

KI-Realitäten/AI Realities

Richard Groß & Rita Jordan

KI-Realitäten

Der vorliegende Band präsentiert verschiedene Perspektiven auf Aspekte der realweltlichen Einbettung von maschinellern Lernen (ML). Die hier versammelten Beiträge gehen auf Vorträge zurück, die auf einer vom Schaufler Lab@TU Dresden ausgerichteten interdisziplinären Online-Tagung im Dezember 2021 gehalten wurden. Der Titel des Bandes, *KI-Realitäten*, verweist einerseits auf das Ausmaß, in dem ML-basierte Technik als schon seit mehr als einem Jahrzehnt dominanter Ansatz sogenannter Künstlicher Intelligenz (KI)¹ die Wirklichkeit zu prägen begonnen hat. Andererseits spielt er auf die reflexiven Herausforderungen an, die sich theoretischen Bemühungen um ein adäquates sozial- und geisteswissenschaftliches sowie philosophisches Verständnis von KI stellen. ML stellt in Form praktischer Anwendungen wie auch als Gegenstand spekulativer Imagination einen wichtigen Faktor und zugleich ein Produkt der Realitäten dar, in die es als Technik und Vorstellung eingebettet ist. Die folgenreiche Integration von auf ML-Verfahren beruhender Technik in bestehende Realitäten macht es erforderlich, die einschneidenden und teils widersprüchlichen Effekte ihrer Verbreitung in den Blick zu nehmen. Das als Titel gewählte Kompositum soll daher nicht allein als Ausdruck der gegenwärtigen Größenordnung realweltlicher Auswirkungen von ML-Anwendungen verstanden werden, sondern vor allem auf die in den Auswirkungen sichtbar werdende Vielzahl unterschiedlicher Aspekte von ML als Forschungsgegenstand hinweisen.

1 Den Begriff Künstliche Intelligenz nutzen wir im Folgenden, sofern nicht anders angegeben, synonym mit maschinellern Lernen, da ML-Ansätze maßgeblich verantwortlich sind für die wesentlichen Fortschritte jüngerer Datums in der KI-Forschung wie auch in der Anwendung von KI.

Aufgrund dieser Vielzahl solcher, sich mitunter zudem wechselseitig bedingender Aspekte scheint uns eine behutsame Annäherung an die theoretische Auseinandersetzung mit ML ratsam zu sein. So birgt die Auseinandersetzung mit einer – noch dazu von lautstark vernehmbaren Kontroversen begleiteten – vermeintlichen ›Schlüsseltechnologie‹ die Gefahr, einem Gegenstand, der als ergiebiges Spekulationsobjekt immerzu von interessengeleiteten Projektionen überzogen und dabei womöglich mit vielem verwechselt wird, vorschnell bestimmte Eigenschaften zuzuschreiben. Eine Einbettung in historisch gewachsene gesellschaftliche Realitäten bedeutet für jedes noch so technisch fixiert anmutende Phänomen, dass es in dynamischen Beziehungen wechselseitiger Abstimmung und Beeinflussung mit anderen sozialen Entitäten steht. Diese Annahme legt für geistes- und sozialwissenschaftliche Forschung nahe, von der praktischen Situiertheit maschinellen Lernens auszugehen, um deren Bedeutung innerhalb gelebter Realitäten nachzuvollziehen und damit technikwissenschaftlich gewonnene Gegenstandsbestimmungen kritisch reflektierend zu ergänzen. An solche Bestimmungen angelehnt beruht ML – schematisch formuliert – auf dem ›Training‹ von Algorithmen durch die Verarbeitung von (zumeist sehr umfangreichen) Datensätzen. Dieser Vorgang bezeichnet die stochastische Modellierung von Daten, die in der Folge auf Basis des erstellten Modells weitere Daten(-analysen) produzieren können. Die Bedeutung der so produzierten Daten ist oft schwer zu erschließen. Dies gilt selbst und gerade auch dann, wenn die Eigenschaften des Datensatzes im Sinne formaler Verteilungen und Regelmäßigkeiten als bekannt und verstanden gelten (Dourish 2016: 7). Die mitunter überraschenden und unheimlich anmutenden Outputs der Modelle (Bucher 2017) sind meist nicht ohne Weiteres erklärbar und lassen sich hinsichtlich ihrer Implikationen aufgrund der Opazität der rechnerischen Modellierung oft auch im Nachhinein nicht erschließen (Burrell 2016).

Eine wesentliche Herausforderung für sozial- und geisteswissenschaftliche wie philosophische Forschung besteht daher in der adäquaten Interpretation von mit ML-Verfahren generierten Outputs wie etwa Texten und Bildern, besonders im Hinblick auf ihre Genese. Als sinnhafte Beiträge zu gesellschaftlichen Realitäten werden diese häufig zum Gegenstand von Kontroversen. Kaum nachvollziehbare Resultate von Mustererkennungen und -generierungen verweisen auf die nicht intuitiv erschließbare operative Logik der Verfahren. Die auf dieser Logik beruhende nichttriviale Transformation digitaler Inputdaten in sinnförmig rezipierte Outputs vollzieht sich auf eine dem menschlichen Denken fremde Weise. Dies stellt eine Hürde für die geis-

tes- und sozialwissenschaftliche Annäherung an KI dar, die sich in mancherlei Hinsicht als »alien subject« (Parisi 2019) erweist.

Daran anschließend besteht eine weitere Herausforderung für die Forschung zu ML im weiteren Sinne in der Bestimmung der Beziehung zwischen gesellschaftlichen Realitäten und technischen Arrangements, die als KI bezeichnet werden. In angewandter Form ist ML-basierte Technik innerhalb der Realitäten, die sie hervorgebracht haben, allgegenwärtig geworden und trägt nun ihrerseits zur Veränderung wie Reproduktion dieser Realitäten bei. Dies heißt für Theorien maschinellen Lernens, dass sie zugleich Theorien über die Wirklichkeit voraussetzen und im Umkehrschluss auch selbst als solche lesbar sind. Gegenwärtige ML-Anwendungen sind dazu in der Lage, datenförmig verfügbare Text- und Bildwerke zu verarbeiten wie auch selbst zu generieren. Sie vermitteln Sinn und mithin Wirklichkeit, was zur Schaffung neuer Welten und zu neuen Formen des Umgangs mit diesen beitragen kann. Diesem Umstand haben sozial- und geisteswissenschaftliche Theorien maschinellen Lernens Rechnung zu tragen.

KI-Pathologien

Innerhalb dieser Gemengelage gibt es zahlreiche Stimmen, die zu Recht auf die ideologischen Funktionen hinweisen, die mystifizierende Deklarierungen von KI als für Menschen unergründlich und nicht verstehbar erfüllen. Eine solche Rhetorik, so etwa Galloway (2021), überhöhe die Komplexität und Leistungsfähigkeit der technischen Verfahren und lenke von der Tatsache ab, dass ML letztlich nicht viel mehr sei als ein »fancy way to calculate an average« (ebd.). Der anhaltende Hype stelle im Grunde einen »total scam« (ebd.) dar, der die tatsächlichen sozialen Probleme, die sich mit der Verbreitung von ML-Anwendungen ergeben, in den Hintergrund treten lasse.

»KI« ist das Resultat kontingenter und umkämpfter historischer Entwicklungen und kann als Ausdruck derselben verstanden werden. Die realen Auswirkungen ihrer Anwendung werden unter anderem von ökonomischen Interessen bestimmt, die aus den Strukturen und Dynamiken einer bestimmten Gesellschaftsordnung erwachsen. Da Outputs aus Modellen generiert werden, die auf durch menschliche Arbeit produzierten Datensätzen beruhen, meint ML praktisch die Extraktion bestehenden Wissens (vgl. Joler/Pasquinelli 2020), nicht zuletzt auch in der Extrapolierung darauf beruhender Prädiktionen. Es setzt sich in soziologischer Betrachtung aus einer Reihe von materiellen Praktiken zusammen, deren Organisation in mancher Hinsicht

dem für den Industriekapitalismus typischen Fließbandprinzip entspricht (vgl. ebd.: 2). ML-Produkte verweisen auf die mühsame Arbeit, die mit der Erhebung und Aufbereitung von Daten einhergeht. Eine solche Sichtweise unterstreicht zudem den Ausbeutungscharakter des soziotechnischen Arrangements, das für das Training und die Anwendung eines Modells vorausgesetzt werden muss. In diesem Sinne operiert ML akkumulativ und stellt keine radikale Transformation dar (vgl. Mackenzie 2017). Seine soziale Einbettung impliziert die Abhängigkeit von und die Integration in gesellschaftliche Strukturen, innerhalb derer es Wirksamkeit entfalten kann.

ML-Anwendungen, so lässt sich beobachten, wirken in bestehenden Realitäten dabei häufig derart, dass sie deren pathologische ideologische Tendenzen reproduzieren, wenn nicht sogar verstärken. Zu diesen Tendenzen gehören die Diskriminierung marginalisierter sozialer Gruppen (vgl. Noble 2018; Apprigh et al. 2018; Chun 2021), die Konzentration von Kapital und politischer Macht bei Plattformen, über die Oligarchen walten (vgl. Whittaker 2021), die Zunahme vernetzter datenbasierter Formen sozialer Kontrolle, die sich mit dem Bild des distribuierten »Polyopticon« (Sherman 2022) beschreiben lassen, sowie die Missachtung von Nachhaltigkeitsfragen, die sich in den von ML-Technik verursachten Umweltbelastungen äußert (vgl. Strubell et al. 2019; Dhar 2020; Bender et al. 2020). Diese Probleme – wenn auch nicht auf ML beschränkt, sondern gleichermaßen für andere digitale Technologien charakteristisch – sind zentral für die realen Auswirkungen von ML-Anwendungen auf soziale Welten und deren biologische, geologische wie klimatische Umwelten auf dem Planeten Erde. »Your computer« – so »smarter auch erscheinen mag – »is on fire« (Mullaney et al. 2021), und dies nicht erst seit gestern.

Maschinenlogik und ihre Folgen

Unser Fokus in diesem Band liegt auf theoretisch herausfordernden und mitunter bislang unterbelichteten Aspekten maschinellen Lernens. Uns interessieren insbesondere jene Merkmale des Phänomens, die es als eigenständigen Forschungsgegenstand auszeichnen; dies im Gegensatz zu Facetten von ML-Anwendungen, die wesentlich strukturell prädestiniert sind und sich etwa primär aus den wirtschaftlichen und politischen Dynamiken erklären, die den »KI-Hype« der letzten zehn Jahre angetrieben haben. Von derartigen Auseinandersetzungen versprechen wir uns Perspektiven, die ein erweitertes

Verständnis und neue Potenziale für den Umgang mit ML-basierter Technik aufscheinen lassen. Dieser Band stellt eine Reihe solcher Perspektiven vor.

Einige der Beiträge des Bandes befassen sich mit Fragen nach einem angemessenen Verständnis der sich in ML-Verfahren manifestierenden »computational reason« (Cavia 2022) sowie den praktischen Auswirkungen ihrer realen Umsetzung. Über theoretische Bemühungen um eine begriffliche Abgrenzung dieses Konzepts von gängigen, in der Regel anthropozentrischen Verständnissen von Vernunft hinaus finden sich ebenso Analysen der praktischen Rolle und Bedeutung konkreter materieller Manifestationen »komputationaler Vernunft«, insbesondere im Hinblick auf ihre nichtintendierten Nebenfolgen. Effekte, die von den Absichten und Erwartungen der Anwendungsentwickler:innen und -nutzer:innen abweichen, sind angesichts der Vielzahl und Größenordnung von Kontexten, in denen ML-Anwendungen stattfinden, keine Seltenheit, sondern ein wesentliches Charakteristikum (vgl. Broussard 2018). ML-Anwendungen haben ihren Anteil an der Emergenz neuartiger kultureller Dynamiken, die nicht zuletzt auf dem »epistemischen Schock« beruhen, den die Eigenlogiken maschinellen Lernens im Zuge ihrer Verbreitung gerade deshalb auslösen, weil sie zur Vermittlung und Produktion symbolvermittelter Sinngehalte bzw. von Bedeutung eingesetzt werden (vgl. Roberge/Castelle 2021: 2, 7).

Unser Interesse gilt insbesondere Entwicklungen im Zusammenhang mit ML-Anwendungen, die verändernd auf Handlungs-, Organisations- oder Denkweisen, mithin auf die Möglichkeitsbedingungen menschlicher Vergesellschaftung einwirken. In seiner Keynote im Rahmen jener Tagung des Schaufler Lab@TU Dresden Ende 2021, auf die dieser Band zurückgeht, schlug Matteo Pasquinelli vor, maschinelles Lernen als Wissensmodell im Sinne einer politischen Epistemologie zu verstehen. Seinen Ausführungen zufolge trägt ML zur Errichtung eines »epistemischen Gerüsts« bei, das die ideologische Form, die logische Form, die technische Form und die soziale Form in einen integrierten Zusammenhang bringe. Pasquinellis Vortrag unterstrich die Multidimensionalität maschinellen Lernens als ein gesellschaftliches Phänomen, das sich aus dem dynamischen Zusammenspiel von Mythologie, kollektiver Imagination, statistischem und mechanischem Denken, Rechenprozessen und der Automatisierung von Arbeit sowie der Überwachung und Kontrolle von sozialem Verhalten speise. Als realweltlicher Zusammenhang stelle es uns vor teils grundlegende Veränderungen in der Art und Weise, wie Wissen produziert, vorgestellt, navigiert und verteilt wird.

Diese Veränderungen haben einschneidende praktische Konsequenzen. Wie Louise Amoore (2020: 40f.) ausführt, sind ML-Anwendungen dazu in der Lage, subtile, manchmal gänzlich unsichtbare Veränderungen in der menschlichen Weltwahrnehmung zu bewirken. Ihr zufolge verschieben algorithmusbasierte experimentelle Datenverarbeitungsverfahren die Schwellen des für menschliche Beobachtung Zugänglichen. ML-basierte »Computer Vision« könne zwar nicht selbst im eigentlichen Sinne »sehen«, doch spiele sie eine wichtige Rolle dabei, was für Menschen sichtbar ist. Die Weltwahrnehmung hänge daher zunehmend davon ab, was von einem ML-vermittelten Sichtbarkeitsregime für menschliche Beobachtung verfügbar gemacht wird. Diese Überlegung zielt weniger auf den Generalverdacht, dass mit Computer Vision in erster Linie ein Verschleierungsinstrument am Werk sei, vielmehr geht es Amoore um die Feststellung, dass technisch vermitteltes Sehen (und mithin auch Erkenntnis im weiteren Sinne) das Ergebnis selektiver Prozesse ist, die zu kontingenten und mitunter schwer verstehbaren Resultaten führen.

Maschinelles Lernen als theoretisches Problem

Unser Interesse an ML ergibt sich ebenso aus dem Umstand, dass die realweltlichen Effekte von dessen Anwendung die Ausrichtung einiger Grundbegriffe der Geistes- und Sozialwissenschaften infrage zu stellen vermögen. Betroffen sind davon solch zentrale Konzepte wie Subjektivität, Interaktion, Kognition oder Kommunikation. Wie Luciana Parisi (2019) feststellt, handle es sich bei KI um ein »alien subject«, das zugleich unsere Vorstellung davon, was unter menschlichen Subjekten zu verstehen sei, auf die Probe stelle. Die zunehmende Automatisierung kognitiver Arbeit, bei der ML-Verfahren eine zentrale Rolle spielen, problematisiere das menschliche Selbstverständnis. Seinem Privileg der alleinigen Entscheidungsgewalt in vielen sozialen Zusammenhängen beraubt, werde das menschliche Subjekt Zeuge der Ausdehnung eines »alien space of reasoning« (ebd.: 30), die in einer »crisis of conscious cognition« (ebd.: 28) resultiere. Und die mit »machine thinking« einhergehenden Unwägbarkeiten vervielfachen sich gewissermaßen, wenn vernetzte Maschinen in einem ihnen eigenen »space of communication opaque to human vision« (ebd.: 31) aktiv werden. Fraglich wird vor diesem Hintergrund nicht nur, was »Denken« impliziert, welche Rolle es in sozialen Prozessen spielt und wessen Privileg es ist, zu denken, sondern auch geläufige Vorstellungen von Kommunikation geraten unter den Verdacht, irreführenderweise zu anthropomorph konzipiert zu sein.

Elena Esposito (2022) versucht mit ihrer systemtheoretischen Konzeption »künstlicher Kommunikation« den merkwürdigen Charakter der Begegnung von Menschen und ML-basierten sozialen Agenten analytisch adäquat zu beschreiben. In ihrer Keynote auf unserer Tagung im Dezember 2021 führte sie die wesentlichen Grundannahmen dieses Ansatzes aus. Es sei, so Esposito, eine fehlgeleitete Vorstellung, dass ML die Eigenschaften menschlichen Verhaltens, Denkens oder Kommunizierens reproduziere. Stattdessen seien dessen Anwendungen gerade dann am erfolgreichsten, wenn sie für Aufgaben eingesetzt werden, in denen es nicht um die Imitation menschlichen Bewusstseins oder Denkens geht. ML-Technik könne sich an Kommunikation beteiligen, indem sie Kontingenz »virtuell« verarbeite. Wenn Nutzer:innen mit »ausgereiften« Bots interagierten, sähen sie sich mit einer Kontingenz konfrontiert, die sie sich nicht selbst zurechnen können; andererseits lasse sich das kommunikative Geschehen nicht hinreichend als alleinige Projektion der User:innen erklären. Zugleich, so Esposito, stünden ML-Beiträge ebenso wenig für die Kontingenz der Maschine, da diese allein auf fremde Referenzen verweise. Ein ML-Modell wäre nichts ohne einen (auf menschlicher Kommunikation beruhenden) Satz von Trainingsdaten, deren Bedeutung und Ursprung oft unklar seien. Das Modell präsentiere den Nutzer:innen intransparent verarbeitete und daher auf kaum nachvollziehbare Weise reflektierte Perspektiven anderer Nutzer:innen, mithin deren Kontingenz, die vom Modell verarbeitet und damit virtualisiert werde. Virtuelle Kontingenz konstituiert Esposito zufolge daher einen Modus von Kommunikation, der insofern ohne distinkte Alterität operiert, als sich hinter ML-Beiträgen nicht ohne Weiteres eine kompakte Adresse ausmachen lässt, der etwa Autor:innen-schaft und damit verbundene Absichten oder Verantwortung zugeschrieben werden könnten. Dies deutet auf eine Verschiebung der Grundbedingungen von Sozialität hin. »Artificial Communication« steht für einen sozialen Operationsmodus, der bisher als Privileg der Beziehungen zwischen Menschen galt: sprachbasierte Kommunikation, aber eben »künstliche«, wie Esposito es ausdrückt – ein Attribut, das hier in erster Linie für das Fehlen eines vertrauten Elements von Kommunikation steht, nämlich einer Adresse, der Kommuniziertes als Handlung zugeschrieben werden kann.

Es zeigt sich hier ein doppeltes Problem, denn die Verstehbarkeit von Kommunikation hängt allgemein davon ab, dass sie einer kompakten sozialen Adresse – etwa einem menschlichen Individuum – zugerechnet werden kann, was wiederum Voraussetzung für die Möglichkeit der Zuschreibung von Verantwortung ist. Adressierbarkeit – in technischer Hinsicht ein wesent-

liches Merkmal aller Computertechnik (vgl. Dhaliwal 2022) – ist angesichts der ML-spezifischen Opazität auf sozialer Ebene demzufolge keine Selbstverständlichkeit mehr (vgl. Burrell 2016). Es verwundert daher nicht, dass viele Bemühungen in der KI-Forschung dem Anliegen gewidmet sind, diese »responsibility gap« (Matthias 2004) zu schließen. So soll etwa das Problem mangelnder Transparenz durch die Entwicklung von »Explainable-AI«-Methoden und -Verfahren gelöst werden, die den Bedürfnissen der an den ML-Anwendungen beteiligten »stakeholders« nach einer besseren Interpretierbarkeit der Outputs gerecht zu werden versuchen (vgl. Zednik 2021). Bei allen Fortschritten in diesem Forschungsfeld ist bis heute bei vielen – wenn nicht den meisten – ML-Anwendungen selten klar, was ein Modell dazu veranlasst hat, einen bestimmten Output zu erzeugen. Diese womöglich konstitutive Lücke lässt Raum für Projektionen und Spekulationen. KI – weniger in der gegenwärtigen, deutlich häufiger in antizipierten, zukünftigen Formen – wird nicht selten ein sublimier Status zugeschrieben (vgl. Ames 2018); man denke nur an die verschiedenen Varianten des »Superintelligenz«-Diskurses (vgl. etwa Bostrom 2014). Andere Autor:innen wiederum amüsieren sich angesichts der Banalität des Scheiterns vieler ML-Anwendungen über Spekulationen nahender technischer Singularität; Anekdoten über »Artificial Unintelligence« (Broussard 2018) oder »Artificial Stupidity« (Steyerl 2017; vgl. auch Mackinnon 2017) gibt es zuhauf.

Wir sehen diese vielfältigen, sich teils widersprechenden Positionen mit Blick auf die in diesem Band versammelten Annäherungen an verschiedene Aspekte maschinellen Lernens als Ausdruck von dessen Vielschichtigkeit und Ambivalenz als Forschungsgegenstand. In diesem Sinne scheint es naheliegend, ML als eine Chimäre zu verstehen, deren Wesen jenseits konventionell geltender Dichotomien liegt, wie Ilan Manouach und Anna Engelhardt (2022: 9) einleitend über das Ungeheuer aus der griechischen Mythologie schreiben, das für die von ihnen vorgelegte »Inventory of Synthetic Cognition« titelgebend ist. Maschinelles Lernen kann gleichzeitig Verarbeitung von Daten, Produktion von Sinn, Verhandlung sozialer Normen und Machtausübung meinen. Es findet in so vielen Bereichen Anwendung, nimmt dabei so variable Gestalten an und erfüllt derart unterschiedliche Funktionen, dass es angesichts der Vielfalt all dieser Aspekte naheliegt, grundlegende Fragen seiner Bestimmung als Forschungsgegenstand zu reflektieren. Es ist uns mit diesem Band ein Anliegen, einen Schritt zurückzutreten, um ML in realweltlich eingebetteter Form im Hinblick auf die benannten Herausforderungen gegenstandsangemessen theoretisch begegnen zu können.

ML-basierte Technik hat sich in so vielen Bereichen verbreitet und ist ein solch integraler Bestandteil des gesellschaftlichen Lebens geworden, dass sie in ihrer weitgehend selbstverständlichen und zumeist als unproblematisch erlebten Allgegenwärtigkeit heutzutage »unsichtbar« zu werden scheint. Eine solche »Transparenz«, um eine Einsicht aus Susan Leigh Stars und Karen Ruhleders (2017) instruktiver Forschung zu »großangelegten Informationsräumen« aufzugreifen, ist für die soziale Funktionalität jeder technischen Infrastruktur charakteristisch – solange sie funktioniert. Der zunehmend infrastrukturelle Status von ML-Technik spricht für ihre weitgehende Akzeptanz, was angesichts der damit einhergehenden Normalisierungs- und Habitualisierungsprozesse die Frage aufwirft, wie ML als Gegenstand kritischer Beobachtung und Analyse – Voraussetzungen jedweder adäquaten Theoriebildung – verfügbar gemacht und gehalten werden kann.

Dimensionen maschinellen Lernens: Modelle – Praktiken – Topologien

Für einen angemessenen theoretischen Umgang mit den oben dargelegten Herausforderungen schlagen wir einen analytischen Rahmen vor, der drei verschiedene Dimensionen der geistes- und sozialwissenschaftlichen wie philosophischen Auseinandersetzungen mit ML identifiziert. Die Unterscheidung von Modellen, Praktiken und Topologien ermöglicht es uns, einige wesentliche Aspekte der realweltlichen Einbettung und Situierung von ML in den Blick zu nehmen. Wir nutzen dieses Differenzierungsschema als den vorliegenden Band strukturierende Heuristik, um Besonderheiten der in den Beiträgen entwickelten Perspektiven hervorzuheben. Gleichmaßen erlaubt uns dieses Schema, Gemeinsamkeiten der unterschiedlichen Ansätze zu unterstreichen, die womöglich auf verallgemeinerbare Charakteristika des Phänomens verweisen.

Die Betrachtung von *ML als Modell* – bezogen auf dessen empirische Beobachtung wie auch auf die Reflexion des sich darin vollziehenden Umgangs mit Wissen – lenkt den Blick auf epistemologische Implikationen. ML modelliert realweltliche Phänomene datenförmig und vermittelt damit die Realitäten, deren Teil es zugleich selbst ist. Um die zentrale Bedeutung dieses Aspekts auf eine pointierte Formel zu bringen: »The model is the message« (Bratton/Agüera y Arcas 2022). Nach den spezifischen Möglichkeitsbedingungen und Implikationen von ML-basierter Modellbildung zu fragen heißt, sich wortwörtlich dem

Charakter des ›Lernens‹ wie auch dessen ›maschineller‹ Logik theoretisch anzunähern.

Die Auswirkungen von ML-Anwendungen hängen andererseits von der konkreten praktischen Situierung der Modelle ab (vgl. Groß/Wagenknecht 2023). *ML als Praxis* erfordert eine eigenständige Betrachtungsebene, die zwar mit den Charakteristika der Modelle verknüpft ist, über diese jedoch gleichwohl hinausweist. Maschinelles Lernen als Praxis zu verstehen, erlaubt Einblicke in die Differenz zwischen Modellannahmen und ihrer praktischen Umsetzung. Es ist genau diese Unterscheidung, anhand derer sich beschreiben lässt, was Einbettung von ML in bestehende kulturelle, wirtschaftliche und politische Realitäten heißt und welche Dynamiken und Interdependenzen dabei zum Tragen kommen.

Im Bemühen um ein adäquates Verständnis der oben beschriebenen nicht-menschlichen und mithin »alien« (Parisi 2019) Gegenstandseigenschaften erachten wir schließlich *ML-Topologien* als eine dritte wichtige Analysedimension der Theoriebildung. Rechenoperationen in vieldimensionalen Vektorräumen – in denen sich Modelle als Ausdruck der Wechselbeziehungen zwischen einzelnen Datenpunkten herausbilden – haben ihre eigenen Dynamiken und folgen spezifischen Logiken. Ein angemessenes Verständnis ihrer Eigenschaften kann vermitteln, was das »Denken« oder die »Kognition« mutmaßlich intelligenter Maschinen ausmacht. Tatsächlich könnte der Vektorraum in rechnerischer Hinsicht als das Medium maschinellen Lernens verstanden werden, da seine Eigenschaften die Möglichkeitsbedingungen von ML-Technik bestimmen. Aus diesem Grund sehen wir in der Auseinandersetzung mit »computational topologies« einen wesentlichen Ansatzpunkt, um die Annahme, dass die operative Logik von ML-Verfahren prinzipiell unverstanden bleiben müssen, überwinden zu können. Wir sind zuversichtlich, dass sich ein angemessenes Verständnis der topologischen Implikationen maschinellen Lernens produktiv auf geistes- und sozialwissenschaftliche wie philosophische Auseinandersetzungen mit KI auswirken wird.

Praktische Relationalität, strukturelle Interdependenz, kooperative Generativität

Die Beiträge des Bandes durchziehen – über die drei Betrachtungsdimensionen Modell, Praxis und Topologie hinweg – mehrere Leitmotive zur Charakterisierung der realweltlichen Einbettung maschinellen Lernens.

Ein erstes wiederkehrendes Motiv ist die Betonung der Relationalität heterogener sozialer Entitäten in ML-Praktiken. Die wechselseitige Beeinflussung von Menschen, Maschinen, Dingen, Infrastrukturen und anderen für die jeweilige Praxis relevanten sozialen Entitäten erfolgt jeweils unter konkreten situativen Bedingungen, die spezifische Dynamiken mit sich bringen. **Hannah Link** (S. 143-168) stellt mit Blick auf den Umgang von Robotikforscher:innen mit ihren Prototypen fest, dass sich nicht nur die Leistungsfähigkeit von Robotern weiterentwickelt, sondern sich dabei auch eine Form von »posthumane[r] Interaktion zwischen Wissensobjekt und -subjekt« (S. 143) entsteht, die sich im Rückschluss auf die Vorstellungen von Menschlichkeit auswirkt. Der Frage nach einem adäquaten theoretischen Ansatz zur Beschreibung der Entwicklung von Kommunikationsfähigkeit zwischen Menschen und Maschinen begegnen **Yaoli Du und Nadine Schumann** in ihrem Beitrag (S. 193-207) mit dem Vorschlag, die interaktive Herausbildung gemeinsamer pragmatischer Handlungsmuster in einer geteilten Welt als zentral zu betrachten. Auch **Jonathan Harth und Maximilian Locher** nehmen mit dem Programm der Relationsmuster praktische Beziehungsaspekte von Mensch und Maschine in den Blick und beleuchten, wie diese auf die Identitäten der beteiligten Entitäten zurückwirken und sie kontinuierlich neu generieren (S. 169-191). Ein empirisches Beispiel für ein solches interaktives Verhältnis ist der »e-rater«, den **Jan Georg Schneider und Katharina Zweig** untersuchen (S. 93-111), um anhand dieses ML-basierten Bewertungssystems die Potenziale und Kontingenzen der automatisierten Beurteilung von Texten auszuloten. Die Beiträge begegnen sich im gemeinsamen Bezug auf postanthropozentrische Formen des Sozialen und explorieren Erweiterungen von Sozialtheorien angesichts neuer mit ML verbundener Problemlagen.

Die kritische Reflexion der Interdependenzverhältnisse innerhalb der wirtschaftlichen und politischen Zusammenhänge, in denen ML-basierte Technologien entwickelt und eingesetzt werden, stellt ein weiteres Leitmotiv dar, das wiederkehrend thematisiert wird. Modellbasierter ML-Technik – etwa automatisierten Sprachverarbeitungssystemen – sind auf folgenreichere Weise Herrschaftsverhältnisse eingeschrieben. Diese Beobachtung nimmt **Christian Heck** zum Anlass, *adversarial hacking* als subversive ästhetische Strategie für die partizipative Öffnung von hegemonial strukturierten KI-Sprachmodellen praktisch vorzustellen (S. 235-286). **Jan Fuhrmann** nähert sich dem Problem systemtheoretisch und findet eine Übersetzungslücke zwischen den für die Grammatiken der Diskriminierung blinden autopoeitischen algorithmischen Systemen und den ihrer Umwelt entspringenden semanti-

schen Gehalten (S. 115-141). Mit einem kritischen Blick auf die Entwicklung sogenannter Large Language Models (LLMs) durch Google/Alphabet verhandeln **Jonathan Roberge und Tom Lebrun** die hermeneutischen Implikationen der jüngeren Entwicklungen im Feld automatisierter Textgenerierung sowie die Machtverhältnisse, innerhalb derer diese Entwicklungen stattfinden (S. 39-65). Die Autoren kommen zu dem Schluss, dass die ökonomischen Rahmenbedingungen kaum Anreize für eine hinreichende Sorgfalt bei der qualitativen Kuratierung von Datensätzen schaffen und stattdessen die schnellstmögliche Skalierung von Sprachmodellen befördern, was häufig eine Verstärkung von nichtintendierten Effekten – nicht selten solchen diskriminierender Art – zur Folge hat. Einen Versuch, die Schnittstellen von Ethik, Epistemologie und Politik zu bestimmen, unternimmt **Catriona Gray** am Beispiel von Empfehlungssystemen, die zur Unterstützung staatlicher Entscheidungspraxis genutzt werden (S. 67-92). Ihr Beitrag untersucht die Umstände und Konsequenzen des Einsatzes von ML-Anwendungen in der öffentlichen Verwaltung und erörtert, wie dadurch vorherrschende Gerechtigkeitskonzepte auf den Prüfstand gestellt werden.

Ein dritter Aspekt, den die Beiträge des Bandes aufgreifen, ist die praktische kooperative Generativität von Menschen und ML-Verfahren. Dieses Leitmotiv greift Diskussionen zur Autor:innenschaft in soziotechnischen Systemen auf und lenkt den Fokus auf spezifische Dynamiken, die kreative Prozesse unter Beteiligung von ML-Technik kennzeichnen. **Miriam Akkermann** nähert sich dem Zusammenhang anhand einer Untersuchung der musikalischen Freiräume, die die Einbindung unterschiedlicher Formen von KI in Kompositionsprozesse erlaubt (S. 315-329). Der Beitrag von **Jakob Claus und Yannick Schütte** (S. 211-233) widmet sich den Dynamiken in der literarischen Ko-Produktion von Autor:in und automatischem Textgenerierungsverfahren am Beispiel des Buches *Pharmako-AI*, das K Allado-McDowell (2021) gemeinsam mit dem Sprachmodell GPT-3 verfasst hat. **Michael Klipphahn-Karge** untersucht die Nuancen verschiedener maschineller Ästhetiken aus kunstwissenschaftlicher Perspektive und entwickelt ein analytisches Kategorienschema zur differenzierten Auseinandersetzung mit sogenannter KI-Kunst (S. 287-314). **Fabian Offert** schlägt mit Blick auf zeitgenössische ML-basierte Kunst vor, diese aufgrund des subtraktiven Charakters der bei ihrer Entstehung genutzten multi-modalen Bildsynthese als Skulptur zu verstehen (S. 273-286).

Die letzten beiden Beiträge des Bandes setzen einer reduktionistischen Sicht auf »computational reason« Positionen entgegen, die Potenziale topologisch informierter Perspektiven auf das Verständnis der generativen Ei-

genlogik maschinellen Lernens erörtern. **AA Cavia und Patricia Reed** (S. 351-363) folgen ausgehend von Ansätzen der konstruktivistischen Mathematik der topologischen Annahme, dass unter der nichteuklidischen Voraussetzung der Möglichkeit räumlicher Pluralität jeder Raum eine in ihn eingebettete begleitende Struktur mit sich führt, ohne dieser vorausgehen zu müssen. Die Autor:innen loten sodann die Möglichkeiten eines begrifflichen Vokabulars aus, das die Domänen des Diskreten und des Kontinuierlichen – der Algebra und der Geometrie – zusammenführt, um eine Erweiterung des »inferential toolkit available to computational reason« (S. 361) zu ermöglichen. **Lukáš Likavčan und Carl Christian Olsson** (S. 333-349) setzen sich in ihrer topologischen Annäherung an ML kritisch mit Immanuel Kants Analogie zwischen der Orientierung im Denken und der geografischen Orientierung auseinander. Ihr Aufsatz nimmt die topologischen Merkmale von künstlichen neuronalen Netzen zum Anlass, unser Verständnis von Denken mit Blick auf die Möglichkeiten der sich in Deep Learning-Verfahren manifestierenden räumlichen Konstellationen zu reflektieren.

AI Realities

The title of this volume alludes to the impact that technologies based on machine learning (ML) have had across countless domains of social life in recent years. *AI Realities* also hints at the reflexive challenges in endeavors to theorize so-called artificial intelligence (AI) as a real-world phenomenon.² We chose this title as an expression of the ambivalent tensions inherent to ML both as a transformative technology and as a challenging research subject. ML's increasing practical integration into social life necessitates focusing on its varied and partially conflicting characteristics as both a factor and a product of the realities in which it is embedded. Real-world embeddedness implies relationships of mutual influence and attunement and hence asks us to consider how ML is shaped as a codependent entity in relation to other entities and structures. This volume is dedicated to exploring the theoretical challenges that stem from the difficulties in understanding ML in terms of its multifaceted ways of appearing in the world. Its chapters are based on talks given at an interdisciplinary conference organized by Schaufler Lab@TU Dresden that took place online in December 2021.

ML's various forms of real-world appearance include speculations as well as practical applications. Its impact is driven by its functional capacities as a technology just as much as by its potential for imagination. In this vein, we approach theorizing ML carefully in a way that avoids misattributing properties to it that actually explain a different phenomenon or the dynamics of the interplay between the two. This points to the challenge of determining the nature of the relationship between AI and the realities that produce it. In applied form, ML has become a ubiquitous part of the realities it has emerged into, now contributing to the reproduction and change of those realities. In its capacity to make meaning via language or image processing and generation, ML can contribute to the world in generative ways and allows for the making of new worlds and new forms of worlding. On the one hand, this disposition implies that theorizing ML presupposes theories of the worlds it has started to inhabit. On the other hand, it suggests theories of ML in the humanities should pay close attention to its practical situatedness to grasp its relevance and effects within lived experiences. Such an approach puts research on ML in the humanities in

2 We use the terms machine learning and artificial intelligence interchangeably in the following – unless indicated otherwise – to describe the technology in question, given ML's dominance in AI research and application in the past decade.

a position to critically engage with conceptualizations of the subject in disciplines such as data science and engineering as well as the adjacent technology industries.

Drawing on these conceptualizations, ML – schematically put – relies upon the ‘training’ of algorithms to build a model through the processing of a dataset compiled to recognize patterns contained in the data. ML-based pattern recognition essentially produces stochastic data analyses. While these analyses are often well understood in terms of the formal patterns and regularities of a given dataset, they are frequently hard to grasp, if not unknowable, in terms of the domain the data represents (Dourish 2016, 7). Hence, the often surprising and “uncanny” output (Bucher 2017) produced by such models is not easy to make sense of. This is not least due to the opacity of the computational modeling: “When a computer learns [...], it does so without regard for human comprehension” (Burrell 2016: 10). From this follows another challenge that sits in the adequate interpretation of ML’s contributions to the world. ML-generated outputs can be harmful in their real-world effects. Therefore, the models frequently become subject to controversies, in part because a model’s ‘reasoning’ can not be fully accounted for. ML’s hard-to-grasp non-human characteristics continue to complicate its understanding in the humanities and beyond.

Pathologies of AI

Many researchers in the humanities tend to dismiss the relevance of ML’s alienness and highlight instead how emphasizing its complexity mystifies the ideological nature of such rhetoric. They point out, for instance, that it basically stands for “a fancy way of calculating an average” and is, in effect, a “total scam” (Galloway 2021) that distracts from the real challenges resulting from the proliferation of applications associated with “AI”. ML has not spontaneously emerged into the world and its real-world impact is, among other factors, driven by economic interests expressing the structures of a particular social order. It performs “knowledge extractivism” (Joler/Pasquinelli 2020) by generating information from the models its productivity depends upon. The knowledge contained in these models is based on datasets produced by human labor. ML, as a set of material practices, takes the shape of an assembly line (p. 2), resembling a form of organizing labor typical of industrial capitalism precisely because it relies on that labor. ML often depends on laborious efforts that allow data to be gathered and compiled into a dataset from which a model is generated that can finally be applied for a dedicated purpose. In this sense, ML

“is an accumulation rather than a radical transformation” (Mackenzie 2017: 5) – its social embeddedness implies dependence on and integration into the social forms it appears in.

From this, it follows that ML leaves its mark on reality by often reproducing, if not amplifying, some of the latter’s pathological political and ideological tendencies. Such tendencies include the discrimination against and exclusion of marginalized people (Noble 2018; Apprigh et al. 2018; Chun 2021), the concentration of power in the hands of oligarch-owned platforms (Whittaker 2021), the rise of networked, data-centric forms of social control, constituting a distributed “polyopticon” (Sherman 2022), and the disregard of environmental issues (Strubell/Ganesh/McCallum 2019; Dhar 2020; Bender et al. 2020). These problems, though not exclusive to ML but characteristic of many computational technologies, are central to its impact on social worlds and the planet. “Your computer” – however intelligent it may seem – “is on fire” (Mullaney et al. 2021), and has been for a long time.

Machine logic and its consequences

Our specific focus in this volume is on the challenging and perhaps hitherto neglected aspects of ML that distinguish it as a genuine research subject in and of itself, as opposed to one whose characteristics are essentially predetermined by structures that precede it, e.g., the political economy that has given rise to the recent ‘AI Hype’ of the past ten years. We are interested in the theoretical challenges that ML confronts us with and how we can approach their resolutions – which includes identifying affordances implied in new perspectives. This volume presents a variety of such approaches and continues discussions that started at our conference in December 2021.

Some of the contributions explore issues related to aspects of “computational reason” (Cavia 2022) as manifested in ML and how they play out in practice. Beyond philosophical efforts to distinguish this notion from conventional conceptions of reasoning, this includes assessing the consequences of ML’s practical deployment specifically with regard to unintended consequences in its use. To the extent that it is employed to generate and distribute information, and hence meaning, effects that differ from the intentions of the technology’s makers and users abound (Broussard 2018). ML, in some respects, introduces new cultural dynamics based on the “epistemic shock” it brings about, based on its “claim to meaning itself” (Roberge/Castelle 2021: 2, 7).

Machine learning as a theoretical problem

Our interest is in these new ways in which ML as a computational technology feeds back to human ways of acting, organizing, classifying, and reasoning. In his keynote at our conference, Matteo Pasquinelli suggested conceiving of ML as a knowledge model in terms of a political epistemology. According to his conceptualization, there is an “epistemic scaffolding” at work that simultaneously connects ideological form, logical form, technical form, and social form. Pasquinelli’s talk highlighted ML’s multidimensionality as a real-world phenomenon. Having emerged from the dynamic interplay of mythology, collective imagery, statistical and mechanical thinking, computation, and the automation of labor, as well as surveillance and control of social behavior, ML presents us with a shift in how knowledge is produced, navigated, distributed, and modeled.

This shift has practical consequences. As Louise Amoore points out concerning “apparently coherent technologies of observation” (2020: 40), ML is capable of causing subtle and sometimes invisible changes down to the level of human perception and sensorium. She argues that “experimental algorithmic techniques” act “in and through data to modify the threshold of perceptibility itself” (p. 41). While ML-based “computer vision” cannot actually ‘see’, it does play a significant part in *what* becomes visible, and hence intelligible, to human observers. How humans perceive the world increasingly depends on what algorithmic systems make available for observation. This is not to argue per se that there is obfuscation at work. Rather, it reveals that technologically mediated vision results from a selective process that can lead to contingent, potentially confusing outcomes.

Our interest in ML is stimulated by how difficult it seems to grasp in theoretical terms and how its real-world impact, in turn, challenges some quite elementary concepts of the humanities, such as notions of cognition, subjectivity, and communication. As Luciana Parisi observes, AI is an “alien subject” (2019) that questions what we consider to be human subjects. To the extent that it plays a major part in the automation of cognitive tasks, ML problematizes our understanding of ourselves. Deprived of its privileged access to decision-making, the human subject experiences the “expansion of an alien space of reasoning” which corresponds to a “crisis of conscious cognition” (pp. 30, 28). And “machine thinking” gets multiplied, so to speak, when connected machines become active in their own “space of communication opaque to human vision”

(p. 31). Therefore, it is not only thinking but also communication that becomes questionable in its human-centered theoretical form.

Elena Esposito introduces the concept of “artificial communication” (2022) to theorize the strange encounters between humans and ML-based agents. In her keynote at our conference, she spelled out the implications of this concept. It would be wrong to assume, Esposito argued, that ML reproduces the characteristics of human behavior, reasoning, or communication. Instead, it is most successful at tasks where it is not deployed to imitate human consciousness or reasoning (p. 2). ML can take on the form of a social agent by processing “virtual contingency”. When users interact with ML-based agents, such as language models, they face a contingency that is not their own nor is it really the contingency of the machine because the model’s output is entirely based on (often human-generated) training data whose meaning and actual origin are in themselves contingent (pp. 9–10). Hence, as Esposito explained in her talk, the algorithmic machine presents the users with processed – or reflected, so to speak – perspectives of other users and their contingent perspectives. She concluded that virtual contingency constitutes a mode of communication without distinct alterity. This indicates a shift in the very conditions of sociality. ML potentially allows humans to engage with machines in ways that were previously considered a privilege of human-to-human relationships. It is language-based communication, but “artificial” – to use Esposito’s term here to indicate its lack of a familiar element, namely an address to which one can attribute received information.

The trouble here is twofold, as the possibility of the interpretability of ML depends on its addressability as a distinct entity which, in turn, is a precondition for the attribution of accountability. Addressability, a core feature of computational technologies in general (Dhaliwal 2022), is not a given in the face of troubles with opacity in ML (Burrell 2016). There have been notable research efforts aimed at closing this “responsibility gap” (Matthias 2004) and at solving issues related to opacity by the development of methods directed toward “explainable AI” to satisfy the explanatory requirements of different stakeholders involved in ML applications (Zednik 2021). Yet to this day, in many, if not most applications of ML, it is not always clear what caused a model to produce a certain output. This seemingly constitutive gap leaves room for projections. Some have ascribed a sublime status to ML (Ames 2018) – think of the various versions of the discourse on “superintelligence” (Bostrom 2014) – while others have ridiculed it for its actual triviality in the face of the apparent rise of singu-

larity; anecdotes of “artificial unintelligence” (Broussard 2018), if not outright “artificial stupidity” (Steyerl 2017; see also Mackinnon 2017) abound.

With regard to the approaches assembled in this volume, we see these diverse and ambivalent ideas of and views on ML as expressions of the subject’s multifaceted social nature. In this sense, perhaps it is best to think of it as a chimera, a “trickster of the natural order” whose nature is beyond “imposed dichotomies”, as Ilan Manouach and Anna Engelhardt write in the introduction to their *Inventory of Synthetic Cognition* (2022: 9). ML can simultaneously be a way of processing information, of generating meaning, of negotiating social norms, and of exercising power. It is not a limited-range, domain-specific technology. Rather, it extends into myriad fields and takes on many forms so that ultimately, given the variety of its potential functions in the world, we feel a need to reflect on questions concerning its conceptualization.

By taking a step back, we intend to examine how ML can be plausibly theorized in the face of the challenges relating to its real-world embeddedness. In summary, these characteristics include its ambivalent nature, owing to the dynamics of its entangled situated relationships with other social entities (Groß/Wagenknecht 2023), the alien nature of its (partly pathological) contributions to the world, which are sometimes difficult to account for, and its invisibility as well as its questionable addressability, which leave space for speculation and projection. ML-based technology has become such an integral part of social life that it seems to be becoming “invisible” in its largely unquestioned ubiquity in lived experiences. Such “transparency”, to reiterate a notion from Susan Leigh Star and Karen Ruhleder’s formative research on “large information spaces” (Star/Ruhleder 1996), is characteristic of any social infrastructure. Acknowledging the infrastructural status of ML-based technologies speaks to their social acceptance which, in its normalizing and habitualizing effects, raises the question of how to maintain critical observation and analysis.

Dimensions of machine learning: models – practices – topologies

To approach ML’s real-world embeddedness, we suggest a tripartite framework to highlight what we consider to be some of the most crucial dimensions of its theorization. We propose this schema as a heuristic to help relate different theoretical perspectives on ML to one another.

Viewing ML as *models* – of real phenomena as well as of knowledge (production) itself – highlights the epistemological stakes implied in its use. ML relies on models of worldly phenomena and is as much a product of this world as a

contributor to it: “the model is the message” (Bratton/ Agüera y Arcas 2022). To inquire into the specific conditions and characteristics of modeling is to inquire into ML’s reason and epistemology.

The impact of its model-based contributions, on the other hand, depends upon how they are put to use *in practice*. The practical dimensions of technology in action, while related to the model aspects, deserve attention in their own right. It is through practices of machine learning that we can gain insights into its characteristics as a real-world phenomenon, including its embeddedness in existing cultural, economic, and political realities, its social and material interdependencies, and the dynamics of its production as well as its application across different domains.

Finally, in an effort to approach the non-human, alien characteristics inherent in ML as distinct from other technologies, we regard the *topologies* of its computational operations as an important aspect. In particular, we suggest that computational operations in many-dimensional vector spaces come with their own dynamics that reveal insights into what characterizes the ‘thinking’, ‘cognition’, or ‘reasoning’ of ‘intelligent’ machines. These operations represent the interrelations of the individual data points in a neural network. In fact, vector space can be considered the very medium of artificial neural networks in computational terms. Its properties set the conditions for, as well as the horizon of, their possibilities and limitations.

Practical relationality, structural interdependence, cooperative generativity

A recurring analytical theme across this volume is the relationality of social entities in ML practices. The mutual influence of humans, machines, infrastructures, institutions, and other material entities on how systems, actions, and communications gain shape is embedded within specific dynamics that result in more or less stable relationships. Observing how roboticists deal with their material prototypes, **Hannah Link** (pp. 143-168) notes that advances in this field of research are accompanied by specific modes of posthuman interaction between knowledge objects and knowledge subjects. These novel modes of interaction in turn also affect the researchers’ perceptions of their own humanity. **Yaoli Du and Nadine Schumann** (pp. 193-207) address the question of an adequate theoretical approach for the development of communication between humans and machines, focusing on the interactive formation of shared pragmatic action patterns that allow heterogeneous social entities to integrate and

mutually attune. In a similar vein, **Jonathan Harth and Maximilian Locher** (pp. 169-191) examine the relational patterns in human-machine interaction and their implications for social theory. Inspired by systems theory and theories of social networks as well as ethnomethodology, they spell out how ML relates to the dynamics in the continuous ontogenesis of social identities. A detailed case study of one such human-machine relationship is presented by **Jan Georg Schneider and Katharina Zweig** (pp. 93-111) on the *e-rater*, a ML-based grading system used to predict test scores in the evaluation of written essays. Relying on the analysis of speech act felicity conditions, the authors analyze the potentials, limits, and contingencies in the automated grading of written text. All these chapters negotiate the practical relationalities of ML within different theoretical frameworks, yet complement one another in their emphasis on posthumanist perspectives on interaction in their search for possibilities to expand the anthropocentric focus that still characterizes many strands of social theory.

Another guiding theme of the various chapters results from critical reflection on the economic and political contexts in which ML-based technologies are developed and deployed. In such a view, ML does not appear to be an autonomous entity but rather an element of interdependent social structures. Examining political patterns of domination inscribed into ML-based systems of text generation, such as large language models (LLMs), **Christian Heck** (pp. 235-286) presents the practice of *adversarial hacking* as a subversive tool to tackle the conservative cultural tendencies and hegemonic ideological leanings of such systems. **Jan Fuhrmann** (pp. 115-141) approaches the problem of interdependence through systems theory and identifies a constitutive translation gap between algorithmic systems, which are blind to the social grammars of discrimination precisely because of their autopoiesis, which implies operational closure, and the structure of the datasets their models are based on produced by their environment. Focusing on the development of LLMs at Google/Alphabet, **Jonathan Roberge and Tom Lebrun** (pp. 39-65) highlight the hermeneutical implications of the rapid developments in the field of automated text generation and the power structures within which this development is taking place. The authors conclude that the economic framework of prestigious research projects at companies like Google provides little incentive to sufficiently care for the qualitative curation of datasets but rather encourages scaling to increase sales, ultimately resulting in a largely uncontrolled amplification of unintended adverse – and often discriminatory – effects. **Catriona Gray** (pp. 67-92) explores the intersection of ethics,

epistemology, and politics by analyzing examples of ML systems used in governmental decision-making practices. Her investigation shows which novel forms of rights or status violations – obviation, diminishment, impugment – become pressing issues in the increasing entanglement of ML in sovereign decision-making and practices of administrative judgment.

A third and final noteworthy point of emphasis negotiated across different chapters of this volume is the cooperative generativity in creative practices featuring ML technologies. Touching on discussions regarding the authorship of technological systems of automation, the question of generativity revolves around the specific dynamics of practical cooperation in creative processes involving ML. **Miriam Akkermann** (pp. 315–329) approaches the topic by examining affordances, both real and imagined, that stem from the integration of different ML methods (and earlier forms of AI preceding them) in musical composition. In their contribution **Jakob Claus and Yannick Schütte** (pp. 211–233) explore the dynamics of literary co-production between human author K. Allado-McDowell and OpenAI’s large language model GPT-3. This co-production culminated in the publication of the book *Pharmako-AI* (Allado-McDowell 2021), which is the subject of analysis here. **Michael Klippbahn-Karge** (pp. 287–314) addresses the ways in which ML, and AI generally, can become the subject of artworks. He suggests an analytical framework to distinguish the different ways in which contemporary artists refer to ML in their works – as a technology actively used to create an artwork, as a subject of critique, or as an occasion for the artist to develop their work. Likewise, regarding the conceptual questions surrounding “AI art”, **Fabian Offert** (pp. 273–286) suggests that ML-based artworks can be conceived of as sculptures owing to the subtractive nature of the image generation methods involved in their making.

The last two contributions of the volume argue against misguided reductionist and trivializing views of ML, and instead, examine the potential of topologically informed understandings of the generative aspects of ML. Exploring the notion of “computational reason” (Cavia 2022) in terms of its topological implications, **AA Cavia’s and Patricia Reed** (pp. 351–363) advance a concept of pointless topology based on constructive mathematics. Their argument considers the topological view that “all space comes with an attendant structure” (p. 353) and proceeds to spell out the affordances of a conceptual vocabulary that unifies the realms of the discrete and the continuous – the algebraic and the geometric – in order to expand computational reason’s “inferential toolkit” (p. 361). Similarly, in their approach to reason via topology, **Lukáš Likavčan and Carl Christian Olsson** (pp. 333–349) critically evaluate Immanuel Kant’s analogy

of orientation in thinking and geographical orientation to propose a topological account of thinking. They take the discussion of artificial neural networks' topological characteristics as an opportunity to reflect on the notion of human thinking in the mirror of the spatiality of so-called deep learning. As computational topologies are of crucial importance in making sense of the alienness and incomprehensibility of ML, an adequate understanding of their relevance and implications will, we hope, contribute to its theorization, and perhaps also to new approaches to the philosophy of computation.

Bibliography

- Allado-McDowell, K. 2020. *Pharmako-AI*. UK: Ignota Books.
- Ames, Morgan G. 2018. Deconstructing the algorithmic sublime. *Big Data & Society* 5(1):1–4. <https://doi.org/10.1177/2053951718779194>.
- Amoore, Louise. 2020. *Cloud Ethics: Algorithms and the attributes of ourselves and others*. Durham and London: Duke University Press.
- Apprich, Clemens, Wendy Hui Kyong Chun, Florian Cramer and Hito Steyerl. 2018. *Pattern Discrimination*. Lüneburg and Minneapolis, MN: meson press/Minnesota University Press.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>.
- Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Bratton, Benjamin and Blaise Agüera y Arcas. 2022. The Model Is The Message. *Noema Mag*. <https://www.noemamag.com/the-model-is-the-message/>. Last access: 12 December 2022.
- Broussard, Meredith. 2018. *Artificial Unintelligence: How Computers Misunderstand the World*. Boston, Mass.: MIT Press.
- Bucher, Taina. 2017. The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society* 20(1): 30–44. <https://doi.org/10.1080/1369118X.2016.1154086>.
- Burrell, Jenna. 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society* 3(1):1–12. <https://doi.org/10.1177/2053951715622512>.
- Cavia, AA. 2022. *Logiciel: Six Seminars on Computational Reason*. Berlin: &&&.

- Chun, Wendy Hui Kyong. 2021. *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*. Boston, Mass.: MIT Press.
- Dhaliwal, Ranjodh Singh. 2022. On Addressability, or What Even Is Computation? *Critical Inquiry* 49(1):1–27. <https://doi.org/10.1086/721167>.
- Dhar, Payal. 2020. The carbon impact of artificial intelligence. *Nature Machine Intelligence* 2:423–425. <https://doi.org/10.1038/s42256-020-0219-9>.
- Dourish, Paul. 2016. Algorithms and their others: Algorithmic culture in context. *Big Data & Society* 3(2): 1–11. <https://doi.org/10.1177/2053951716665128>.
- Esposito, Elena. 2022. *Artificial Communication: How Algorithms Produce Social Intelligence*. Cambridge, Mass.: MIT Press.
- Galloway, Alexander. 2021. Questions. Answers. <http://cultureandcommunication.org/galloway/questions-answers>. Last access: 12 December 2022.
- Groß, Richard and Susann Wagenknecht. 2023. Situating machine learning – On the calibration of problems in practice. *Distinktion. Journal of Social Theory*. <https://doi.org/10.1080/1600910X.2023.2177319>.
- Mackenzie, Adrian. 2017. *Machine Learners: Archaeology of a Data Practice*. Cambridge, Mass.: MIT Press.
- Mackinnon, Lee. 2017. Artificial Stupidity and the End of Men. *Third Text* 31(5–6):603–617. <https://doi.org/10.1080/09528822.2018.1437939>.
- Manouach, Ilan and Anna Engelhardt. 2022. Preface. In *Chimeras: Inventory of Synthetic Cognition*, Eds. Ilan Manouach and Anna Engelhardt, 9–13. Athens: Onassis Foundation.
- Matthias, Andreas. 2004. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology* 6:175–183. <https://doi.org/10.1007/s10676-004-3422-1>.
- Mullaney, Thomas S., Benjamin Peters, Mar Hicks and Kavita Philip (eds.). 2021. *Your Computer is On Fire*. Boston: MIT Press.
- Noble, Safiya Umoja. 2018. *Algorithms of oppression: How search engines reinforce racism*. New York: NYU Press.
- Parisi, Luciana. 2019. The Alien Subject of AI. *Subjectivity* 12(1):27–48.
- Pasquinelli, Matteo and Vladan Joler. 2020. The Nooscope Manifested: AI as Instrument of Knowledge Extractivism. *AI & Society* 36:1263–1280. <https://doi.org/10.1007/s00146-020-01097-6>.
- Roberge, Jonathan and Michael Castelle. 2021. Toward an End-to-End Sociology of 21st-Century Machine Learning. In *The Cultural Life of Machine Learning: An Incursion into Critical AI Studies*, Eds. Jonathan Roberge and Michael Castelle, 1–20. Cham: Palgrave Macmillan.

- Sherman, Stephanie. 2022. The Polyopticon: A diagram for urban artificial intelligences. *AI & Society*. <https://doi.org/10.1007/s00146-022-01501-3>.
- Star, Susan Leigh and Karen Ruhleder. 1996. Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research* 7(1):111–134. <https://doi.org/10.1287/isre.7.1.111>.
- Star, Susan Leigh and Karen Ruhleder. 2017 (engl. 1996). Schritte zu einer Ökologie von Infrastruktur. Design und Zugang für großangelegte Informationsräume. In Susan Leigh Star, *Grenzobjekte und Medienforschung*, Eds. Sebastian Gießmann and Nadine Taha, 359–401. Bielefeld: transcript.
- Steyerl, Hito. 2017. The Nation-State System: „Gott ist doof.“ On Artificial Stupidity. *Now is the Time of Monsters. What Comes After Nations* (Haus der Kulturen der Welt, Berlin, 23–25 March 2017). <https://soundcloud.com/hkw/now-is-the-time-of-monsters-2>. Last access: 12 December 2022.
- Strubell, Emma, Ananya Ganesh and Andrew McCallum. 2019. Energy and Policy Considerations for Deep Learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–3650, Florence: Association for Computational Linguistics. <http://dx.doi.org/10.18653/v1/P19-1355>.
- Whittaker, Meredith. 2021. The Steep Cost of Capture. *interactions* 28(6):50–55. <https://doi.org/10.1145/3488666>.
- Zednik, Carlos. 2021. Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *Philosophy & Technology* 34:265–288. <https://doi.org/10.1007/s13347-019-00382-7>.

