeber(innen) haben. Und diese Fragen sind – wie gezeigt wurde – aufgrund der *langen Ketten von Verantwortlichkeiten* nicht leicht zu beantworten. Aber es gilt auch: Die grundsätzliche Annahme, dass individuelle oder institutionelle Akteure in diesem Setting Verantwortungssubjekte und die Maschine Verantwortungsobjekt bleibt, wird nicht zur Disposition gestellt.

Dies beginnt sich erst zu ändern, wenn man *algorithmiengetriebene* und *algorithmien determinierte* Entscheidungen in den Blick nimmt. Denn es werden zunehmend KI-Systeme entwickelt, die eben nicht mehr ausschließlich auf vorgegebenen Rechenregeln basieren, sondern auf lernenden, sich eigenständig weiter entwickelnden Algorithmen. Solche selbstlernenden Systeme, mit denen (noch) „vornehmlich in Forschungslaboren“²³ großer Technologiefirmen experimentiert wird, verändern sich in einer hochdynamischen und ab einem gewissen Punkt für Menschen nicht mehr nachvollziehbaren Weise. Bereits jetzt sind zumindest (teil)autonome Systeme im Einsatz, bei denen sich der Algorithmus „vielschrittig, kontextuell und kollaborativ, meist in Trial-and-Error- und Copy-and-Paste-

²⁰ Vgl. etwa G. Ropohl, Neue Wege, die Technik zu verantworten, in: H. Lenk/G. Ropohl (Hg.), Technik und Ethik, Stuttgart ²1993, 149-176.

²¹ K. Zweig, Die KI war's. Von absurd bis tödlich: Die Tücken der künstlichen Intelligenz, München 2023, 149.

²² Vgl. Gutachten der Datenethikkommission der Bundesregierung (Kurzfassung), Berlin 2019, 17, abrufbar unter: www.datenethikkommission.de [Zugriff: 24.05.2024].

²³ A. Pentenrieder, Algorithmen im Alltag. Eine praxistheoretische Studie zum informierten Umgang mit Routenplanern, Frankfurt 2020, 33.

Verfahren aus vorhandenen Programmbibliotheken²⁴ entwickelt. Das Potenzial dieser auch als *Machine Learning* bezeichneten Vorgänge kann sich „insbesondere dort optimal entfalten, wo datenvolumenstarke Aufgaben auf der Grundlage einer großen Menge unstrukturierter Daten zu bewältigen sind. Die Erfolge zahlreicher Internetpioniere, die ein rasant wachsendes Unternehmen mit disruptiven Technologien aufgebaut haben, gründen zu einem guten Teil auf solchem maschinellen Lernen als Fortschrittmotor.“²⁵

Im Zuge des *Machine Learning* wird ein Grad an Automatisierung und Autonomie erreicht, der es nicht nur Adressat(innen) oder Nutzer(innen) von Anwendungen, sondern auch hochspezialisierten Entwickler(innen) von Computerprogrammen erschwert bis unmöglich macht, diese sich immer weiter verselbständigenden Prozesse noch nachzuvollziehen und damit letztlich auch, sie noch in irgendeiner Weise zu kontrollieren. Der vielfach vorgenommene Vergleich mit einer *Black Box* scheint daher gerechtfertigt. Es lässt sich nicht (mehr) überprüfen, nicht von Softwareentwickler(innen) und schon gar nicht von Laien, nach welchen Kriterien und Konstruktionslogiken Daten ausgewertet werden oder aber auch, welche Daten überhaupt zur Grundlage von Berechnungen herangezogen werden bzw. wurden.

5. Expandierende Verantwortungslücken

Zwischen KI-Systemen und Menschen entsteht durch die im vorangegangenen Abschnitt thematisierten Entwicklungen eine *distributed agency*²⁶, eine „distribuierte Handlungsträgerschaft“²⁷. So zeigt Valentin Rauer auf, wie Menschen etwa an Entscheidungen im Kontext digitaler Waffensysteme „immer noch beteiligt“²⁸ seien, es sich aber bereits abzeichne, dass sie „zunehmend aus dem Entscheidungskreislauf herausgenommen“²⁹ würden. Die Opazität selbstlernender Systeme begünstige die Entstehung „subjektloser Prozesse“, die letztlich die Ursache für eine „diffuse Verantwortungslücke und Verantwortungsunterbrechungen“³⁰ seien, so Rauer weiter unter Verweis auf autonome Sicherheitstechnologien.

Dieser Logik folgend können nun weder Mensch noch Maschine verantwortlich gemacht werden. Zum einen wird der Mensch, wenn er Verantwortung für ein sich seiner Kontrolle immer stärker entziehendes KI-System übernehmen soll, überfordert. Zum anderen verfügt das KI-System aber auch nicht über die für moralische Urteilskraft notwendigen Eigenschaften wie Rationalität und freien Willen, so dass das System nicht in die Bresche springen und Verantwortung übernehmen könnte. Je mehr Entscheidungen also durch schwer zu kontrollierende selbstlernende Algorithmen getroffen werden, umso lückenhafter wird die Zuschreibung von Verantwortlichkeit.³¹ Damit aber steht der zu erwartende zunehmende Einsatz von selbstlernenden KI-Systemen in der realen Gefahr, dass sich die hier entfaltete Expansionsdynamik zuspitzt und an vielen Stellen immer mehr Verantwortungslücken

²⁴ K. Braun/C. Kropp, Einleitung. In digitaler Gesellschaft: Herausforderungen, Risiken und Chancen einer demokratischen Technikgestaltung, in: Dies. (Hg.), In digitaler Gesellschaft. Neukonfigurationen zwischen Robotern, Algorithmen und Usern, Bielefeld 2021, 7-34, 23.

²⁵ Martini, Black Box Algorithmus (s. Anm. 5), 22.

²⁶ Vgl. W. Rammert, Distributed agency and advanced technology. Or: how to analyse constellations of collective inter-agency, in: J.-H. Passoth /B. M. Peuker/M. W. J. Schillmeier (Hg.), Agency without actors. New approaches to collective action, London 2012, 89-112.

²⁷ V. Rauer, Distribuierte Handlungsträgerschaft. Verantwortungsdiffusion als Problem der Digitalisierung sozialen Handelns, in: PVS, Sonderheft 52 (2017) 436-453.

²⁸ Rauer, Distribuierte Handlungsträgerschaft (s. Anm. 24), 444.

²⁹ Ebd.

³⁰ Ebd. 445.

³¹ Vgl. P. Kellmeyer, Ethische Fragen bei Brain-Computer Interfaces und anderen Neurotechnologien, in: K. Liggieri/O. Müller (Hg.), Mensch-Maschine-Interaktion: Handbuch zu Geschichte - Kultur - Ethik, Stuttgart 2019, 316-324, 320.

entstehen. Diese drohende Entwicklung ist in den Mittelpunkt politischer, ethischer und rechtlicher Reflexion zu rücken, gerade auch im Hinblick auf ihre möglichen Auswirkungen auf die Demokratie im digitalen Zeitalter.

6. Verantwortungslücken in theologisch-ethischer Reflexion

Die theologische Ethik räumt der vielschichtigen Kategorie Verantwortung seit einigen Jahrzehnten eine zentrale Stellung ein. Der Begriff der *Verantwortungsethik* wird vielfach als Synonym für Autonome Moral im christlichen Kontext genutzt. Der Grundgedanke moralischer Selbstbestimmung findet seinen Widerhall im Verantwortungsbegriff, denn Subjekt der Moral ist der für seine Freiheit verantwortliche Mensch. Der theologische Ethiker Karl Wilhelm Merks spricht deshalb auch von einer „autonomen Verantwortungsethik“³².

Auch diese autonome Verantwortungsethik wird durch die Komplexität der Verantwortungszuschreibungen durch regelbasierte KI-Systeme einerseits und durch die Expansion von Verantwortungslücken durch selbstlernende KI-Systeme andererseits mit neuen Fragen konfrontiert. Will sich theologische Ethik in die Diskussionen zum Verhältnis von Mensch und Maschine einbringen, so wird sie sich notgedrungen auch mit den Herausforderungen konfrontieren müssen, die mit der Digitalisierung und ihren „unterschiedlichen Neukonfigurationen des Miteinanders von Algorithmen, Robotern, digitalen Medien und menschlichen Nutzer*innen“³³ einhergehen.

In der Vergangenheit wurde der Ruf nach Verantwortung umso vehementer, je komplexer und schwieriger die Zuschreibung von Verantwortung wurde.³⁴ Nun lässt sich fragen, ob das angesichts der beschriebenen Entwicklungen digitaler Gesellschaften auch für Gegenwart und Zukunft gilt. Oder aber ist nicht vielmehr zu befürchten, dass mit der Expansion der Verantwortungslücken auch ein expansiver Rückzug aus Verantwortlichkeiten einhergeht, was schwerwiegende Folgen für Gesellschaft und Demokratie hätte. Eine der zentralen mit der weiteren Etablierung von selbstlernenden KI-Systemen einhergehenden ethischen und rechtlichen Herausforderungen besteht deshalb darin, die durch digitale Handlungsprogramme entstehenden subjektlosen Verantwortungslücken ins Blickfeld zu rücken, zu analysieren und zu bearbeiten, was zu einer „der dringenden politischen Aufgaben in der Zukunft“³⁵ gehören wird.

An dieser Aufgabe sollte sich eine theologische Verantwortungsethik beteiligen. Was könnte sie stark machen? So wichtig und notwendig es angesichts der Komplexität unserer Lebenswelt einerseits erscheint, Verantwortung gerade nicht auf eine individuaethische Dimension zu reduzieren, sondern den Blick für ihre systemischen Aspekte immer weiter zu schärfen, so elementar bleibt es zugleich, daran zu erinnern, dass der Mensch sich seiner Verantwortung nicht leichtfertig entledigen kann und darf, auch nicht unter Verweis auf die Komplexität der Welt oder die drohende Übermacht der Maschinen. Es gilt, auch im Zeitalter der Digitalisierung für eine menschenwürdige Welt, die auch eine menschenwürdige Welt der Mensch-Maschine-Interaktionen umfasst, einzutreten. Angesichts der zunehmenden Algorithmisierung der Lebenswelt wird es theologischer Ethik deshalb zentral darum gehen müssen, individuelle und institutionalisierte Akteure zu befähigen, die immer mächtiger

³² K.-W. Merks, *Theologische Fundamentaethik*, Freiburg 2020, 159. Verantwortung fasst Merks als ein „Zuständigkeits- und Verpflichtungswissen, die Erwartung und Forderung, auf diese Wissen durch die Tat angemessen zu reagieren, und die Rechtfertigungsbereitschaft gegenüber einer hierfür berechtigten Instanz“ S. 151.

³³ K. Braun/C. Kropp, Einleitung. In *digitaler Gesellschaft* (s. Anm. 22), 10.

³⁴ Vgl. G. Banzhaf, *Der Begriff der Verantwortung in der Gegenwart: 20.-21. Jahrhundert*, in: L. Heidbrink/C. Langbehn/J. Loh (Hg.), *Handbuch Verantwortung*, Wiesbaden 2017, 149-167.

³⁵ Rauer, *Distribuierte Handlungsträgerschaft* (s. Anm. 24), 450.

werdenden Mechanismen selbstlernender Maschinen, die systematisch subjektlose Verantwortungslücken produzieren, zu entmachten.