

# Metamorphosen maschinellen Lernens 2012 – 2026

---

Innerhalb der vergangenen sechs Jahre, in denen diese Studie Gestalt angenommen hat, hätte ich für einleitende Bemerkungen zum gesellschaftlichen Stellenwert und der wissenschaftlichen Relevanz des Themas meiner Untersuchung zu verschiedenen Zeitpunkten sehr unterschiedlich ansetzen können. Dies gilt nicht nur in dem offensichtlichen Sinne, dass die letzten sechs Jahre den Zeitraum markieren, in dem die hier präsentierten Resultate meiner Forschung erst sukzessive Form angenommen haben. Im gleichen Zeitraum hat deren Gegenstand eine solche Vielzahl an möglichen Aufhängern für eine anekdotisch vermittelte Einführung in diese soziologische Studie geliefert, dass mir irgendwann gerade die Erwähnung dieses Umstand den besten Ansatz für einen Einstieg darzustellen schien. Maschinelles Lernen (im Folgenden: ML) als seit geraumer Zeit maßgebliche Variante sogenannter Künstlicher Intelligenz (KI) hat sich als soziales Phänomen in den vergangenen Jahren nicht nur beträchtlich verändert, sondern ist dabei auch kontinuierlich Gegenstand von inner- wie außerwissenschaftlichen Kontroversen gewesen. Diese betreffen Diskussionen um den aktuellen ›Stand der Technik‹ und vor allem Beurteilungen hinsichtlich der zu erwartenden – zu erhoffenden wie zu befürchtenden – Folgen ihrer Verbreitung und Weiterentwicklung.

Die Eigenheiten des Gegenstandes dieser Studie nachvollziehbar zu machen und ebenso dessen Dynamiken als soziales Phänomen Rechnung zu tragen, ist Anliegen dieser Einleitung. Damit sollen gleich zu Beginn die Spezifika des Themas vermittelt und insbesondere für die damit einhergehenden Problematisierungsweisen der Untersuchung sensibilisiert werden. Im Folgenden werde ich daher zunächst eine Skizze der jüngeren Entwicklungen um ML anfertigen und Tendenzen des diese Entwicklungen begleitenden sozial- und geisteswissenschaftlichen Diskurses nachzeichnen. Dies wird mir dazu dienen, den Gegenstand dieser Studie sowie meine Perspektive auf diesen zu spezifizieren, bevor ich schließlich den grundlegenden Ansatz und das Vorgehen meiner Untersuchung vorstelle.

## Einleitung

›Vorläufig rückblickend‹ aus dem Jahre 2026 lässt sich eine bereits mehr als ein Jahrzehnt andauernde Phase ausmachen, in der KI beständig ökonomische, wissenschaftliche und zunehmend politische Aufmerksamkeit erfahren hat. Bereits Anfang 2020 – als die Arbeit an meinem Dissertationsvorhaben am Schaufler Lab@TU Dresden begann – waren intensive öffentliche Debatten um KI als vermeintlich die Geister scheidende »Schlüsseltechnologie des 21. Jahrhunderts« (Roth 2020) anhaltend an der (post-)massenmedialen<sup>1</sup> Tagesordnung.<sup>2</sup> Die diskursive Prominenz von KI war (und ist) derart hoch, dass sich bisweilen der Eindruck aufdrängte, soziologische KI-Forscher:innen könnten allein mit der Aufgabe, den inmitten des Hypes sich verbreitenden Halb- bis Unwahrheiten um utopische wie dystopische KI-Szenarien mit wissenschaftlicher Auf- bzw. Abklärung beizukommen, mehr als genug Beschäftigung finden. Dies lag auch daran, dass sozial- und geisteswissenschaftliche Forschung zu ML mittlerweile eigene, den Präntentionen der an der (Weiter-)Entwicklung von ML maßgeblich beteiligten Unternehmen gegenüber durchaus kritische Perspektiven etabliert hatte. Zunehmend wurde grundlegende Kritik an ML in seiner gegenwärtigen Form vernehmbar: »If 2012 was the year in which the Deep Learning revolution began, 2019 was the year in which its sources were discovered to be vulnerable and corrupted« (Pasquinelli & Joler 2021, S. 1267). ML ist als soziales Phänomen beschreib- und problematisierbar geworden. Einsichten in seine gesellschaftliche Einbettung offenbaren etwa Abhängigkeitsverhältnisse, die Beobachter:innen gleichermaßen für wissenschaftliche Forschung interessant wie politisch problematisch erscheinen.<sup>3</sup>

- 
- 1 Die vorangestellte Vorsilbe markiert die Ergänzung der ›klassischen‹ Massenmedien um internetbasierte und insbes. sog. Soziale Medien.
  - 2 Alternative Formeln zur Unterstreichung (bzw. Bewerbung) des epochalen Charakters von KI vergleichen diese (oder semantisch nah: [Big] Data) mit mutmaßlich vitalen Energie- oder Wertressourcen, etwa Elektrizität, Gold oder Öl, vgl. Sudmann 2018b, S. 57, Höinghaus 2015. Wenig überraschend ziehen solche Vergleiche nicht selten Kritik und Widerspruch nach sich, vgl. etwa Spitz 2017.
  - 3 »Beobachter:innen« verwende ich hier identisch mit »Beobachtern« im systemtheoretischen Sinne (mit bestimmten Unterscheidungen bezeichnende und unterscheidende Beobachtungsinstanzen wie soziale und psychische Systeme, vgl. Luhmann 1984, S. 63 sowie Luhmann 1997, S. 67f.). Die von mir verwendete geschlechtsneutrale Form könnte daher gewissermaßen als irreführend verstanden werden, erschien mir allerdings im Sinne konsistenter sprachlicher Konventionen aus pragmatischen Gründen als beste Lösung, zumal geschlechtsbezogene Unterscheidungen sich bekanntermaßen als zentral für die die Selbstbeobachtung psychischer Systeme erweisen. Zudem erlaubt mir diese Wahl, im Folgenden ›binnendifferenzierend‹ »Beobachter« von »Beobachterinnen« und wiederum »Beobachter:innen« zu unterscheiden.

»The Great AI Awakening« (Lewis-Kraus 2016) schien sich daher bereits am Anfang meiner Arbeit an der vorliegenden Studie in mancherlei Hinsicht längst vollzogen zu haben. Unter diesem Titel hatte schon Ende 2016 das New York Times Magazine eine ausführliche Reportage des Journalisten Gideon Lewis-Kraus über den *state of the art* in Sachen KI publiziert. Den Anlass dazu schuf der damals als bahnbrechend wahrgenommenen Umstellung von Such- und Übersetzungsfunktionen bei Google auf Machine Learning-Verfahren. Bei diesem Vorgang, der viele Millionen Nutzer:innen vermutlich erstmalig mit auf gegenwärtigen Formen sogenannter Künstlicher Intelligenz beruhender Software in Kontakt kommen ließ, handelt es sich um ein Kernereignis in der Geschichte des neuerlichen Aufstiegs von »konnektionistischen« bzw. »subsymbolischen« Ansätzen der KI-Forschung und -Entwicklung, die mit dem Label ML verbunden werden (vgl. Sudmann 2018b).

Diese »KI-Schule« verfolgt einen Ansatz, der anders als bei früheren, rückblickend häufig »good old-fashioned AI« (GOFAI) genannten Strömungen nicht (mehr) auf regelbasierte Programmierung für »künstliche« Agenten setzt (vgl. etwa Pasquinelli 2017). Deren Leistungsfähigkeit – etwa sogenannter Expertensysteme der 1980er und 1990er Jahre – misst sich am Komplexitätsniveau dieses Skripts von symbolisch vermittelten Anweisungen. Die Programmierung derartiger Verfahren besteht dabei – schematisch gesprochen – aus diskreten Regeln zur Transformation definierter Inputs (vgl. etwa Muhle 2018, S. 18). Demgegenüber beruht ML – in erster Annäherung – darauf, dass ein in einem Datensatz als repräsentiert angenommener Phänomenzusammenhang algorithmisch modelliert wird. Dem »ground-truthing« (vgl. Jatón 2021, Kap. 1 & 2, Jatón 2024), d.h. der Festlegung einer Datenbasis, die den interessierenden Zusammenhang abbilden soll, folgt dann eine »Trainingsphase«, in der ein Algorithmus eine Funktion errechnet, die in bestmöglicher Annäherung die Struktur des Datensatzes, d.h. der Verteilung einzelner Datenpunkte in einem Vektorraum von Datenbeziehungen beschreiben kann. Hier sind Input und Output in Form eines Datensatzes gegeben und eine »Maschine lernt« (daher ML) Transformationsregeln zwischen beiden. Dieser Ansatz steckt grundsätzlich hinter allen populären als Durchbrüche gefeierten Neuerungen im Feld der KI-Entwicklung innerhalb der letzten 15 Jahre – von *Computer Vision* bis zu *Large Language Models* – und verhalf die Rede von Algorithmen und KI weit über Fachkreise heraus zu popularisieren.<sup>4</sup>

4 Interessanterweise werden in aller Regel nicht etwa maßgebliche Fortschritte im Bereich mathematischer oder informatischer Forschung für die rapiden Entwicklungen im Bereich ML seit 2012 verantwortlich gemacht. Wesentliche theoretische Grundlagen für die Entwicklung künstlicher neuronaler Netze der Art, wie sie in populären Verfahren zum Einsatz kommen, waren bereits in den 1980er Jahren geschaffen. Effiziente und davor zunächst überhaupt viable technische Implementierungen derselben wurden jedoch erst möglich unter der Voraussetzung von hinreichend verfügbarer Rechenleistung. Der technische »Kniff«, die Berechnungen der Netze von der linear rechnenden CPU (*central processing unit*) auf die parallel ope-

Weil die vorliegende Arbeit eine soziologische Auseinandersetzung mit gegenwärtigen Formen von KI zu leisten beansprucht, handelt sie von ML, was die Wahl des Titels dieser Studie erklärt. Die bis zu dieser Stelle synonyme Verwendung von KI und ML soll indes die semantische Unschärfe im diskursiven Gebrauch beider Bezeichnungen spiegeln. Ohne Rekurs auf KI lässt sich kaum einführend über ML schreiben und eine gewisse begriffliche Konfusion kann wohl als symptomatisch für die diskursiven Dynamiken um KI und ML gelten. Ich erwähne diesen Umstand auch deshalb, weil gegenwärtig schon in Fragen der Begriffswahl in einer maßgeblich von der Kommerzialisierung der in der Forschung entwickelten Produkte geprägten Lage soziologisch mitreflektiert werden muss, dass sich schon auf der Ebene der Gegenstandsbezeichnung Interessen manifestieren. Unter diesen Bedingungen ist es dann wenig überraschend, dass Forschende und anders praktisch mit KI Agierende – dies auch im Vorgriff auf die Ergebnisse meiner empirischen Fallstudie – den Begriff sogar explizit ablehnen, und eher noch mit ML in Zusammenhang gebracht werden wollen, andererseits in öffentlichen Auftritten oder Anträgen selbst wiederum darauf zurückgreifen. Im Folgenden wird zumeist von ML, gelegentlich von KI, mitunter aber auch von KI/ML die Rede sein. Bezieht sich erstere Bezeichnung explizit auf den Gegenstand meiner Untersuchung, verstehe ich KI im Kontext dieser Untersuchung als eine seit nunmehr etwa 70 Jahren etablierte Semantik zur Bezeichnung einer besonderen Art technischer Phänomene, denen ML zugerechnet wird und deren Eigenart im Folgenden erörtert werden soll. KI/ML hingegen verwende ich für Kontexte, deren Bezugsrahmen im angedeuteten Sinne uneindeutig ist oder aber explizit KI wie auch ML einschließt.

Für das »Great AI Awakening« als Aufstieg des ML-Paradigmas in KI-Forschung und -Entwicklung der jüngeren Zeit lässt sich pragmatisch ein Anfangspunkt im Jahre 2012 ausmachen. Die Veröffentlichung eines Fachaufsatzes von Alex Krizhevsky, Ilya Sutskever und Geoffrey Hinton (2012) lieferte den Nachweis der Leistungsfähigkeit sogenannter neuronaler Netze im Hinblick auf automatische Objekterkennung in Bildern.<sup>5</sup> Es führte erstmalig vor, dass eine technisch umsetzbare algorithmische

---

rierende GPU (*graphics processing unit*) umzulagern, war dabei von zentraler Bedeutung und zugleich ein Grund dafür, warum ML-Anwendungen zunächst aufgrund der technischen Anforderungen an GPUs kaum auf handelsüblichen PCs zum ›Laufen‹ zu bringen waren, vgl. Sudmann 2018a, S. 22ff. sowie kritisch Whittaker 2021.

- 5 Seine Beiträge in der Forschung zu künstlichen neuronalen Netzen seit den 1980er Jahren brachte Geoffrey Hinton 2019 den Turing-Award und zuletzt den Nobelpreis für Physik sowie den inoffiziellen Titel des »Godfather of AI« ein, vgl. etwa Ranosa 2019, Rothman 2023. Dass dieser Titel durchaus ambivalent besetzt ist und Hintons Rolle in Vergangenheit und Gegenwart der KI-Forschung kontrovers verhandelt werden, zeigt etwa die kritische Berichterstattung der öffentlichkeitswirksamen Aufgabe seines Postens bei Google im Jahre 2023, die er damit begründete, dass er infolge dieser freier über Risiken in Verbindung mit aktuellen Entwicklungen in Sachen KI/ML sprechen könne, vgl. Marx 2023.

mische Modellierung eines hinreichend quantitativ wie qualitativ umfassenden Datensatzes von Bildern es erlaubt, das Modell nach Abschluss der sog. Trainingsphase neue Bilder automatisch »erkennen« zu lassen.<sup>6</sup> Die Bilder im Trainingsdatensatz sind mit den Bildinhalten korrespondierenden Tags (»Hund« und über weitere 1000 Klassen) versehen und der Erfolg der Forschung bestand darin, ein Modellierungsverfahren vorlegen zu können, dass die Inhalte von bisher ungekennzeichneten Bildern zuverlässig klassifizieren, d.h. die korrekten Tags produzieren kann. Das Verfahren lieferte signifikant bessere Ergebnisse als vergleichbare Ansätze der damaligen Zeit. Noch im gleichen Jahr erschien ein weiteres Paper, das weit über wissenschaftlichen Kreise hinaus in kompakter Form der Schlagzeile für Furore sorgte, dass KI nun »völlig autonom« Katzen in YouTube-Videos Katzen erkennen könne. Die Studie einer Google-Forschungsgruppe wies nach, dass ihr auf einem Verfahren des sog. *unsupervised learning* beruhendes Modell »von sich aus« begonnen hatte, Bilder, in denen Katzen zu sehen sind, konsistent einer Klasse zuzuweisen, obwohl das algorithmische Modell mit nicht vorab klassifizierten Bilddaten erstellt worden war und daher Katzen gar nicht »kennen« konnte (vgl. Quoc et al. 2012).

Anschließend Forschung ebnete über die nächsten Jahre hinweg den Weg zu einem Zuverlässigkeitsgrad bei visueller Objekterkennung, der dem menschlicher Proband:innen in vergleichbaren Tests entspricht. Die Prominenz von sog. *human parity* als Maßstabsgröße zur Beurteilung von Verfahren maschinellen Lernens erstreckte sich bald auf weitere Anwendungsbereiche. Bald sollte *human parity* mutmaßlich auch in den Bereichen Spracherkennung und -übersetzung erreicht worden sein.<sup>7</sup> Seit der Verbreitung von sog. Large Language Models (LLMs), etwa der GPT-Reihe von Open AI, ist diese Lösung wieder und wieder neu spezifiziert und variiert worden.

Die bislang erwähnten Beispiele beziehen sich allesamt allein auf die Verarbeitung von Sprach- und Bilddaten und Anwendungen für Endnutzer:innen. Es sollte keineswegs übersehen werden, dass Verfahren maschinellen Lernens auch in wirtschaftliche, journalistische und anderen professionellen Praktiken Einzug gehalten haben. Die diversen Einsätze von ML in dieser Vielzahl von Kontexten sind unlängst auch zum Gegenstand soziologischen Interesses geworden. So liegen mittlerweile in vielfältiger Form Studien zur praktischen Nutzung von ML vor, z.B. im Kontext etwa von Polizeiarbeit, in der Justiz, im Versicherungswesen, in der (Präzisions-)Medizin, im Personalmanagement, am Finanzmarkt oder in der Wissenschaft.<sup>8</sup> Besondere publizistische Aufmerksamkeit erhielt zudem schon früh die Forschung zu »Computer Vision« im militärischen Kontext des Einsatzes

6 Konkret handelt es sich um ein sog. *Deep Convolutional Neural Network*, vgl. ebd.

7 Vgl. kritisch zu diesen Behauptungen Beaver 2022.

8 Vgl. dazu entsprechend der Reihenfolge der Themenaufzählung Text exemplarisch: Andrejevic 2018, Egbert, Kaufmann et al. 2019, Esposito & Heimstädt 2021, Završnik 2021, Cevolini

von Drohnen in Kriegsszenarien, nicht zuletzt aufgrund der von dieser und weiterer Berichterstattung ausgehenden Skandalisierung der beobachteten Praktiken (vgl. etwa Richardson 2019). Überhaupt war die sprunghafte Zunahme öffentlichen Interesses an ML zuerst bei Bilderkennungsverfahren zu beobachten – auch oder gerade, weil deren Aufstieg zu diskursiver Prominenz von brisanten Enthüllungen begleitet wurde, die problematische Aspekte der Anwendungen offenbarten.<sup>9</sup>

An Bestimmungen aus der Informatik angelehnt beruht ML – schematisch formuliert – auf der algorithmischen Verarbeitung von (zumeist sehr umfangreichen) Datensätzen. Im Zuge dieses auch als *Training* bezeichneten Vorgangs wird eine Modellierung des Datensatzes errechnet. Dieses Modell hat die Form einer mathematischen Funktion, die in möglichst präziser Annäherung an den Datensatz dessen Strukturen – die Beziehungen einzelner, als Input und Output definierter Datenpunkte zueinander – beschreiben soll. Die errechnete Funktion ist im beschriebenen Sinne eine Datenanalyse, die in der Folge weitere Daten(-analysen) produzieren kann (vgl. Dourish 2016, S. 7). Die Bandbreite dieser Klasse statistischer Analysen ist groß und reicht von linearer Regression bis zu neuronalen Netzen, die Milliarden von Gewichtungen umfassen – etwa die bereits erwähnten Sprachmodelle.<sup>10</sup> Zentral für die Einordnung als Verfahren maschinellen Lernens ist die datenbasierte Modellierung eines Datensatzes als (partiell) automatisierte Errechnung einer Funktion, die von einem bestimmten Algorithmus »gefunden« wird. Dieses »function finding« (Mackenzie 2017, S. 96) impliziert, dass die Funktion inputsensibel, d.h. in Abhängigkeit von der Datenbasis variabel ist und ggf. das Modell unter Berücksichtigung neuer Daten »weiterrechnen«, d.h. adaptiv reagieren kann. Dies unterscheidet klassische Machine Learning-Algorithmen wie *k-means* (vgl. MacQueen 1967) oder *support vector machines* (vgl. Cortes & Vapnik 1995) von anderen algorithmischen Verfahren, weshalb Algorithmen für sich nicht prinzipiell mit ML gleichzusetzen sind. Wenngleich sich hier eine diskursive Amalgamierung beobachten lässt – so weit,

---

& Esposito 2020, Hofmann 2023, Mateescu & Nguyen 2019, Lange, Lenglet & Seyfert 2016, Both 2020.

9 Man denke etwa automatische Klassifizierung Schwarzer Menschen als Gorillas in der Foto-App von Google im Jahre 2015. Der Mechanismus konnte offenbar auch drei Jahre später nur durch Entfernen der gesamten Klasse »Gorilla« – also Label wie auch zugehörige Bilder von Gorillas unterbunden werden, vgl. Vincent 2018.

10 Die im Falle sogenannter generativen Verfahren ergänzt werden um Vorhersagefunktionen im Sinne von »Auto-Vervollständigen« durch Errechnung der höchstwahrscheinlichen nächsten Elemente des Modells. So können die generative von den »diskriminierenden« Funktionen in ML unterschieden werden, wobei das Attribut allein auf den Bezug der Funktion auf die Datengrundlage und nicht etwa auf die Beurteilung von Outputs bezogen ist. Probabilistisch modellierende Bayes'sche Netze stellen ein klassisches Beispiel für generative Verfahren dar, wohingegen Support Vector Machines (vgl. Cortes & Vapnik 1995), eingesetzt als Klassifikator oder Regressor, häufig zur Erklärung von diskriminierenden Funktionen herangezogen werden, so etwa in Jebara 2004, S. 5–12.

sodass auch in den für diese Studie maßgeblichen sozial- und geisteswissenschaftlichen Diskussionen nicht immer trennscharf zwischen ML, KI und Algorithmen unterschieden wird –, halte ich terminologische Differenzierung in diesem Zusammenhang für wichtig. Folglich hat diese Studie nicht etwa Algorithmen an sich oder KI überhaupt, sondern explizit ML zum Gegenstand.<sup>11</sup>

Mit Blick auf die für ML typischen algorithmischen Modellierungsverfahren ist charakteristisch, dass die derart produzierten Datenanalysen sich aufgrund der zugrundeliegenden Errechnungsmethode für Beobachter:innen in ihrem Sinngehalt oft als schwer nachvollziehbar erweisen, auch wenn sie sinnförmig verfasst (Text, Bild) sind. LLMs ›verstehen‹ nicht etwa eine Frage und ›überlegen‹ sich daraufhin eine Antwort, sondern errechnen für eine Eingabe (›Prompt‹) einen wahrscheinlichen Output auf Basis stochastischer Analysen der ihm zugrundeliegenden (häufig unüberschaubar großen) Datensätze.<sup>12</sup>

Dieses Problem gilt in aller Regel selbst (und gerade auch) dann, wenn die Eigenschaften des Datensatzes als *ground truth* des Modells im Sinne formaler Verteilungen und Regelmäßigkeiten als bekannt wie auch im Sinne der sinnhaften Bedeutung der Daten als ›verstanden‹ vorausgesetzt werden können (vgl. Dourish 2016, S. 7). Die mitunter überraschenden, wenn nicht gar unheimlich anmutenden Outputs der Modelle (vgl. Bucher 2017) sind daher nicht ohne Weiteres in ihrer Genese ›erklärbar‹ und lassen sich hinsichtlich ihrer Implikationen meist auch im Nachhinein nicht sinnvoll erschließen (vgl. Burrell 2016).<sup>13</sup>

In den vorangehenden Sätzen klang bereits die wesentliche Rolle von Datensätzen an, die auch in der Hinsicht nicht zu unterschätzen ist, dass die Verfügbarkeit hinreichend quantitativ großer wie qualitativ gehaltvolle Datensätze eine der wesentlichsten Voraussetzungen von ML und damit auch von dessen rasantem Aufstieg innerhalb der letzten etwa zehn Jahre ist. Vor dem Hype um ML war der Hype um ›Big Data‹ – etwa in Form des schon erwähnten *ImageNet*. Dieser Datensatz wurde 2006 von der Informatikerin Fei-Fei Li konzeptualisiert als Verknüpfung des damals schon existierenden *WordNet*<sup>14</sup> mit Bildern. Innerhalb von zweieinhalb Jahren

11 Florian Jatón berichtet in seiner empirischen Studie zur sozialen Genese von Algorithmen davon, dass interessanterweise unter praktisch mit der Entwicklung von algorithmischen Verfahren Beschäftigten bisweilen oft als unklar gilt, welche Art von Algorithmen als »real machine learning« gelten können, vgl. Jatón 2021, S. 269.

12 Dessen exakte Komposition zudem nicht öffentlich bekannt ist.

13 Unter dem Label »Explainable AI« oder »XAI« wird seit längerer Zeit daran gearbeitet, diesen Umstand zu beseitigen, bislang mit nützlichen, aber überschaubaren Ergebnissen, vgl. Zednik 2021 sowie Sudmann 2020, S. 194ff.

14 Einem seit den 1980ern entwickelten lexikalischen Datensatz, der Wörter der englischen Sprache in gegenwärtig 117000 distinkten »Synsets« eingruppiert. Dabei werden die semantischen Beziehungen der Synsets zueinander erfasst und jeder Eintrag enthält zudem eine Kurzdefinition, siehe Princeton University 2010.

gelang es Lis Team, eine 3,2 Millionen (manuell in 5247 »Synsets« klassifizierte [siehe Fn. 14]) Bilder umfassende Sammlung aufzubauen (vgl. Jatón 2021, S. 272) – mittlerweile umfasst ImageNet mehr als 14 Millionen Bilder und 21841 »Synsets« (vgl. ImageNet 2021). Bemerkenswert an diesem Projekt ist nicht zuletzt, dass für die händische Annotation der Bilder Amazons Dienst »Mechanical Turk«<sup>15</sup> in Anspruch genommen wurde, weil das Bewältigen dieser Aufgabe allein durch studentische Hilfskräfte wohl Jahrzehnte gedauert hätte (vgl. Fei-Fei 2010). Die Inanspruchnahme des Crowdsourcing-Dienstes für *ImageNet* ist damit vermutlich eines der frühesten, wenn nicht das früheste Beispiel für sogenannte »ghost work« im Zusammenhang mit ML.<sup>16</sup> Mit diesem Konzept bezeichnen Mary Gray and Siddharth Suri diejenige von Menschen – zumeist (auch im Falle von AMT) aus dem Globalen Süden – verrichtete Arbeit, ohne die sogenannte Künstliche Intelligenz in ihrer gegenwärtigen Form nicht realisiert werden könnte (vgl. Gray & Suri 2019, S. IX).

Dieser Umstand berührt gleichsam unmittelbar den eben erwähnten »Durchbruchaufsatz« (vgl. Krizhevsky et al. 2012). Der in Anlehnung an den Vornamen des Erstautoren Alexander Krizhevsky als *AlexNet* bekannte Ansatz nutzt nämlich den ImageNet-Datensatz (zu dem Zeitpunkt 1,2 Mio. Bilder mit 1000 Klassen) als »Ground Truth« und stellte seine Leistungsfähigkeit im Rahmen der »ImageNet Large Scale Visual Recognition Challenge« (vgl. ImageNet 2021b) im Jahre 2012 unter Beweis, die *AlexNet* mit signifikantem Abstand zum nächstplatzierten gewann (vgl. Wei 2019). Aufgrund dieser Verstrickung erscheint es nicht vermessen, die umfangreiche Annotationsarbeit für *ImageNet* als so wichtigen wie ungewürdigten Beitrag zum Erfolg von *AlexNet* zu verstehen. Es deutet sich hier bereits an, dass die soziologische Betrachtung von ML instruktive Einsichten womöglich weniger aus der Beobachtung der Nutzung von ML, sondern vor allem auch der Genese entsprechender technischer Anwendungen gewinnen kann. So präsentiert dieser erste kursorische Überblick ML als räumlich verteilten Prozess, in dessen Verlauf verschiedene Ressourcen mobilisiert werden. In einer solchen diachronen Perspektive auf die Genese von ML-Anwendungen erscheinen diesen weniger als disruptive radikale Transformationen und vielmehr als Akkumulationsvorgänge (vgl. Mackenzie 2017, S. 5).<sup>17</sup>

15 Zum Hintergrund der Crowdsourcing-Plattform vgl. Crawford 2021. S. 63–69, bes. 67f. sowie ausführlicher zur Geschichte des »Mechanical Turk«: Standage 2022.

16 Ab 2007 waren jährlich wohl ca. 20–30.000 Menschen als Arbeitskräfte an der Klassifizierung der ImageNet-Bilder beteiligt, vgl. Markoff 2012.

17 Diese Betrachtungsebene lässt indes auch den sich in der gleichen Zeit vollziehenden Aufstieg von Sozialen Medien in neuem Licht erscheinen. Versteht man diesen nämlich vor allem als Aufstieg von Data Capture-Technologien und Cloud Services (vgl. Halpern et al. 2022), die Nutzer:innendaten zur weiteren Verwendung (und Verwertung) auf privaten Servern speichern, wird leicht nachvollziehbar, dass Soziale Medien und ML alsbald in einen datenökonomischen Zusammenhang geraten würden.<sup>17</sup> Dieser wurde bald als Plattform- und/oder Überwachungskapitalismus (vgl. etwa Srnicek 2018, Zuboff 2018) beschrieben – und beruht

## Ein dynamischer Gegenstand

Die beschriebenen Entwicklungen seit 2012 (*AlexNet*) bzw. schon 2007 (*ImageNet*) haben schnell sozial- und geisteswissenschaftliches Interesse auf sich gezogen. Mittlerweile liegt ein vielseitiges Korpus sozial- und geisteswissenschaftlicher Forschung zu ML vor, die sich in den letzten Jahren entwickelt, differenziert und konsolidiert hat. Mittlerweile ist die Rede von »Critical AI Studies« als einem – in der Formulierung der Medien- und Literaturwissenschaftler:innen Rita Raley und Jennifer Rhee (2023) – interdisziplinär operierenden »field in formation«. ML bzw. KI scheint also, so ließe sich daraus schließen, neue Ansätze geistes- und sozialwissenschaftlicher Forschung erforderlich gemacht zu haben, die jüngeren Datums sind oder überhaupt erst noch gefunden werden müssen (»... in formation«). Eine solche Proklamation lässt sich als Versuch der Konsolidierung weit verzweigter Forschungen verstehen, der auf Umstände reagiert, die auch für die vorliegende Studie von zentraler Bedeutung sind.

Eine solche Konstatierung belegt zudem, dass nicht immer völlig klar scheint, was unter ML/KI zu verstehen sei und wie sich die Forschung – theoretisch wie methodisch – auf ihren Gegenstand beziehen sollte. Dies verweist wiederum darauf, dass »Künstliche Intelligenz« als diskursive Referenz nicht einheitlich verwendet wird. Dies verwundert wenig, da schon ein historischer Blick auf KI die Geschichte eines *moving target* offenbart (vgl. etwa Fazi 2021, S. 6). Bei diesem handelt es sich um ein kontext-, situations- und interessenabhängig variabel bestimmbares Phänomen, dessen wechselnde Verständnisse das Spannungsfeld zwischen dem technisch Möglichen – »künstlich« erzeugt – und dem vermeintlich das Menschliche Auszeichnenden – häufig als Maßstab für »Intelligenz« – und beständig aufs Neue ausgelotet haben.<sup>18</sup>

Schon die Hintergründe der ersten bekannten Verwendung von »Artificial Intelligence« im Jahre 1955 belegen, dass ein wesentliches Anliegen hinter dieser

---

im Hinblick auf Wertschöpfung nicht unwesentlich auf global distribuierte »ghost work«. Wobei gewissermaßen fraglich scheint, ob (»digitaler« [Staab 2019]) »Kapitalismus« zur Beschreibung der »Informationsökonomie« überhaupt noch die richtige Bezeichnung ist. Maßgeblich dafür ist die Beurteilung der Frage, ob Daten i.S.v. Information als Kapital verstanden werden sollten, vgl. Wark 2019.

- 18 Sinnbildlich dafür gilt der dem Informatiker Larry Tesler zugeschriebene Satz: »AI is whatever hasn't been done yet«, so etwa populär etwa Hofstadter 1979, S. 601. Tesler jedoch hat eigenen Angaben zufolge stattdessen Folgendes von sich gegeben: »Intelligence is whatever machines haven't done yet.« Many people define humanity partly by our allegedly unique intelligence. Whatever a machine – or an animal – can do must (those people say) be something other than intelligence« (Tesler [o.J.]). Was sich viel eher als Kritik eines Anthropozentrismus liest, der Menschlichkeit mit Intelligenz schlechthin liest, wurde daher vermutlich nicht ganz im Sinne des Urhebers zum geflügelten Wort.

Wortschöpfung darin bestand, einen für Finanzierungsanträge attraktiven Titel für das mit Mitteln auszustattende eigene Forschungsvorhaben präsentieren zu können (vgl. Moor 2006). Anlass dieser ersten prominenten Proklamation war bekanntlich das »Dartmouth Summer Research Project on Artificial Intelligence« und damit zunächst dessen Finanzierungsantrag, ausgearbeitet und eingereicht von John McCarthy, Marvin Minsky, Nathaniel Rochester und Claude Shannon. Die im Antragstext des für den Sommer 1956 angesetzten (und durchgeführten) achtwöchigen Workshops gesetzten abstrakten Ziele lesen sich auch heute noch nachvollziehbar. Das nach wie vor prominente Schlagwort »Learning« ist ebenso enthalten – wie interessanterweise auch ein expliziter Bezug auf das Medium Sprache, das im Kontext aktueller Debatten um LLMs in aller Munde ist:

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. (McCarthy et al. 1956, S. 1)

Aus diesem Ursprungsszenario heraus hat sich für das Feld der KI-Forschung eine temporale Paradoxie hinsichtlich des durch »KI« Bezeichneten ergeben. »KI« (und gleichsam die Spezifizierung »ML«) bezieht sich immerzu auf aktuell die Forschung beschäftigende Probleme, die (noch) nicht technisch lösbar sind. Es handelt sich bei KI/ML daher in wesentlicher Hinsicht häufig primär im Wortsinne um eine Projektion.

Versuche, qua der Markierung eines wissenschaftlichen »field in formation« Übersicht zu schaffen, hängen von gegenwärtigen Visionen über zu Erreichendes, aber ebenso vom jeweils aktuellen gegebenen faktischen »Stand der Technik« ab. Dieser entwickelt sich mit Blick auf KI – aktuell mit besonderem Fokus auf ML – ebenso dynamisch. Dies haben die vergangenen zehn Jahre Technikentwicklung wohl unbestreitbar gezeigt. Als *moving target* erweist sich KI/ML daher mindestens in zweifacher Hinsicht.

»Durchbrüche« innerhalb dieses Feldes können die soziale Bedeutung und Relevanz von KI maßgeblich verändern, wenn etwa die Veröffentlichung neuer Produkte innerhalb kurzer Zeit die Defizite der vorigen vergessen macht. Sofern technische Neuerungen nicht zugleich terminologische Verschiebungen mit sich bringen, sondern unspezifisch als »KI« verhandelt werden, kann dies den Stand wissenschaftlicher Forschung schnell altern lassen. Dies gilt nicht nur für Vorhersagen über die Potenziale wie auch die Risiken von KI, die beständig Gefahr laufen, schnell widerlegt zu werden. Auch weniger an Spekulationen interessierte Forschungen stehen dem Problem gegenüber, den Anspruch auf generalisierbare Aussagen zum untersuchten Phänomen mit dessen permanenter Vorläufigkeit ver-

mitteln zu müssen. KI-Forschung ist daher als ein »field in formation« zu verstehen – ebenso wie ihr Gegenstand. Dieser ist nicht nur semantisch ein *moving target*, sondern muss auch in seiner je aktuellen technisch-materialen Gestalt immerzu als Prototyp verstanden werden.<sup>19</sup> Wichtig scheint mir dabei die Beobachtung, dass sich abhängig vom jeweiligen »state of the art« auch der soziale Bezugsrahmen für ML verändert. Rückblickend etwa lässt sich sehen, dass im Bereich visueller Kunst die für bildgebende ML-Verfahren seit 2015 so charakteristische »visual indeterminacy« (Hertzmann 2020) vor allem Ausdruck der Möglichkeiten und Grenzen von sogenannten GAN-Architekturen war. Seit dem Einzug von Verfahren wie *Stable Diffusion* (im Jahre 2022), die nahezu fotorealistische Bilder zu generieren imstande sind, kann nicht mehr davon die Rede sein, dass visuelle »KI-Kunst« gerade durch eine charakteristische Unschärfe der Bildelemente gekennzeichnet sei.<sup>20</sup>

Mit Blick auf jüngere Entwicklungen könnte auch über das Feld der Kunst hinaus schnell der Eindruck entstehen, als hätte es allein seit 2020 eine Reihe von so folgenreichen Innovationen in der Entwicklung von KI-Verfahren und -Produkten gegeben, dass fraglich scheinen könnte, inwiefern die hier primär aus empirischer Feldforschung gewonnenen Beschreibungen überhaupt in irgendeiner Form generalisierbare Gehalte zur Verfügung stellen können. Vor allem im Feld der Verarbeitung und Generierung von Text- und Bildinhalten – zunehmend gilt dies auch für auditive Formen – zeigt sich die technische Realität gegenüber dem, was vor wenigen Jahren noch als innovativ angesehen wurde, als eine entschieden andere. *CLIP*, *DALL-E*, *Stable Diffusion*, *Midjourney*, *ChatGPT*, *Gemini*, *Grok*, *Copilot*, *DeepSeek*, *Claude* (Anthropic) – seit 2020 wurde eine ganze Reihe von Bild- und Sprachmodellen veröffentlicht, denen nachgesagt wird, ihre jeweiligen »Vorgänger« in einer Vielzahl von Hinsichten übertroffen zu haben, was Leistungsfähigkeit und damit verbunden mutmaßlich auch Nützlichkeit in Anwendungsform angeht. Im Umkehrschluss heißt dies, dass vormals aktuelle Verfahren schnell in Vergessenheit geraten können.

Mit dieser Einsicht, dass bestimmte Formen von ML mitunter »schnell altern«, möchte ich sensibilisieren für potenzielle Differenzierungslinien in Bezug auf den Gegenstand dieser Studie. Der ausführlichen Auseinandersetzung mit dieser Problematik gewissermaßen vorausseilend soll deren Erwähnung an dieser frühen Stelle dem Erwartungsmanagement dienen: Empirischer Gegenstand dieser Studie sind spezifische Formen von ML zu einer gegebenen Zeit, als Praktiken darüber hinaus

19 Vgl. zur technischen Bedeutung des Prototyping bereits vor 40 Jahren Floyd 1984 sowie allgemeiner zum gesellschaftlichen Stellenwert des Phänomens Dickel 2019.

20 Ausführlich zu dieser Periodisierung mit der Unterscheidung GAN/Post-GAN im Hinblick auf bildgebende Verfahren im Kontext von zeitgenössischer KI-Kunst vor/seit 2019 vgl. Offert 2023.

auf spezifische Art und Weise situiert. Abstrahierende Generalisierungen zu maschinellem Lernen oder gar Künstlicher Intelligenz sind nicht das primäre Anliegen der Untersuchung und können nur unter bestimmten Voraussetzungen gewonnen werden. Damit ist zugleich freilich nicht ausgeschlossen, dass sich über verschiedene Generationen von ML hinweg konsistente Problembezüge in den mit ML verbundenen sozialen Praktiken ergeben – etwa für den beschriebenen Zeitraum seit 2012 – wie an späterer Stelle zu zeigen sein wird. Dennoch scheint mir dieser Hinweis zur Vorsicht an dieser frühen Stelle durchaus geboten, besonders mit Blick auf mediale Diskursdynamiken, innerhalb derer definitive Festlegungen zu den langfristigen Potentialen, Schwierigkeiten und Gefahren im Zusammenhang mit ML an der Tagesordnung sind.

Im Rahmen dieser kursorischen thematischen Sondierung, der in den Abschnitten zur Soziologie künstlicher Intelligenz eine systematische Erörterung des Untersuchungsgegenstandes folgen wird, soll nun nach dieser ersten Annäherung an ML mittels der Schilderung wichtiger Ereignisse und Dynamiken der letzten etwa zehn Jahre die Perspektive gewechselt werden.

## KI/ML-Diskurse

Innerhalb der letzten Jahre ist KI/ML innerhalb des sozial- und geisteswissenschaftlichen Diskurses zu einem prominenten Thema avanciert. Schon in quantitativer Hinsicht beeindruckt das rasant gestiegene Interesse, das sich unter anderem in der Zahl einschlägiger Artikel, Bücher und Schwerpunktheften, sowie – auf institutioneller Ebene – neu ins Leben gerufener Fachzeitschriften und spezifisch denominierter Professuren, wenn nicht ganzer Forschungseinrichtungen ausdrückt.<sup>21</sup>

In diesem Abschnitt versuche ich einführend eine knappe Sondierung prägnanter Tendenzen dieses Diskurses in schematisierender Form, die heuristisch drei distinkte Positionen identifiziert. In erster Annäherung an die sozial- und geisteswissenschaftliche Literatur zu KI/ML werde ich nicht unmittelbar auf eine systematische soziologische Auswertung der bisherigen Forschung abzielen. Stattdessen geht es mir zunächst darum, einen Eindruck der lebhaften Debatten zu vermitteln, wie sie für den gegenwärtigen Diskurs um KI/ML dies und jenseits fachwissenschaftlicher Grenzen charakteristisch sind. Damit verfolge ich die Absicht, für die Eigenheiten eines Forschungsgegenstandes zu sensibilisieren, der nicht nur, wie beschrieben, schon immer ein *moving target* gewesen ist, sondern darüber hinaus auch gegenwärtig so anhaltend wie lautstark kontrovers verhandelt wird. Wenngleich sich

---

21 In der erwähnten Einführung zum Feld der »Critical AI Studies« findet sich eine Auflistung vor allem US-amerikanischer (geistes-)wissenschaftlicher Einrichtungen, vgl. Raley & Rhee 2023, S. 198f.

soziologische Forschung von solchen Dynamiken jenseits der eigenen Fachgrenzen, vor allem in Absicht auf normative Beurteilungen von KI/ML, nicht übermäßig oder gar etwa in der eigenen Theoriearbeit beeinflussen lassen sollte, so sind diese doch zumindest auf angemessene Weise als potenziell relevante Bestandteile des Forschungsgegenstandes einzuordnen. Dies gilt besonders vor dem Hintergrund der Vermutung, dass entsprechende diskursive Dynamiken auch vor empirischen Szenarien, wie ich sie im Kontext dieser Studie untersuche, kaum Halt machen werden. Und darüber hinaus ist im Sinne einer reflexiven methodologischen Herausforderung wohl allemal anzuerkennen, dass außerfachliche Diskurse zu innerfachlich verhandelten Gegenständen den soziologischen Diskurs vielmehr durchaus informieren und nicht etwa völlig unberührt lassen. Schließlich gilt, dass auch wir Soziolog:innen das, was wir über unsere Gesellschaft wissen, doch vor allem aus den (Post-)Massenmedien wissen (vgl. Luhmann 1996, S. 9).

Zur Differenzierung diskursiver Positionen zu KI/ML beziehe ich mich auf Umberto Ecos Unterscheidung von »Apokalyptikern« und »Integrierten« (Eco 1984a [1962]), die dieser im Zuge einer Studie zur Kritik an der Massenkultur der 1960er Jahre entwickelt hatte. Mit diesen Bezeichnungen charakterisiert Eco insbesondere – in Auseinandersetzung mit der Frankfurter Schule (vgl. Eco 1984b, S. 11) – Modi der kritischen Bezugnahme auf die moderne Massenkultur. Erscheint diese Apokalyptiker:innen grundlegend als in einem Zerfallsprozess befindlich, den sie als letzte echte »Kulturmenschen« (ebd.) bezeugen, erleben Integrierte sie vielmehr als sich entwickelnd, erweiternd und auf zu begrüßende Weise wandelnd. Ecos Schema betrifft den normativen wie operativen Charakter diskursiver Bezugnahmen auf einen Gegenstand im Modus der Kritik bzw. Affirmation: Integrierte seien die Produzent:innen von Texten der Massenkultur, Apokalyptiker:innen hingegen schrieben lieber *über* sie. Im Zuge der Distanznahme qua Kritik<sup>22</sup> stifteten sie ihren Leser:innen dennoch Trost, da ihre Positionierung Identifikationspotenzial biete: »er läßt ihn, vor dem Hintergrund der drohenden Katastrophe, die Existenz einer Gemeinschaft von ›Übermenschen‹ erahnen, die sich über die Banalität und den ›Durchschnitt‹ zu erheben vermögen [...]« (ebd.); eine »ausgewählte *community*« (Eco 1984a [1962], S. 17) derer, die es besser wüssten.

Während die Apokalyptiker gerade dadurch überleben, dass sie Theorien über den Zerfall ausbilden, versagen sich die Integrierten weitestgehend der Theoriearbeit; sie erzeugen und übermitteln ihre Botschaften in unbefangener Leichtigkeit, tagtäglich auf allen Ebenen. Die Apokalypse ist eine Besessenheit der dissenters, des Andersdenkenden; die Integration ist die konkrete Realität derjenigen, die nicht abweisen, nicht anderer Meinung sind. (ebd., S. 16)

22 Etwa mittels solcher begrifflicher Konstruktionen wie »Kulturindustrie« (vgl. ebd., S. 19).

Nun handelt es sich bei der Massenkultur, wie Eco sie sich zum Thema seiner Studie machte, um einen wesentlich anders verfassten Forschungsgegenstand als ML/KI, der zudem in der Gegenwartsgesellschaft aus einer Vielzahl von Gründen gleichsam anders zu kontextualisieren wäre. Daher ist eine Modifikation des Eco'schen Modells vorzunehmen, um zu einem aussagekräftigen analytischen Schema für die Beurteilung des KI/ML-Diskurses zu gelangen. Statt von zwei soll von drei maßgebliche Positionierungen ausgegangen werden, die entsprechend der Eigenheiten von KI/ML durch erweiterte Beurteilungskriterien bestimmt sein sollen.<sup>23</sup> Diese werden nicht primär hinsichtlich ihres normativen Bezugs (Affirmation/Kritik) auf ML schlechthin unterschieden, sondern in zeitlich differenzierter Form einerseits auf ML in seiner gegenwärtiger gesellschaftlichen Form und andererseits auf ihre Prognosen hinsichtlich der Zukunft von KI/ML bezogen. Diese zwei Kriterien werden darüber hinaus um das der kognitiven Beurteilung von KI/ML ergänzt. In Anlehnung an Niklas Luhmanns Unterscheidung von kognitiven und normativen Erwartungen (vgl. Luhmann 1991, S. 138f.) sowie in Anbindung an das Anliegen der vorliegenden Studie soll festgehalten werden, inwiefern in den referierten Positionen ML als hinreichend verstanden (und mithin als Forschungsgegenstand gewissermaßen ›erledigt‹) oder aber als genuines wissenschaftliches Problem markiert wird und mithin als bislang eben nicht hinreichend verstanden gilt.

### Drei Positionen zu KI/ML

Ein erstes ›Lager‹ halte ich mit dem Schlagwort *Gegenwartsapokalypse* (im Folgenden kurz: GA) für treffend beschrieben. Diesem sollen solche Positionen zugerechnet werden, die im Modus immanenter (und häufig mit aktivistischen Absichten verbundener politisch-engagierter) Kritik auf grundlegende Probleme im Zusammenhang mit aktuellen Anwendungen von ML und deren Auswirkungen hinweisen. Apokalyptischen Charakter bekommen diese Perspektiven dabei weniger durch die Imagination einer zur willkürlichen Setzung eigener Handlungsziele fähigen, übermächtigen Form von KI/ML, die ›von sich aus‹ böswillige Anliegen entwickelt und verfolgt. Vielmehr geht es um die Rolle von KI/ML im Kontext bestehender gesellschaftlicher Gewalt- und Ungerechtigkeitsregime und drohender, wenn nicht schon manifester Klimakatastrophen. Als rechenintensive und daher ressourcenkonsumierende Technik sei von KI/ML in den Händen der Falschen eine weitere Verschlechterung der gesellschaftlichen (und planetarischen) Zustände zu

---

23 Wie zu sehen sein wird, weisen die folgenden Positionen bei allen maßgeblichen Unterschieden allerdings eine Gemeinsamkeit darin auf, dass sie KI/ML bzw. (Digital-)Technik nicht grundlegend kategorisch ablehnen. Derartige gewissermaßen technophobe Positionen mag es auch geben, doch spielen sie im KI/ML-Diskurs m.E. keine nennenswerte Rolle.

erwarten. Die Ergänzung des Zeitbezuges soll indes darauf hinweisen, dass GA-Positionen nicht notwendigerweise eine unaufhaltsame Verfallsgeschichte erzählen, sondern einerseits von nostalgischen Bekundungen absehen und andererseits die Zukunft insoweit als offen verstehen, dass sie nicht vollends resignieren und jedenfalls die Möglichkeit einer ›besseren‹ Form von KI/ML für denkbar halten. Wenngleich die Fortsetzung aktueller Tendenzen katastrophale Folgen mit sich brächte, gäbe es doch noch alternative Trajektorien für KI/ML. Ich rechne diesem Lager eine Reihe von Positionen zu, wie sie insbesondere im erwähnten Feld der geistes- und sozialwissenschaftlichen Feld der »Critical AI Studies« (Raley/Rhee 2023) zu finden sind. Wohl nicht gänzlich überraschend weisen die entsprechenden Positionen eine Affinität zu Kritischer Theorie im Sinne der Frankfurter Schule und feministischen, ökologischen, dekolonialen sowie/bzw. antirassistischen Ansätzen aus.<sup>24</sup> Als politisch engagierte Forschungsansätze stellen sie der gegenwärtigen Realität von KI/ML ein quasi-apokalyptisches Zeugnis aus und explizieren oft selbst Wege in Richtung einer besseren zukünftigen Lage, im Kontext derer sie wohl integriert gelten könnten, wenn diese Wege auch als fragil zu gelten haben und der Horizont einer nicht-apokalyptischen Zukunft zunehmend außer Reichweite zu geraten drohe (vgl. Burrell 2023).

Diesem Positionenbündel steht das Lager der *Gegenwartsintegration* (GI) gegenüber, innerhalb dessen die jüngeren Innovationsdynamiken um ML prinzipiell gutgeheißen werden. Dieses nicht selten technikindustrienahen Positionen zeigen sich angetan von bisherigen Entwicklungen um KI/ML und werden unter anderem durch das Führungspersonal von an diesem Prozess beteiligter Unternehmen wie Elon Musk, Eric Schmidt, Shane Legg oder Sam Altman vertreten. Ebenso wird es durch prominente (Technik-)Wissenschaftler wie etwa Yann LeCun oder die bereits erwähnten Geoffrey Hinton (bis 2023 Google, vgl. Fn. 26) und Ilya Sutskever (OpenAI)<sup>25</sup> öffentlichkeitswirksam repräsentiert. Diese Positionierung leuchtet nicht zuletzt als Ausdruck der Interessen einer Allianz von industrieller und wissenschaftlicher Forschung ein. Dabei ist – wie oben bereits angedeutet – zu beobachten, dass wirtschaftliche (Plattform-)Unternehmen zunehmend als Akteure der KI-Forschung in Erscheinung treten und aufgrund ihrer finanziellen Möglichkeiten gerade in Bereichen der Forschung erfolgreich sind, die hohe Anforderungen an Rechenleistung stellen, etwa an den erwähnten LLMs.

24 Vgl. etwa mit Bezug auf Aspekte von Ausbeutung, Überwachung und Klassengesellschaftlichkeit etwa Pasquinelli 2024, Noble 2018, Apprich et al. 2018, Whittaker 2021; mit Fokus auf Ökologie und Klima Bender et al. 2021, Mullaney et al. 2021; sowie bzgl. postkolonialer und rassistische Aspekte in KI/ML etwa Amaro 2022, Chun 2021, Gebru & Torres 2024.

25 Hinton und Sutskever sind infolge des Aufschwungs um ML-Ansätze seit Veröffentlichung ihres oben vorgestellten Aufsatzes im Jahre 2012 schließlich in die Forschungsabteilungen privatwirtschaftlicher Unternehmen gewechselt.

Tatsächlich spielen derart die gegenwärtige gesellschaftliche Realität von ML normativ affirmierende orientierte Positionen im geistes- und sozialwissenschaftlichen Diskurs um KI/ML keine relevante Rolle und kommen innerhalb der Fachdebatten in der Regel nur als Teil der Beschreibung des ML/KI-Feldes vor. Das Lager der *GI* soll an dieser Stelle dennoch genauer in den Blick genommen werden, da es gegenwärtig eine interessante Metamorphose zu durchleben scheint. Offensichtlich verfügt es über öffentlichkeitswirksame Kanäle, weil es nicht durch zuletzt das Führungspersonal von maßgeblich in der KI-Entwicklung tätigen Plattformunternehmen verkörpert wird. Aus ebendiesem Milieu heraus wird seit mehreren Jahren auf eindringliche Weise vor den Gefahren der Weiterentwicklung aktueller ML-Verfahren unter den gegenwärtigen Vorzeichen gewarnt. Diese könne zur Folge haben, dass autonome *AGI* (*Artificial General Intelligence*) oder andere Formen von ›Superintelligenz‹ (vgl. Bostrom 2014) entstünden, die sich nicht mehr kontrollieren ließen und möglicherweise gegen ›die Menschheit‹ richten könnten. Es ist deshalb mitunter die Rede von ›existenziellen‹ Risiken in der Weiterentwicklung von KI/ML die Rede, die es einzudämmen gelte. Solche Ansichten werden gegenwärtig zuhauf und großem Medienecho publiziert, etwa in Form offener Briefe. Derartiges geschah etwa im Frühjahr 2023, als im März unter anderem von Elon Musk, dem Informatiker Yoshua Bengio und dem Historiker Yuval Noah Harari zum Aussetzen der Forschung an LLMs für ein halbes Jahr aufgefordert wurde, da diese ein »profound risk to society and humanity« (Bengio et al. 2023) darstellen könnten. Zwei Monate später sahen sich unter anderem Open AI-CEO Sam Altman wie auch der ›AI Godfather‹ Geoffrey Hinton und Ilya Sutskeyer (zu dem Zeitpunkt *Chief Scientist* bei OpenAI) zu einem »Statement of AI Risk« bewegt (Hinton et al. 2023).<sup>26</sup> Ich bezeichne die *Gegenwartsintegrierten* deshalb als zumindest in Teilen als Lager mit einer Affinität zu Positionen, die als *zukunftsapokalyptisch* verstanden werden können.

Damit sind die Prämissen der Unterscheidung Umberto Ecos umgekehrt und eine widersprüchlich anmutende Dynamik markiert. Es sind nun die ehemals Integrierten, die apokalyptische Szenarien beschwören und dafür ihrerseits von den ›eigentlichen‹ Apokalyptiker:innen (der Gegenwart: *GA*) für ihre falsche Version der Apokalypse kritisiert werden. Im Sinne schismogenetischer Dynamiken könnten diese sich in dieser Konstellation dazu gezwungen sehen, stärker als vorher zu markieren, dass ihr Narrativ gefährlicher Tendenzen im KI/ML-Feld vielleicht doch nicht ganz so apokalyptisch war – oder zumindest unbedingt klarmachen, dass es ihnen um eine andere, gewissermaßen ›authentischere‹ Version von KI-Apokalypse

26 Als Randnotiz ist hierbei festzuhalten, dass Hinton, der vorher kaum als öffentliche Figur in Erscheinung getreten war, vor diesem Bekenntnis seinen Posten bei Google aufgab, um sich laut eigener Aussage als individuelle Person und eben nicht mit Google Affiliierter äußern zu können.

ging.<sup>27</sup> Ihre Kritik an der GI-Position richtet sich daher besonders auch auf deren Instrumentalisierung des apokalyptischen Gedankens.<sup>28</sup> Damit werde die Angst vor dem Einbruch der Apokalypse als Vorwand genutzt, um besser Aussichten auf zukünftige finanzielle Zuwendungen durch Investments zu gewährleisten, aber auch die eigene Position dies und jenseits von KI/ML zu stärken.<sup>29</sup> Auch gegenüber politischen Regulationsversuchen brächten sich ›Neo-Apokalyptiker:innen‹ in Stellung, schließlich könnten sie so für sich in Anspruch nehmen, die drohenden Gefahren im Blick zu haben und glaubwürdig anmutend behaupten, selbst mit größtem Interesse an sicherer und ›sozialverträglicher‹ (schon im Sinne von ›friedlicherer‹ oder ›katastrophale Folgen bestmöglich verhindernder‹) KI zu arbeiten (vgl. etwa Burrell 2023, Gebru & Torres 2024).

Ein drittes Lager diskursiver Positionen scheint mir schließlich mit *Epistemische Krise* (EK) treffend beschrieben. Diese Positionen zeichnet aus, dass sie die kognitive Herausforderung in der Auseinandersetzung mit KI/ML in den Vordergrund stellen. Wenngleich sie positive wie negative realweltliche Effekte durchaus nicht leugnen, entsprechen ihre kritischen Perspektiven auf KI/ML doch zumeist Beobachtungen zweiter Ordnung. So etwa, wenn sie in KI/ML das Potenzial einer weiteren narzisstischen Kränkung der Menschheit im Sinne Sigmund Freuds vermuten – wobei freilich den ›allzu humanistisch‹ Argumentierenden (vor allem

27 Das Konzept der Schismogenese geht auf Gregory Bateson zurück, der dieses im Kontext seiner Forschung zur Beschreibung von sozialen Beziehungsdynamiken »progressiver Differenzierung« (Bateson 1983a, S. 108) innerhalb von (ethnografisch beobachteten) Kollektiven entwickelte: Mitglieder einer Gruppe A verhalten sich gegenüber Mitgliedern einer Gruppe B anders als gruppenintern, und Gleiches gilt umgekehrt. Sofern das Verhalten gegenüber der anderen Gruppe von dieser als provokant wahrgenommen wird – etwa verunglimpfend oder (so bei Bateson) prahlerisch –, kann diese geneigt sein, sich zu verteidigen und den so wahrgenommenen Angriff zu erwidern, woraufhin wiederum komplementäre Reaktionen erwartet werden können usw. Es ergibt sich darin eine Dynamik wechselseitiger Verstärkung, die zu immer schärferen Angriffen, dabei aber zugleich auch zur Ausbildung und Stabilisierung der eigenen Gruppenidentität führt (vgl. ebd.).

28 Diese Überlegung scheint mir parallel zu einem Gedanken Ecos hinsichtlich des Nutzens apokalyptischer Schwarzmalerei zu liegen, wenn dieser nämlich fragt: »Sind die apokalyptischen Visionen möglicherweise das raffinierteste Produkt, das sich dem Massenkonsum darbietet?« (Eco 1984a, S. 17).

29 Das Anliegen, die Menschheit vor KI zu retten, sei daher ein vorgeschobenes. Tatsächlich zeige sich im Lager der GI vor allem eine Affinität zu eugenischem Gedankengut. Das technindustrielle Milieu derer, die an AGI bzw. ›Superintelligenz‹ arbeiten, zeichne sich jenseits der offiziellen Behauptungen des noblen humanistischen Kernanliegens jedoch vor allem durch rassistische, xenophobe, klassistische und sexistische Einstellungen aus (vgl. Gebru & Torres 2024). Das Akronym *TESCREAL* fasst die maßgeblichen Ideologien zusammen, die sich in diesem Milieu anzutreffen seien, zu denen »transhumanism, Extropianism, singularitarianism, cosmism, Rationalism, Effective Altruism, and longtermism« (ebd.) zu zählen wären.

unter den GA) eine Kränkung zugeschrieben wird (vgl. Bratton 2024)<sup>30</sup>. Eine solche sei mit Blick auf KI/ML nicht nur ein wesentliches epistemisches Hindernis, sondern verstelle gleichermaßen den Blick auf praktisch unausgeschöpfte Potenziale, für deren ausbleibende Realisierung ein mangelndes Verständnis derselben infolge eines normativ verankerter Anthropozentrismus verantwortlich gemacht werden können. Die Problemdiagnose des Lagers ED richtet sich damit also vor allem auf gegenwärtige Schwierigkeiten im theoretischen Verständnis von KI/ML. Der Diskurs befinde sich daher in einer Situation, in der für die Beschreibung der gegenwärtig vorgefundenen Lage noch keine adäquaten begrifflichen Mittel vorhanden seien.<sup>31</sup> Entsprechende Einladungen zur Arbeit an einer dem rechen-technischen Stand der Dinge angemessenen »Philosophy of Computation« (vgl. etwa Bratton & Agüera y Arcas 2022, Cavia 2023) stehen dem Eindruck gegenüber, dass im weiteren diskursiven Umfeld die Verständnisschwierigkeiten im Zusammenhang mit KI/ML größtenteils verkannt und KI/ML entsprechend allein als Chiffre verhandelt werden. Das genuin Neue und eigentlich Interessante an KI/ML werde demzufolge überlagert, wenn nicht gar völlig unterdrückt bzw. »begraben« angesichts der durch industrielle Propaganda gesicherte Dominanz althergebrachter semantischer Stereotypen (vgl. Hui 2023).<sup>32</sup>

Umgekehrt jedoch, so wäre wiederum aus Sicht der *Gegenwartsapokalyptiker:innen* kritisch am EK-Lager anzumerken, laufe Rhetorik, die vor allem die unergründliche Komplexität und Schwerverständlichkeit von KI/ML unterstreicht, doch immerzu Gefahr, aus dem Blick zu verlieren, was KI/ML als realweltlicher Phänomenzusammenhang tatsächlich bewirke. Angesichts der zahlreichen bedenklichen Effekte von ML-Anwendungen müssen solche Anliegen wohl als verblendet erscheinen – solange man jedenfalls der Ansicht ist, ML stelle nicht viel mehr als einen »fancy way to calculate an average« (Galloway 2021) dar.

In der Zusammenschau ergibt sich das Bild einer diskursiven Disposition, in der die drei besprochenen Positionenbündel als recht trennscharf unterscheidbare Lager unversöhnlich zu koexistieren scheinen, die jeweils vertretene Ansichten wechselseitig ablehnen. Im Kontext der vorliegenden Untersuchung – zumal an

30 Zu denen dieser bekanntermaßen die kosmologische, biologische und psychologische Kränkung in Verbindung mit den Entdeckungen von Kopernikus, Darwin und zuletzt seinen eigenen zählt, vgl. Freud 1966 [1917].

31 Diese Beschreibung erinnert diagnostisch an William Ogburns Konzept des »cultural lag«, das eine »kulturelle Phasenverschiebung« beschreibt, die eintrete, »wenn von zwei miteinander in Wechselbeziehungen stehenden Kulturelementen das eine sich früher oder stärker verändert als das andere und dadurch das zwischen ihnen bisher vorhandene Gleichgewicht stört« (Ogburn 1972, S. 328, vgl. auch Luhmann 1980, S. 296).

32 Yuk Hui identifiziert dies mit Bezug auf Hans Blumenberg als »eschatological imaginary« (ebd.) in Parallele zur Dysfunktionalität theologischer Semantiken im Kontext der Beschreibung modernen Gesellschaften.

diesem frühen Punkt derselben – möchte ich nicht zu elaborierten Erklärungsversuchen für die offenkundige Uneinigkeit ansetzen. Gleichmaßen geht es mir nicht um begründete Parteinahme im Widerstreit der Positionen. Offensichtlich könnte diese Heuristik als Ausgangspunkt für weiterführende Untersuchungen dienen, schließlich wirft das abgebildete Schema viel mehr neue Fragen auf, als dass es irgendetwas erklären könnte. Eine ›echte‹ Diskursanalyse könnte sicherlich auf produktive Weise genauer die Bedingungen dieser bemerkenswerten Inversion des Eco'schen Schemas sowie dessen Effekte hinsichtlich diskursiver Hegemonieverhältnisse in den Blick nehmen. Dies ist allerdings gerade nicht das Anliegen, das ich hier verfolge. Stattdessen geht es mir an dieser Stelle allein um die Veranschaulichung verschiedener Aspekte und Dynamiken gegenwärtiger KI/ML-Diskurse.

Die vorläufige Sichtung desselben scheint mir im Ergebnis einerseits angesichts eines nicht zu knappen Angebots an Beurteilungen von ML vor allem theoretische Vorsicht (vor allzu verbindlichen konzeptuellen Festlegungen) zu gebieten. Noch vor der systematischen Auseinandersetzung mit vorhandenen sozialtheoretischen Angeboten und erst recht eigener empirischer Forschung zum Untersuchungsgegenstand entnehme ich der vorangehenden Diskussion in erster Linie Hinweise mit Blick auf mein Untersuchungsdesign. Dies sollte derart verfasst sein, dass ich ML darin jenseits der besprochenen diskursiven Abgründe und mithin als Phänomen verstehbar zu machen vermag, das apokalyptisch wie auch integriert und als Fortschreibung bekannter Phänomene wie auch als Bruch und genuine Emergenz beschrieben werden kann.

Darüber hinaus zeigt sich, dass die geschilderten Kontroversen von einem Diskurs zeugen, in dem Beobachtungen erster wie zweiter Ordnung gleichermaßen eine Rolle spielen. Insbesondere die ›soziologienahen‹ Positionsbündel *Gegenwartsapokalypse* und *epistemische Krise* üben sich vor allem in der kritischen Beobachtung von blinden Flecken (wechselseitige Zuschreibung ›ideologische Verblendung‹) und latenten Funktionen (etwa Mobilisierung finanzieller Mitteln oder Entzug politischer Regulierung im Lager der *Gegenwartsintegration*). Es handelt sich mithin um einen Diskurs, der KI/ML bereits längst *soziologisiert* hat. Soziologische Forschung zu ML hat folglich zu berücksichtigen, dass der außerfachliche Diskurs um ihren Gegenstand selbst längst voller soziologischer Argumente ist. Dies heißt für soziologische KI-Forschung einerseits, dass ML als soziales Phänomen in Abhängigkeit von gewählten Perspektiven unterschiedliche Charakteristiken aufweist und diese Perspektiven sich andererseits als von bestimmten Interessen geleitet erweisen können.

Die Konstellation der drei Positionen offenbart gemäß der gewählten Vergleichskriterien eine dreifache Frontstellung. In den gegenwärtigen ist nicht allein umstritten, wie die Gegenwart und Zukunft von ML normativ zu beurteilen ist. Darüber hinaus gibt es keinen Konsens darüber, wie ML überhaupt zu verstehen

ist. In dieser Hinsicht ist der aktuelle Diskurs womöglich auch als Ringen um ein adäquates Verständnis technischer Phänomene ohne eindeutig identifizierbare historische Vorbilder zu verstehen. Die Beobachtung dieses Ringens sorgte bei mir früh mitunter für Verwirrung, bald aber auch für intellektuelle Neugier und weckte in mir den Wunsch, diesem eigentümlichen Untersuchungsgegenstand meine Aufmerksamkeit als Forscher zu schenken. Entsprechend möchte ich in dieser Studie die bestehenden Kontroversen zum Anlass für eine soziologische Untersuchung machen, im Zuge derer ich als Feldforscher die empirische Realität von ML in den Blick nehmen werde, um auf dieser Basis einen eigenen Zugang zu der Frage danach zu gewinnen, wie sich ML mit Mitteln der soziologischen Theorie auf instruktive Weise beschreiben lässt. Zu dieser empirischen Realität von ML zählen dann freilich die Beschreibungen von KI/ML im Feld wie auch die praktische Realisierung des derart Beschriebenen. Der vorangehenden kursorischen Betrachtung des Untersuchungsgegenstandes entnehme ich also vor allem auch den Hinweis, dass mit Blick auf meinen Untersuchungsgegenstand Beobachtungen zweiter wie auch dritter Ordnung desselben bzw. von dessen Beobachter:innen relevant sein werden

In diesem Sinne möchte ich als erstes Zwischenfazit dieser Studie eine Reihe von Aspekten festhalten, die mir mit Blick auf ML für den weiteren Fortgang der Untersuchung – sowohl hinsichtlich Theoriebildung wie auch methodologischer Erwägungen für den empirischen Teil der Studie – besonders relevant erscheinen. In Form einer knappen Auflistung scheint mir insbesondere bemerkenswert, dass KI/ML erstens als *moving target* innerhalb von Projektionen zukünftiger gesellschaftlicher Szenarien fortdauernd als Spekulationsobjekt dient, deshalb zweitens als umkämpft zu gelten hat, was unter KI/ML eigentlich zu verstehen sei und dies schließlich drittens als Grundlage von Kontroversen darum gelten kann, wie KI/ML im Kontext geistes- und sozialwissenschaftlicher Forschung verhandelt wird, was selbstverständlich auch soziologische Analysen des Phänomens wesentlich betrifft. Vor dem Hintergrund dieser Überlegungen werde ich im Folgenden vorstellen, wie ich mich im Kontext dieser Studie ML soziologisch annähern werde.

## Ansatz und Aufbau der Untersuchung

Einer kurzen Darstellung der jüngeren Geschichte im KI-Feld folgend habe ich im vorangehenden Abschnitt vor allem zu zeigen versucht, auf welche Art und Weise KI/ML aktuell diskursiv problematisiert wird. Im Folgenden gehe ich nun davon aus, dass die geschilderten Auseinandersetzungen als Symptom für Desiderate soziologischer Forschung verstanden werden können. Dies gilt mit Blick auf Theoriebildung wie auch empirische Sozialforschung. Die vorliegende Studie möchte ich als Versuch der Reaktion auf diese Probleme verstehen, die im Kern Aspekte von So-

zialtheorie und Methoden qualitativer Sozialforschung betreffen. Worauf genau ist diese Studie dann eine Reaktion?

KI/ML ist als vielgestaltiges soziales Phänomen nicht ohne Weiteres in seiner ›nackten‹ materialen Gestalt erkennbar und eindeutig festlegbar. Stattdessen ist davon auszugehen, dass den zahlreichen Interessenparteien, die als *stakeholders*<sup>33</sup> in unterschiedlicher Form an der Entwicklung, Verbreitung oder Nutzung von praktischen KI/ML-Anwendungen beteiligt sind, eine ebenso zahlreiche Vielfalt an interessengeleiteten Beobachtungsperspektiven auf den Gegenstand entsprechen. Diese Konkurrenz differierender und durchaus konfligierender Ansichten möchte und kann ich nicht auflösen. Stattdessen soll es expliziter Anspruch dieser Studie sein, diesen Umstand vom Explanans zum Explanandum machen – in empirischer Perspektive. Dementsprechend sieht das Vorgehen der Untersuchung von einer übermäßig umfänglichen theoretischen Bestimmung von KI/ML ab, und geht umgekehrt vor, angeleitet von der folgenden grundlegenden Frage: Wie zeigt sich KI/ML in der gesellschaftlichen Wirklichkeit? Im Zusammenhang mit den vorangehenden Überlegungen stellt sich sodann unmittelbar eine Reihe verschiedener reflexiv gerichteter soziologischer Folgefragen: An welche sozialen Bedingungen ist die jeweils beobachtete Erscheinungsform geknüpft? Welchen Vorstellungen von KI/ML spielen im konkreten praktischen Geschehen eine Rolle? Und ebenso: Suggestieren empirische Beobachtungen Widersprüche zwischen im Feld geäußerten Vorstellungen des Gegenstand und dessen tatsächlicher praktischer Bedeutung?

Kaum zufällig finden sich in verschiedenen, auch hinsichtlich ihrer disziplinären Hintergründe differierenden Studien Hinweise darauf, dass ML *vielen* sei bzw. sein könne. Die Literatur- und Medienwissenschaftlerinnen Rita Raley und Jennifer Rhee etwa machen »Critical AI Studies« – wie eingangs erwähnt – als ein eigenes Forschungsfeld aus, innerhalb dessen eine Vielzahl unterschiedlicher Disziplinen wichtige Diskursbeiträge liefern. Den Autorinnen zufolge spiegle dieser Umstand nicht zuletzt das auch historisch nachweisbare interdisziplinäre Wesen der KI-Forschung (vgl. Raley/Rhee 2023, S. 188). Ihr Verständnis von KI als »assemblage of technological arrangements and sociotechnical practices, as concept, ideology, and dispositif« (ebd.) ruft eine ganze Reihe gängiger soziologischer Konzeptströmungen auf und wird ergänzt um Verweise auf Korporationen, Geräte, Infrastrukturen, Industrien, Umweltkosten und Ausbeutung (ebd., S. 189). Der Text, erschienen im Frühjahr 2023 als Einleitung eines Sonderhefts von *American Literature*, liest sich als Erkundung der Frage, was literaturwissenschaftliche Forschung im engeren

33 Ein analytisches Modell ausgehend von an ML beteiligten *stakeholders* entwirft Carlos Zednik basierend auf der einleuchtenden Annahme, dass die in *Explainable AI (XAI)* implizierte Erklärbarkeit von ML-Anwendungen als jeweils relativ zur Position der beteiligten Parteien zu verstehen ist, vgl. Zednik 2021.

Sinne über ihre Disziplinengrenzen hinaus an Fachliteratur *zusätzlich* zu berücksichtigen hätte, wenn sie sich auf gegenstandsadäquate Art und Weise mit KI auseinandersetzen will. Es ist naheliegend, dass diese Intuition gleichermaßen auch für soziologische Forschung zu gelten hat. Die Vielfalt der in dieser Studie referierten Literatur wird dies – hoffentlich auf gewinnbringende Weise – demonstrieren können. Einer Vielzahl an Phänomenaspekten sollte bestenfalls eine reichhaltige, diese beleuchtende Literatur entsprechen, deren soziologischer Gehalt allerdings mitunter nicht unmittelbar offensichtlich ist und deshalb expliziert werden muss.

Mit dem Facettenreichtum des Untersuchungsgegenstands dieser Studie geht sogleich einher, dass diese kaum den Anspruch verfolgen kann, ein vollständiges Panorama von ML in seiner phänomenalen Mannigfaltigkeit herzustellen. Es gilt für diese Studie geradezu gilt das Gegenteil: gerade ein mikrologischer (bzw. -skopischer), auf die Entdeckung instruktiver Zusammenhänge im vermeintlich Kleinen abzielender methodischer Zugang wird, so hoffe ich, zu Einsichten führen, die sich womöglich in der abstrahierenden Theoriediskussion noch gar nicht finden lassen und diese anregen können.<sup>34</sup> In Form einer ethnografischen Fallstudie verspreche ich mir Einblicke in ML als empirisches Phänomen zu erlangen und dadurch Perspektiven auf ML zu gewinnen, die bislang womöglich wenig beachtete Aspekte des Untersuchungsgegenstandes in den Blick bringen. Dies scheint mir eine vielversprechende, bislang kaum eingeschlagene Ausrichtung von soziologischer Forschung zu ML darzustellen. Noch 2024 halten Roger Häußling, Claudius Härpfer & Marco Schmitt fest:

Für eine intensive soziologische Auseinandersetzung mit der neueren Künstlichen Intelligenz fehlen bislang zum einen noch Möglichkeiten einer empirischen Auseinandersetzung mit dem Thema jenseits öffentlicher oder wissenschaftlicher Diskurse und zum anderen auch ein theoretischer Zugang der techniksoziologisch, wie sozialtheoretisch überzeugt. (Häußling et al. 2024, S. 10)

Habe es in den letzten Jahren zweifellos instruktive soziologische Beiträge zu KI/ML gegeben, so seien es doch zumeist eher Studien mit ähnlicher Ausrichtung wie mein vorangehender Versuch einer fragmentarischen Diskursanalyse, die soziologische Diskussion um KI/ML geprägt hätten. Dementgegen beabsichtige ich mit dieser Studie, einen Beitrag hinsichtlich der explizit identifizierten Forschungslücken zu leisten, indem ich ethnografisch in Form einer vergleichenden Fallstudie detaillierte Einblicke in ML als soziale Praxis zusammentrage, um die aus deren Analyse gewonnenen Einsichten dann zum Ausgangspunkt einer Diskussion verschiede-

---

34 Was im Grunde einem Plädoyer für die Produktivität qualitativer empirischer Forschung gleichkommt; vgl. expliziter dazu im Kontext von Technikforschung Suchman, Gerst & Krämer 2019.

dener Aspekte einer Sozialtheorie maschinellen Lernens zu machen. Zuvor sollen jedoch zunächst der Gegenstand dieser Untersuchung auf systematische Weise soziologisch erschlossen und die mobilisierten theoretischen, methodologischen und methodischen Mittel vorgestellt werden. Diesem Anliegen sind die folgenden zwei Kapitel gewidmet. Um einen Überblick zum im Verlauf der vorliegenden Studie zu Erwartenden zu schaffen, möchte ich zuvor den Aufbau der Untersuchung vorstellen.

Das anschließende zweite Kapitel zur dient der Rekapitulation des soziologischen Diskurses zu ML im engeren sowie mitunter auch KI im weiteren Sinne. In einem ersten Kapitelabschnitt werde ich ML in der Auseinandersetzung mit in verschiedenen geistes- und sozialwissenschaftlichen Fachdisziplinen verankerten Phänomenbeschreibungen als einen schwer zu fassenden und ambivalent verfassten Forschungsgegenstand bestimmen. Im Zentrum dieser gewissermaßen sozialphänomenologischen Annäherung an ML steht dann schließlich die mit dessen Einsatz typischerweise einhergehende Gleichzeitigkeit widersprüchlicher Eindrücke. Beobachter:innen kann ML als zugleich vertraut wie auch befremdlich erscheinen. Diese Spannung haben instruktive sozialtheoretische Beiträge zu ML häufig explizit aufgenommen und auch mir wird diese Disposition als Ausgangspunkt soziologischer Theoriebildung dienen. Diese Beobachtung werde ich zunächst zum Ausgangspunkt einer Erörterung sog. Lernalgorithmen als operationslogischem Kern von ML-Verfahren machen. Dabei werde ich insbesondere ML von sog. GOF AI (*good old fashioned AI*) unterscheiden und zu plausibilisieren versuchen, welche soziologischen Implikationen mit *Maschinellenlernen* als technischem Kognitionsvorgang einhergehen. Dies wird der Spezifizierung meines Untersuchungsgegenstandes mit Blick auf die anschließenden – empirischen wie theoretischen – Analysen dienen und zugleich klare Unterscheidungskriterien wie auch Ähnlichkeiten zu verwandten (digital-)technischen Phänomenen offenbaren lassen. Daraufhin werde ich mittels einer techniktheoretischen Erörterung konstitutiver Ambivalenzen in ML erste Elemente für einen sozialtheoretischen Beschreibungsrahmen entwickeln. Ausgehend von der Operationslogik algorithmischer Modellierung wird die Erörterung markanter Eigenheiten von ML gleichermaßen darauf eingehen, in welcher geschichtlichen Linie die sich in ML manifestierende Epistemologie steht und inwiefern folglich auch die mit ML emergente Sozialität nicht ohne historische Vorbilder ist. Daran anschließend – und bereits in Annäherung an die folgende empirischen Operationalisierung– werde ich dessen Dimensionen und Extensionen von ML als soziales Phänomen in den Blick nehmen. Wie zu zeigen sein wird, handelt es sich bei ML um ein räumlich distribuiertes wie zeitlich differenziertes Phänomen, das sein technisches Potenzial über verschiedene Phasen hinweg akkumuliert, wobei sich die Phasen ihrerseits auf eine Vielzahl von Ressourcen angewiesen zeigen werden. Diese Ressourcen

werden als unverzichtbare Infrastrukturen dessen erscheinen, was schließlich in technischer Anwendung ML genannt wird.

Auf diese Überlegungen aufbauend findet eingangs des dritten Kapitels zunächst eine explizite sozialtheoretische Problematisierung von ML statt. Unter maßgeblichem Rückgriff auf systemtheoretische Überlegungen zielt diese Problematisierung darauf ab, ML als soziales Phänomen präzise beschreibbar zu machen. In kritischer Auseinandersetzung mit Elena Espositos Begriff »künstlicher Kommunikation« (vgl. Esposito 2017, 2022) soll nicht nur eine soziologische Bestimmung des Untersuchungsgegenstands geleistet werden, sondern ebenso eine argumentative Grundlage für dessen empirische Operationalisierung als soziotechnische Praxis geschaffen werden. Infolge der sozialtheoretischen Betrachtung und der empirischen Operationalisierung sollen im weiteren Verlauf der Untersuchung zwei Konzeptualisierungen von ML unterschieden werden. Mit Bezug auf erstere wird ML als Theorieproblem erscheinen, dessen Diskussion bislang für eine Reihe instruktiver Beschreibungen in Form von Oxymora gesorgt hat, die auf Schwierigkeiten eines angemessenen Verständnisses rechentechnischer Kontingenz verweisen. Mit Blick auf die empirische Untersuchung hingegen wird ML als eine soziotechnische Praxis zu verstehen sein, die durch die Beteiligung der beschriebenen rechentechnischen Verfahren wie auch entsprechender praktischer Selbstbeschreibungen als solche zu identifizieren ist. Diese methodologische Konzeptualisierung wird zusätzlich eine situationstheoretische und -analytische Unterfütterung erfahren. In Vorbereitung des konkreten methodischen Untersuchungsdesign werde ich John Deweys Verständnis der Situation vor allem hinsichtlich seiner methodologischen Potenziale mit Blick auf empirische Analysen von Situationen unter Beteiligung von ML vorstellen. In Auseinandersetzung mit weiteren grundlegenden situationstheoretischen Überlegungen (Goffman, Boltanski & Thévenot, Mische & White) und Ansätzen, die technische Vermittlungs- und Ausstattungskonzepte von Situationen in den Blick nehmen (Knorr-Cetina, Marres), werde ich einen methodologischen Rahmen für die Analyse der für diese Studie titelgebenden »Situationen maschinellen Lernens« entwickeln. Als deren analytische Angelpunkte werden sich die initiale Unbestimmtheit und der konstitutive Problembezug einer jeden Situation sowie der Kooperationscharakter ihrer praktischen Transformation erweisen. Zuletzt stelle ich im dritten Kapitel mein konkretes methodische Vorgehen einer empirischen Situationsanalyse vor und führe abschließend in die vergleichende ethnografische Fallstudie ein, die das Fundament meiner empirischen Untersuchung darstellt.

Die Ergebnispräsentation der ethnografischen Studie im vierten Kapitel wird im Anschluss an eine kurze Einführung zur Darstellungsweise fallweise vorgenommen. Zunächst werde ich meine Fallstudien der Entstehung eines Werkes des Medienkünstlers Fabian Hampel sowie der bienenbiologischen Forschung in einem biorobotischen Labor jeweils überblicksartig vorstellen. Daran anschließend werde

ich in Form von je drei prägnanten ethnografischen Vignetten und unmittelbar anschließenden Kurzanalysen die maßgeblichen analytischen Motive meiner in der Feldforschung gewonnenen Beobachtungen präsentieren. Die Analysen richten sich insbesondere darauf, welche situativen Problematisierungen sich in den beobachteten Fällen unter bestimmten Bedingungen zeigen. Um eine ergiebige Analyse im Sinne meiner Fragestellung nach ML-spezifischer Sozialität zu ermöglichen, wird mir dabei insbesondere um die Sichtbarmachung der situativen Dynamiken mit Blick auf verschiedene Aspekte von ML gehen.

Das anschließende fünfte Kapitel »Experimentelle Problemkalibrierungen« erörtert die Resultate der empirischen Studie in Form einer vergleichenden Analyse. Eng am empirischen Material werden die beobachteten situativen Probleme mit und Problematisierungen von ML, aber auch die praktischen Strategien zur produktiven Bewältigung und Transformation der vorgefundenen Situationen hinsichtlich ihres soziologischen Gehalts diskutiert. Dabei wird sich eine charakteristische doppelte Unbestimmtheit, einerseits von ML und andererseits der Situation seines Einsatzes, als Grundlage der praktischen Probleme im Umgang mit ML erweisen. Die konkreten empirischen Artikulationen diese Disposition werden maßgeblich als Resultate epistemischer Unsicherheit in der Anbahnung von ML als praktischer Kooperation verstanden. Im Anliegen der Überwindung dieser Unsicherheit werden sich Bemühungen um vermittelnde Übersetzungen von Datenanalysen, interdisziplinäre Problemaufklärung und akkumulierende Infrastrukturierung zum Aufbau eines praktisch verfügbaren Gedächtnisses ausmachen lassen. Mit Blick auf die spezifischen Umstände der untersuchten Fälle in Kunst und Wissenschaft wird sich zeigen, dass die beobachteten Praktiken aufgrund ihrer experimentellen Anliegen Situationen nicht nur geradezu heraufbeschwören, sondern diese auch im Sinne der Anliegen für sich einzuhegen wissen. Mit den Konzepten der *Problemkalibrierung* und der *schwachen Situationalität* beschreibe ich daraufhin, welche maßgeblichen Mittel die Praktiken zur Bewältigung ML-spezifischer technischer Kontingenz aufbringen. Abschließend betrachte ich die erarbeiteten Befunde im Hinblick auf ihre Generalisierbarkeit und beurteile im Lichte des gewählten methodologischen Rahmens insbesondere hinsichtlich des analytischen Stellenwerts von Unbestimmtheit und Problemen ihre Implikationen für die maßgeblichen Fragestellungen der Untersuchung.

Das sechste Kapitel »Soziotechnische Intelligenz mit stochastischen Kontingenzmaschinen« verknüpft dann schließlich die empirischen Ergebnisse mit sozialtheoretischen Reflexionen. Die beobachteten *Situationen maschinellen Lernens* werden, verstanden als *soziotechnische Intelligenztests*, dabei zur Grundlage von theoretischen Überlegungen bezüglich des ambivalenten Status von Intelligenz. Im Anschluss an die Diskussion verschiedener Facetten eines soziologischen Intelligenzbegriffs stelle ich dann schließlich mein im Fortlauf der Untersuchung erarbeitetes Verständnis von ML systematisch vor. ML wird darin als *stochastische*

*Kontingenzmaschine* bestimmt, deren soziale Funktion die *rechentechnische Transformation sozialer Unbestimmtheit in kontingente Beobachtungen* ist.

Im abschließenden Kapitel zur »Realität wahrscheinlicher Fiktionen« werde ich die Ergebnisse der Studie kompakt zusammenfassen und von diesen ausgehend wissenssoziologische und gesellschaftstheoretische Anknüpfungspunkte an die Untersuchung aufzeigen. Dabei werde ich insbesondere auf epistemologische Hürden im theoretischen Umgang mit ML eingehen und maßgebliche Implikationen des in dieser Studie erarbeiteten Verständnisses von ML im Lichte laufender Theoriediskussionen veranschaulichen.