INHALT

Einleitung 9

1 Erde 31

2 Arbeit 61

3 Daten. 99

4 Klassifizierung 135

5 Affekt 165

6 Staat 197

Schluss: Macht 229

Coda: Weltraum. 249

ANHANG

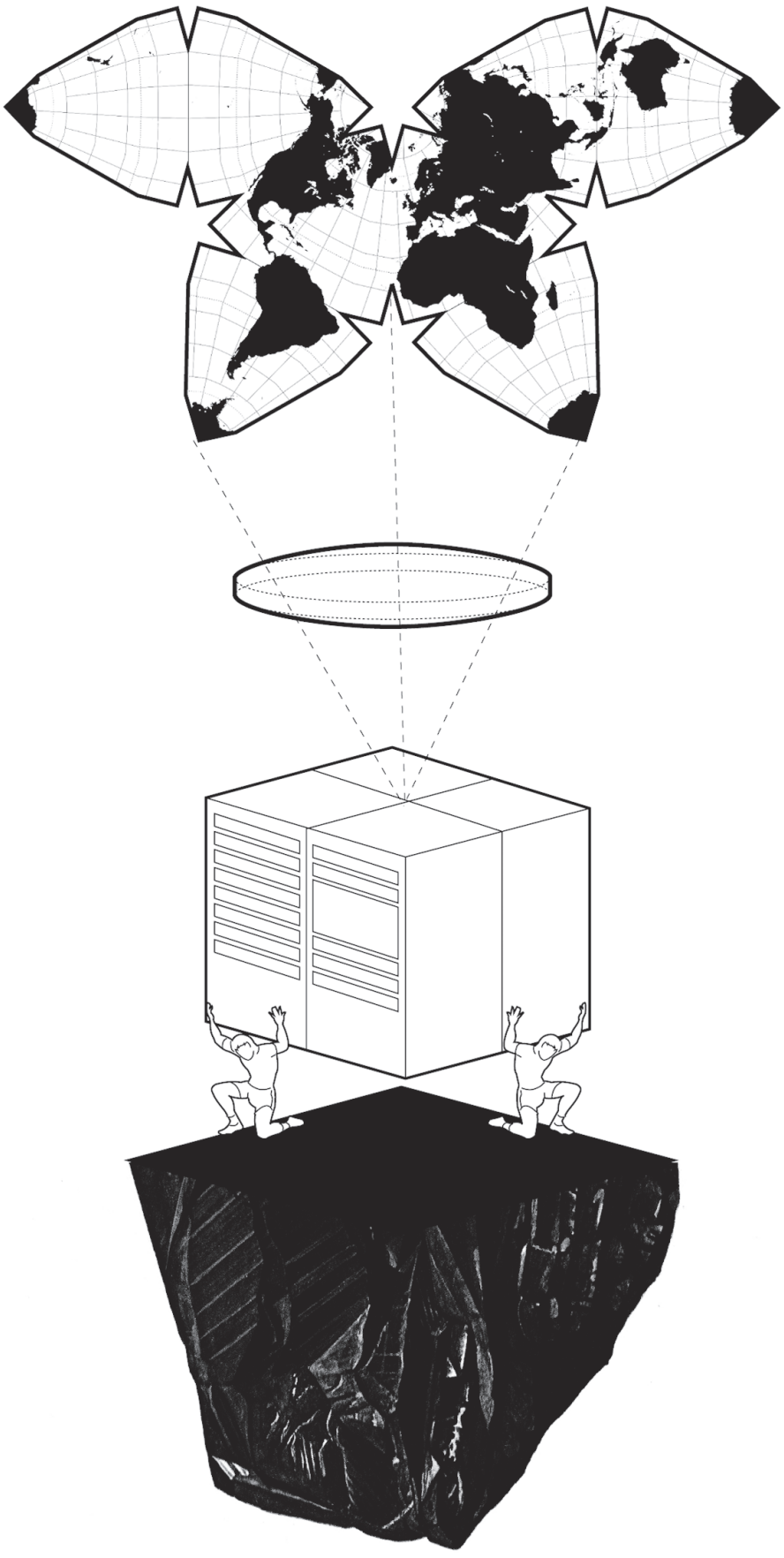
Danksagung 259

Bibliografie 265

Anmerkungen 307

Bildnachweis 332

Register 333



EINLEITUNG

Das klügste Pferd der Welt

Europa war am Ende des 19. Jahrhunderts gefesselt von einem Pferd namens Hans. Der «Kluge Hans» war ein wahres Wunder: Er konnte mathematische Aufgaben lösen, die Zeit angeben, Tage auf dem Kalender markieren, musikalische Töne unterscheiden sowie Wörter und Sätze buchstabieren. Die Leute rissen sich darum, den deutschen Hengst dabei zu beobachten, wie er durch das Klopfen mit seinen Hufen Antworten auf schwierige Probleme diktierte und dabei zuverlässig richtiglag. «Was ist zwei plus drei?» Hans stampfte eifrig fünfmal auf dem Boden auf. «Welcher Wochentag ist heute?» Das Pferd stampfte mit den Hufen auf, um jeden einzelnen Buchstaben auf einer zu diesem Zweck präparierten Tafel anzuzeigen, und buchstabierte auf diese Weise die korrekte Antwort. Und Hans konnte sogar noch komplexere Aufgaben lösen, etwa diese: «Ich denke an eine bestimmte Zahl. Von dieser ziehe ich neun ab und lande im Ergebnis bei der Drei. An welche Zahl denke ich?» Bis zum Jahr 1904 hatte der Kluge Hans internationale Berühmtheit erlangt; die *New York Times* feierte ihn mit den Worten «Berlins wundersames Pferd – es kann fast alles außer sprechen».[1]

Hans' Trainer, ein pensionierter Mathematiklehrer namens Wilhelm von Osten, war lange schon von tierischer Intelligenz fasziniert. Er hatte sich bereits erfolglos an der Unterrichtung von Katzen- und Bärenjungen in großer Zahl versucht, doch erst die Arbeit mit seinem eigenen Pferd war erfolgreich. Am Anfang brachte er Hans das Zählen bei, indem er dessen Bein hochhielt, ihm eine Zahl präsentierte und dann die Hufe entsprechend häufig auf dem Boden aufsetzte. Bald schon gab Hans die Ergebnisse einfacher Additionen durch seine Huf-

tritte zutreffend wieder. Anschließend brachte van Osten eine Tafel mit dem Alphabet ins Spiel, sodass Hans nun Buchstaben benennen konnte, indem er so und so viele Male aufstampfte. Nach zwei Jahren Ausbildung war von Osten überrascht von dem tiefen Verständnis komplexer gedanklicher Gebilde, das das Tier an den Tag legte. Daher ging er nun mit Hans auf Tournee, um zu beweisen, dass auch Tiere denken konnten. Damit wurde das Pferd zu einer viralen Sensation der Belle Époque.

Viele Leute waren allerdings skeptisch. Das preußische Kultusministerium richtete eine Untersuchungskommission ein, die von Ostens wissenschaftliche Behauptungen überprüfen sollte. Die «Hans-Kommission» wurde vom Psychologen und Philosophen Carl Stumpf sowie von seinem Assistenten Oskar Pfungst geleitet. Sie bestand unter anderem aus einem Zirkusdirektor, einem pensionierten Lehrer, einem Zoologen, einem Tierarzt und einem Offizier der Kavallerie. Doch auch nach ausführlichen Befragungen des Tiers sowohl in An- als auch in Abwesenheit seines Trainers konnte das Pferd die Liste seiner richtigen Antworten ununterbrochen fortführen und die Kommission keine Belege für etwaige Betrügereien finden. Wie Pfungst später schrieb, stellte Hans seine Künste vor «Tausende(n) von Zuschauern während vieler Monate» unter Beweis – «Pferdekenner, Trick-Kenner ersten Ranges, unter denen nicht ein einziger irgendwelche regelmäßige Zeichen bemerkt», die zwischen Befrager und Pferd ausgetauscht worden wären.[2]

Die Kommission befand, dass die Methoden, nach denen Hans unterwiesen worden war, eher dem «Unterricht von Kindern in der Grundschule» ähnelten als der Abrichtung eines Tiers und daher «der wissenschaftlichen Erforschung lohnten».[3] Doch Stumpf und Pfungst hegten immer noch Zweifel. Einer ihrer Befunde gab ihnen besonders zu denken: Wenn der Befrager selbst die Lösung nicht kannte oder weit entfernt von ihm stand, antwortete Hans nur selten richtig. Dies führte Pfungst und Stumpf zu der Überlegung, ob möglicherweise irgendein unbewusstes Signal die Antworten an Hans übermittelt haben könnte.

Und wie Pfungst in seinem Buch von 1911 beschreibt, hatte sie ihr Gefühl nicht getäuscht: Die Pose des Befragers, sein Atem und seine Mimik veränderten sich auf kaum merkliche Weise, wenn sich das Tier der richtigen Antwort näherte, und brachten es dazu, an der entsprechenden Stelle innezuhalten.[4] Später erprobte Pfungst diese Hypothese

Wilhelm von Osten und der Kluge Hans

auch an menschlichen Testkandidaten und fand sein Ergebnis bestätigt. Am meisten faszinierte ihn an dieser Entdeckung, dass den Fragestellern im Allgemeinen gar nicht bewusst war, dass sie dem Pferd Hinweise gaben. Die Lösung des Rätsels vom Klugen Hans war also, so Pfungst, die unbewusste Anleitung durch den Befrager des Pferds:[5] Das Tier wurde daraufhin konditioniert, die Resultate zu erzielen, die sein Besitzer sehen wollte; das Publikum hingegen fand, dass dies nicht das war, was es sich unter einer außergewöhnlichen Intelligenz vorgestellt hatte.

Die Geschichte vom Klugen Hans ist in vielerlei Hinsicht verlockend: in Bezug auf das Verhältnis von Wünschen, Einbildung und Handeln; mit Blick auf die Ökonomie des Spektakels und auf die Frage, wie wir das Nichtmenschliche anthropomorphisieren; in Bezug darauf, wie ein kognitiver Bias entsteht; und hinsichtlich einer Politik der Intelligenz. Hans regte die Prägung eines Begriffs an, der in der Psychologie für eine bestimmte Art von begrifflicher Falle steht: der Kluge-Hans-Effekt, auch *Observer-expectancy*-Effekt genannt. Mit ihm wird der Einfluss der vom Versuchsleiter seinen Versuchspersonen unabsichtlich vermittelten Hinweise beschrieben. Das Verhältnis zwi-

schen Hans und von Osten macht die komplexen Mechanismen deutlich, durch die kognitive Verzerrungen ihren Weg in Systeme hineinfinden, und weist darauf hin, wie Menschen sich in die Phänomene verstricken, die sie untersuchen. Heute wird diese Geschichte im Kontext des maschinellen Lernens als Mahnung erzählt, dass man nicht immer wissen könne, was ein Modell eigentlich genau aus den Daten lerne, mit denen man es füttert.[6] Selbst ein System, das im Training offenbar spektakulär abschneidet, kann furchtbar schlechte Voraussagen treffen, wenn ihm neuartige Daten aus der Welt vorgelegt werden.

Dies führt uns zu einer zentralen Frage dieses Buchs: Wie wird Intelligenz «gemacht», und welche Fallen können sich dadurch auftun? Die Geschichte vom Klugen Hans handelt auf den ersten Blick davon, wie ein Mensch Intelligenz erschaffen hat, und zwar dadurch, dass er ein Pferd dazu abrichtete, auf Signale zu reagieren und dabei ein menschenähnliches Erkenntnisvermögen zu imitieren. Auf einer anderen Ebene jedoch erkennen wir, dass dieses Verfahren der Intelligenzherstellung noch viel breiter angelegt war. Denn dieses Unterfangen war darauf angewiesen, durch eine Vielzahl von Institutionen validiert und bestätigt zu werden, darunter Universitäten, Schulen, die Wissenschaft, die öffentliche Meinung und das Militär. Und außerdem gab es ja auch noch einen Markt für von Osten und sein besonderes Pferd – emotionale und finanzielle Investitionen, die die Besichtigungstouren, Zeitungsberichte und Vorträge antrieben. Bürokratische Autoritäten wurden versammelt, um die Fähigkeiten des Tiers zu messen und zu testen. In die Konstruktion von Hans' Intelligenz spielte also ein ganzes Konglomerat von finanziellen, kulturellen und wissenschaftlichen Interessen hinein und entschied mit darüber, ob diese tatsächlich außergewöhnlich war.

Wir sehen hier zwei unterschiedliche Mythologien am Werk. Der erste Mythos besagt, dass nichtmenschliche Systeme (seien es Computer oder Pferde) dem menschlichen Geist irgendwie analog wären. Diese Auffassung unterstellt, dass – genügend Training oder Ressourcen vorausgesetzt – eine menschenähnliche Intelligenz aus dem Nichts geschaffen werden könnte, ohne dabei die grundlegende Tatsache zu berücksichtigen, dass Menschen körperliche und soziale Wesen sind, die in größere ökologische Zusammenhänge eingebettet sind. Der zweite Mythos besagt, dass Intelligenz etwas ist, das freischwebend existiert,

so als wäre sie ein natürlicher, von gesellschaftlichen, kulturellen, historischen und politischen Kräfteverhältnissen losgelöster Gegenstand. Tatsächlich hat diese Auffassung von Intelligenz im Laufe der Jahrhunderte einen enormen Schaden angerichtet und diente auch dazu, Herrschaftsverhältnisse von der Sklaverei bis zur Eugenik zu rechtfertigen.[7]

Besonders wirkmächtig sind diese Mythologien auf dem Feld der künstlichen Intelligenz, wo seit Mitte des 20. Jahrhunderts die unumstößliche Überzeugung vorherrscht, dass die menschliche Intelligenz formalisiert und durch Maschinen reproduziert werden könne. So wie Hans' Intelligenz, die vermeintlich so sorgfältig herangezogen worden war wie ein Grundschulkind, mit der Intelligenz eines Menschen vergleichbar galt, so wurden auch KI-Systeme wiederholt als zwar simple, aber dennoch menschenähnliche Formen von Intelligenz beschrieben. 1950 prognostizierte Alan Turing, «daß am Ende des Jahrhunderts der Gebrauch von Wörtern und die allgemeinen Ansichten der Gebildeten sich so sehr geändert haben werden, daß man ohne Widerspruch von denkenden Maschinen wird reden können».[8] Der Mathematiker John von Neumann erklärte 1958, dass das menschliche Nervensystem «*prima facie* digitaler Natur» sei,[9] und der MIT-Professor Marvin Minsky antwortete einmal auf die Frage, ob Maschinen denken könnten: «Natürlich können Maschinen denken; wir können denken, und wir sind ‹Fleischmaschinen›.»[10] Nicht jeder teilte jedoch diese Auffassung. Joseph Weizenbaum, ein früher Erfinder im KI-Bereich und Schöpfer des ersten Chatbot-Programms namens ELIZA, war der Meinung, dass dem Bild vom Menschen als bloßem Informationsverarbeitungssystem ein viel zu vereinfachter Intelligenzbegriff zugrunde liege, der wiederum der «perversen, grandiosen Phantasie» das Wort geredet habe, dass KI-Forscher eine Maschine erschaffen könnten, die lerne «wie ein Kind».[11]

Dies war einer der entscheidenden Streitpunkte in der Geschichte der künstlichen Intelligenz. Im Jahr 1961 veranstaltete das MIT eine wegweisende Vortragsreihe mit dem Titel «Management and the Computer of the Future». Ein herausragendes Aufgebot an Computerwissenschaftlerinnen und -wissenschaftlern nahm daran teil, darunter Grace Hopper, J. C. R. Licklider, Marvin Minsky, Allen Newell, Herbert Simon und Norbert Wiener, und diskutierte über die rapiden Fortschritte, die bei der digitalen Datenverarbeitung erzielt worden waren.

Zum Abschluss brachte John McCarthy die kühne These vor, dass die Unterschiede zwischen menschlichen und maschinellen Aufgaben illusorisch sei. Letztere seien einfach komplizierte menschliche Aufgaben, für die es bloß mehr Zeit brauche, um sie zu formalisieren und von Maschinen lösen zu lassen.[12]

Der Philosophieprofessor Hubert Dreyfus widersprach dem allerdings und äußerte die Sorge, dass die versammelten Ingenieure und Ingenieurinnen «die Möglichkeit nicht einmal in Erwägung ziehen, dass das Gehirn Informationen auf völlig andere Weise verarbeiten könnte als ein Computer».[13] Und in seinem späteren Werk *Was Computer nicht können* macht er deutlich, dass menschliche Intelligenz und Sachverstand stark von vielen un- und vorbewussten Prozessen abhängig sind, während Computer für ihr Funktionieren darauf angewiesen sind, dass alle Prozesse und Daten explizit ausformuliert und formalisiert sind.[14] Weniger formale Aspekte der Intelligenz müssten daher entweder abstrakter gefasst, eliminiert oder auf die Erfordernisse des Computers zugeschnitten werden, was ihn unfähig dazu mache, Informationen über Situationen so zu verarbeiten, wie es Menschen tun.

Auf dem Feld der KI hat sich seit den 1960er Jahren vieles verändert. Unter anderem kam es zu einem Übergang von symbolischen Systemen hin zu jener Welle der Begeisterung für Techniken des maschinellen Lernens, wie wir sie in jüngerer Zeit beobachten können. Die frühen Auseinandersetzungen darüber, was KI eigentlich leisten kann, sind vergessen, und die Skepsis ist verflogen. Seit Mitte der 2000er Jahre hat sich die künstliche Intelligenz in Windeseile sowohl zu einem akademischen Forschungsfeld als auch zu einer ganzen Industrie ausgewachsen. Heute setzen einige wenige mächtige Technologiekonzerne KI-Systeme im planetarischen Maßstab ein, und erneut werden diese als mit der menschlichen Intelligenz vergleichbar oder ihr sogar überlegen angepriesen.

Die Geschichte vom Klugen Hans erinnert uns allerdings daran, wie spezifisch das eigentlich ist, was wir unter Intelligenz verstehen oder als solche anzuerkennen bereit sind. Hans wurde gelehrt, die Lösung von Aufgaben in einem sehr eng umgrenzten Rahmen zu imitieren: addieren, subtrahieren und Wörter buchstabieren. Dies spiegelt eine eingeschränkte Perspektive wider auf das, was Pferde *und* Menschen können.

Hans vollbrachte zwar schon beachtliche Leistungen in den Bereichen der artübergreifenden Kommunikation, der öffentlichen Vorführung und des geduldigen Abwartens, aber dennoch wurden diese nicht als Intelligenz anerkannt. Wie es die Autorin und Ingenieurin Ellen Ullman ausdrückt, habe jener Glaube, dass der Geist wie ein Computer und umgekehrt sei, «Jahrzehnte des Nachdenkens in den Computer- und Kognitionswissenschaften infiziert» und stelle eine Art Erbsünde in diesem Bereich dar.[15] In der künstlichen Intelligenz steckt die Ideologie eines kartesischen Dualismus: Sie wird im engen Sinne als entkörperlichte Intelligenz verstanden, bar jeder Beziehung zur materiellen Welt.

Was ist KI?
Weder künstlich noch intelligent

Stellen wir einmal die trügerisch einfache Frage: Was ist künstliche Intelligenz? Wenn Sie eine Passantin auf der Straße danach fragen, wird sie möglicherweise Siri von Apple nennen oder Amazons Cloudspeicher, die Autos von Tesla oder den Suchalgorithmus von Google. Wenn Sie die gleiche Frage hingegen an Experten auf dem Gebiet des Deep Learning richten, dann geben diese Ihnen wahrscheinlich eine eher technische Antwort, die sich darum dreht, wie neuronale Netze sich aus Dutzenden von Schichten zusammensetzen, gelabelte Daten empfangen, Gewichtungen und Schwellenwerte zugewiesen bekommen und damit Daten auf eine Weise zu klassifizieren in der Lage sind, die noch nicht vollständig erklärt werden kann.[16] Im Zuge einer Analyse von Expertensystemen beschrieb Professor Donald Michie im Jahr 1978 künstliche Intelligenz als eine Verfeinerung des Wissens, bei der «eine Zuverlässigkeit und Kompetenz der Kodifizierung hergestellt werden kann, die über das höchste Niveau weit hinausgeht, welches ein menschlicher Experte ohne externe Hilfsmittel jemals erreicht hat, ja vielleicht auch nur erreichen konnte».[17] In einem der bekanntesten Lehrbücher zum Thema vertreten Stuart Russell und Peter Norvig die Ansicht, dass KI der Versuch sei, intelligente Wesen zu verstehen und zu konstruieren. «Intelligenz hat primär mit rationalem Handeln zu

tun», wie sie behaupten. «Idealerweise vollführt ein intelligenter Akteur die bestmögliche Handlung in einer gegebenen Situation.»[18]

Jeder Versuch, künstliche Intelligenz zu definieren, tut etwas, indem er einen Rahmen dafür festlegt, wie diese aufzufassen, gemessen, bewertet und reguliert werden sollte. Wird KI von Konsumgüterherstellern für ihre konzernweite Infrastruktur definiert, dann bestimmen Marketing und Werbung ihren Horizont. Werden KI-Systeme als zuverlässiger oder rationaler als jeder menschliche Akteur betrachtet, da sie vermeintlich in der Lage sind, die «bestmögliche Handlung» auszuführen, dann legt dies nahe, dass man ihnen auch weitreichendere Entscheidungen auf den Gebieten der Gesundheitsversorgung, des Bildungswesens oder der Strafjustiz überlassen könnte. Und stehen einzig spezielle Algorithmen im Fokus des Interesses, dann suggeriert dieser Umstand, dass nur der ununterbrochene technische Fortschritt zähle und weder die Rechenkosten solcher Ansätze noch ihre weitreichenden Auswirkungen auf einen ohnehin schon unter Stress stehenden Planeten mitberücksichtigt werden.

Im Gegensatz dazu möchte ich in diesem Buch die These darlegen, dass KI weder *künstlich* noch *intelligent* ist. Künstliche Intelligenz ist vielmehr verkörpert und materiell – hergestellt auf der Basis von natürlichen Rohstoffen, Kraftstoffen, menschlicher Arbeitskraft, Infrastrukturen, Logistiken, Geschichten und Klassifikationen. KI-Systeme sind ohne ein vorheriges ausgiebiges und rechenintensives Training mit umfangreichen Datensätzen oder vorgegebenen Regeln und Belohnungen weder autonom noch rational oder auch nur in der Lage, irgendetwas wahrzunehmen. Tatsächlich ist die künstliche Intelligenz, so wie wir sie kennen, in ihrer Existenz sogar vollkommen auf zahlreiche übergeordnete politische und soziale Strukturen angewiesen. Und aufgrund des Kapitals, das für die Entwicklung von KI in großem Maßstab erforderlich ist, und der Arten von Wahrnehmung, die durch sie optimiert wird, sind KI-Systeme letztlich so konzipiert, dass sie den bestehenden herrschenden Interessen dienen. In diesem Sinne ist künstliche Intelligenz ein Register der Macht.

In diesem Buch werden wir der Frage nachgehen, wie künstliche Intelligenz im weitesten Sinne *gemacht* wird, und die ökonomischen, politischen, kulturellen und geschichtlichen Kräfte aufdecken, die sie

prägen. Haben wir die KI erst einmal mit solchen größeren Strukturen und sozialen Systemen wieder verbunden, dann können wir auch die Vorstellung mit guten Gründen zurückweisen, dass künstliche Intelligenz nur eine Sache der Technologie sei. Ganz grundlegend betrachtet sind es technische und soziale Praktiken, Institutionen und Infrastrukturen, Politik und Kultur, die KI ausmachen. Rechnerische Vernunft und verkörperte Arbeit sind zutiefst miteinander verbunden, denn KI-Systeme reflektieren soziale Beziehungen und Weltverständnisse ebenso, wie sie sie hervorbringen.

Es sei erwähnt, dass der Ausdruck «künstliche Intelligenz» in der computerwissenschaftlichen Gemeinschaft Unbehagen hervorrufen kann. Über die Jahrzehnte hinweg war er immer mal wieder ein modisches Schlagwort und wird eher beim Marketing als in der Forschung benutzt. In der technischen Literatur gängiger ist der Ausdruck «maschinelles Lernen». Trotzdem wird die Rede von KI häufig im Rahmen von Forschungsförderungsanträgen verwendet, wenn Risikokapitalgeber ihre Scheckbücher ziehen sollen oder Forschende die Aufmerksamkeit der Presse erheischen wollen, um eine neue wissenschaftliche Entdeckung publik zu machen. Das heißt, sowohl der aktive Gebrauch als auch die Ablehnung des Begriffs sorgen dafür, dass sein Bedeutungsgehalt stets im Fluss ist. Für meine Zwecke verwende ich den Ausdruck KI, um damit jenen gewaltigen industriellen Apparat zu bezeichnen, in den auch Politik, Arbeit, Kultur und Kapital involviert sind. Wenn ich hingegen vom maschinellen Lernen spreche, dann meine ich damit eine Reihe von technischen Verfahren (die aber eigentlich auch sozialen und infrastrukturellen Charakter besitzen, auch wenn sie kaum jemals unter diesem Aspekt thematisiert werden).

Es gibt jedoch gewichtige Gründe dafür, *warum* sich das Feld so sehr auf die technischen Aspekte konzentriert hat – auf algorithmische Fortschritte, zunehmende Produktverbesserungen und größeren Bedienungskomfort. Die Machtstrukturen an der Schnittstelle von Technologie, Kapital und Governance profitieren sehr von einer solch verengten, abstrahierenden Analyse. Um zu verstehen, inwiefern KI grundlegend politisch ist, müssen wir daher über die Betrachtung neuronaler Netze und statistischer Mustererkennung hinausgehen und stattdessen fragen, *was* optimiert wird sowie *für wen* und *wer* da-

rüber entscheiden darf. Anschließend können wir dann die Auswirkungen dieser Entscheidungen nachzeichnen.

KI wie einen Atlas betrachten

Wie kann ein Atlas uns zu verstehen helfen, wie künstliche Intelligenz gemacht wird? Ein Atlas ist ein ungewöhnliches Buch: Er ist eine Ansammlung disparater Teile, mit Karten, die in ihrem Maßstab variieren – von einer Satelitenaufnahme der Erde bis hin zur detaillierten Darstellung einer Inselgruppe. Wenn Sie ihn aufschlagen, dann sind Sie vermutlich auf der Suche nach spezifischen Informationen über einen bestimmten Ort – oder eben auch nicht. Vielleicht lassen Sie sich auch einfach treiben, folgen Ihrer Neugier und stoßen auf unerwartete Pfade und neue Perspektiven. Nach Ansicht der Wissenschaftshistorikerin Lorraine Daston zielen alle wissenschaftlichen Atlanten darauf ab, den Blick zu schulen und die Aufmerksamkeit der Betrachtenden auf bestimmte aufschlussreiche Details und signifikante Merkmale zu lenken.[19] Ein Atlas vermittelt Leser:innen eine bestimmte Sichtweise auf die Welt, versehen mit dem Signum der Wissenschaftlichkeit – in Gestalt von Maßstäben und Verhältnissen, Breiten- und Längengraden –, und einen Sinn für Form und Konsistenz.

Dennoch ist ein Atlas ebenso sehr ein kreativer Akt – eine subjektive, politische und ästhetische Intervention – wie eine Sammlung von wissenschaftlichen Daten. Der französische Philosoph Georges Didi-Huberman betrachtet ihn als etwas, das sowohl im ästhetischen Paradigma des Visuellen als auch im epistemischen Paradigma des Wissens beheimatet ist. Dadurch, dass er hier beide aufführt, unterminiert er die Ansicht, dass Wissenschaft und Kunst jemals vollständig voneinander getrennt sein könnten.[20] Ein Atlas bietet uns vielmehr die Möglichkeit, die Welt neu zu lesen, disparate Teile auf andere Weise miteinander zu verknüpfen und sie «neu zu bearbeiten und wieder zusammenzusetzen, ohne zu glauben, dass wir sie zusammenfassen oder erschöpfen».[21]

Mein vielleicht liebster Ausspruch darüber, inwiefern ein karto-