Einleitung

»Dimidium facti qui coepit habet: sapere aude.«
»Wer begonnen hat, hat schon zur Hälfte gehandelt: Wage zu denken!«
Horaz

Dieses Buch handelt von Ihnen, liebe Leserin, lieber Leser und von mir, dem Autor. Davon, wie wir als denkende menschliche Wesen in einer Zeit leben können, die dieses Denken immer stärker automatisiert und für eine angeblich überlegene Künstliche Intelligenz reklamiert. In der diese Künstliche Intelligenz bei ungebremstem und ungesteuertem Wachstum zu einer der größten Bedrohungen für uns werden kann. Und es handelt davon, was getan werden muss, damit wir und unsere Kinder und Enkel weiter in Freiheit und Selbstbestimmung leben können.

Geht es heute um KI, ist die Rede von einer Intelligenzexplosion der Maschinen, die zu einer künstlichen Superintelligenz führen kann. Ein Szenario, das nicht sicher, aber durchaus denkbar ist. Leider ist derzeit nicht ausgeschlossen, dass vorher die Implosion der menschlichen Vernunft erfolgt. Dafür gibt es jedenfalls schon mehr Anzeichen. Beide Szenarien hängen miteinander zusammen. Sie haben einen gemeinsamen Kern: die Krise der Vernunft, die sich zur Krise der Moderne und zur Krise des Menschen ausgewachsen hat.

Dabei geht es um die Frage, wer in Zukunft entscheidet, und wer die Kontrolle hat. Es geht um die Machtfrage. Eine immer mächtigere KI wirft das Kontrollproblem auf, und das auf zwei Ebenen: Sollte eine superintelligente KI geschaffen werden, was Fachleute zunehmend für möglich halten, steht die Menschheit vor dem Problem, ob sich diese KI überhaupt noch durch Menschen kontrollieren ließe, und wenn ja, wie. Sollte die Antwort darauf nein lauten, ist auch die Frage nach der Möglichkeit menschlicher Freiheit negativ entschieden, denn eine nicht kontrollierbare

KI gewinnt gegen den Menschen jedes Spiel. Das Risiko, dass eine solche Superintelligenz geschaffen wird, mag heute noch gering sein, da ihre Folgen aber derart massiv wären, muss es untersucht und die Entwicklungen heutiger KI entsprechend so gelenkt werden, dass die Risiken möglichst beherrschbar bleiben.

Die zweite Ebene des Kontrollproblems, das KI aufwirft, muss uns ebenso beschäftigen, denn es gefährdet schon heute die Freiheit: Es geht um die Macht, die Unternehmen und Staaten durch KI über Menschen ausüben. Eine Macht, die die Demokratie weltweit bedroht und die eine Gegenreaktion erfordert. Es geht – dramatisch formuliert – um die beiden Szenarien einer potenziellen (Welt-)Herrschaft *durch* KI oder einer potenziellen (Welt-)Herrschaft *der* KI. Beide Ebenen bilden das politische Kontrollproblem der mächtigsten Technologie, die Menschen je entwickelt haben. Es sind Fragen, die wir heute beantworten und Entscheidungen, die wir heute treffen müssen, um morgen überhaupt noch Entscheidungen treffen zu können.

Auch wenn dieses Buch kritisch mit KI ins Gericht geht, so ist es doch nicht radikal technikskeptisch oder gar technikablehnend. KI kann auch viel Gutes ermöglichen: bessere Medikamente, bessere Maschinen, eine bessere Versorgung der Weltbevölkerung vor allem mit Wasser, Lebensmitteln und Bildung. Vielleicht gelingt es sogar, mit ihrer Hilfe viele der großen sozialen und politischen Konflikte unserer Zeit oder auch die größte derzeitige Herausforderung, die Klimakrise, besser zu lösen. Das alles setzt aber voraus, dass die Menschen die Kontrolle über sie behalten.

Die Frage, welche Richtung diese Technologie nehmen wird, ist offen. Fest steht allein, dass heute schon darüber entschieden wird.

Auch durch Sie. Denn Sie nehmen an dieser Debatte teil, ob Sie wollen oder nicht. Sei es als Nutzer der zahlreichen digitalen Dienste oder als Bürger, der sich um eine lebenswerte Zukunft sorgt. In jedem Fall als Mensch, dessen Vermögen, vernünftig zu denken und zu entscheiden, durch KI einem nie gekannten Stresstest ausgesetzt wird.

Deshalb will dieses Buch eine Aufforderung sein zum Denken und zum Engagement in der Demokratie, will Anstöße geben zum menschlichen Denken in Zeiten Künstlicher Intelligenz. Damit Sie und wir alle eine menschliche Zukunft haben können.

1 Die beinahe Machtübernahme

»Die Demokratie setzt die Vernunft im Volke voraus, die sie erst hervorbringen soll.«
Karl Jaspers

Am 6. Januar 2021 hält die Welt den Atem an. Ein aufgebrachter Mob stürmt, angefeuert durch den eigenen Präsidenten, das Kapitol in Washington und damit das Herz der amerikanischen Demokratie. Viele der Teilnehmer filmen live und übertragen den Sturm auf Amerikas Demokratie über Facebook. Einige tragen die Videokameras montiert auf ihren Helmen. Der Facebook-Algorithmus schaltet dazu passende Werbung von Waffen und Militaria.[1] Neben einem Sturmgewehr, einer schusssicheren Weste und Kisten voller Munition erscheint die Werbezeile: *The survival gear that you are missing, Die Überlebensausrüstung, die Ihnen fehlt.* Die Teilnehmer hatten zuvor ihre Pläne ganz offen in den einschlägigen Gruppenchats besprochen, dennoch traf die Eskalation die Polizei für die Weltöffentlichkeit erkennbar unvorbereitet. Der Facebook-Algorithmus zeigte sich weniger überrascht, er sah schon den Bedarf an weiteren Waffen und Material voraus und regierte mit entsprechender Werbung. Ist er doch darauf trainiert, die Absichten der Nutzer aus ihren Datenspuren herauszulesen und ihnen dann entsprechend passende Kaufangebote zu unterbreiten. Er weiß, was Sie brauchen, wenn Sie tagelang mit Gleichgesinnten den Sturm auf die Demokratie geplant haben, schließlich hat er selbst seinen Beitrag dazu geleistet.

Denn er kann noch viel mehr, als nur Werbung verkaufen: Er kann Menschen mit ähnlichen Ansichten erkennen, schon bevor sie selbst wissen, dass sie diese Ansichten haben. Und daraus Absichten machen. Durch das permanente Durchpflügen gewaltiger Datenbestände, das Erkennen von Mustern und die Bildung milliardenfacher persönlicher Profile kann er größeres

Wissen über seine Datenlieferanten gewinnen, als diese selbst über sich haben, deshalb ihre Vorhaben, Wünsche und Gefühle besser verstehen als sie selbst und darum auch besser steuern. Durch gezielt verschickte Nachrichten und Beiträge kann er aus latentem Ärger über einzelne Angehörige von Bevölkerungsgruppen manifeste Fremdenfeindlichkeit machen, aus dem Ärger über bestimmte Politiker einen Hass auf die Demokratie. Bei diesem Prozess spielen Fakten kaum eine Rolle, sie sind sogar hinderlich. Der Algorithmus will Emotionen mobilisieren. Und er ist auf Steigerung aus, weil er Aufmerksamkeit gewinnen und nicht mehr abgeben will. Dabei hilft ihm eine zunehmende Polarisierung. Eine bestimmte psychische Disposition des Nutzers, die er aus dessen Daten herausgelesen hat, kombiniert mit anderen Datenpunkten reicht ihm aus, um ihn einschlägigen Gruppen vorzuschlagen, die ihn dort mit weiterem »passenden« Material versorgen. Ein radikaler Anführer, der den Ton vorgibt, hat sich dort schon längst gefunden: Er hat die meisten *Likes* der anderen Gruppenmitglieder eingesammelt, indem er den Ärger durch immer spektakulärere und unglaublichere »Enthüllungen« und emotionalisierende Aufrufe befeuert. Wenn er noch dazu selbst der Präsident ist, umso besser.

So hat sich die Eskalation der Proteste gegen das Ergebnis einer demokratischen Wahl in der wichtigsten Demokratie, die lange als Anführer der freien Welt galt, vor den Augen eines alles sehenden Algorithmus abgespielt, der dabei keineswegs passiv blieb: Ist er doch darauf trainiert, die Nutzer, die ähnlich ticken, zusammenzubringen, um gleich einer ganzen Gruppe ein passendes Produkt zu empfehlen und sie durch »Anstupsen« dazu zu bringen, es dann auch zu kaufen. An diesem Tag war die Zahl der entschlossenen potenziellen Käufer von Waffen und Militärausrüstung erfreulich hoch. Erfreulich für die Facebook-Algorithmen. Weniger erfreulich für die Demokratie.

Die Werbetreibenden konnten zufrieden sein: Ihre Botschaft erreichte eine zu hundert Prozent passende Zielgruppe, hatte

doch der Algorithmus mitgeholfen, sie passend zuzurichten: Eine Gruppe von Menschen, die buchstäblich zu allem entschlossen war, vor allem dazu, Gewalt einzusetzen. Der Algorithmus hatte seinen unmerklichen, aber wesentlichen Beitrag dazu geleistet, diesen Werbeerfolg zu ermöglichen. *Mission accomplished.* Dass dabei kurzzeitig die Demokratie am Abgrund stand, wurde als Kollateralschaden in Kauf genommen.

Der gesellschaftliche Preis für ein solches Werbetool ist hoch, ihn zahlen am Ende alle. Denn der Algorithmus teilt die Bevölkerung in Gruppen und Grüppchen auf, je nachdem wo ihre Schwächen liegen und sie am leichtesten manipulierbar sind, und bestärkt sie dann in diesem Gruppendenken mit »passenden« Nachrichten. Eine Untersuchung zeigte nach den Ereignissen des 6. Januar, dass in den Tagen nach der Erstürmung die Gruppen der Trump- und Biden-Sympathisanten jeweils zwei völlig unterschiedliche Wirklichkeiten bei Facebook angezeigt bekamen.[2] Das Ereignis wurde so zum Katalysator weiterer gesellschaftlicher Spaltung. Alles Folge eines Werbemodells, das auf Personalisierung setzt und dabei den gemeinsamen Grund von Überzeugungen, der die Basis von Gemeinsinn bildet, gezielt untergräbt. Und Folge von Algorithmen, deren Belohnungssysteme ausschließlich darauf programmiert sind, die Verweildauer der Nutzer und die Klickrate der Werbeanzeigen zu maximieren.

Ein Algorithmus, der alle überwacht und dabei die Strippen zieht, um die Überwachten in seinem Sinne zu lenken. Natürlich ist das eine Metapher. Die aber einen maßgeblichen Kern des neuen Überwachungskapitalismus trifft. So hat die amerikanische Autorin Shoshanna Zuboff das neue digitale ökonomische System bezeichnet, das die Daten der Nutzer sammelt und auswertet, um aus ihnen Verhaltensprognosen zu erstellen, die an den Warenterminmärkten der Aufmerksamkeitsökonomie hoch profitabel gehandelt werden.[3] Doch wer ist der Algorithmus, und wenn ja, wie viele? Hinter diesen hochkomplexen Programmen lassen sich die Ziele der Unternehmen erkennen, die sie

betreiben. Und einige von ihnen sind sehr tückisch. Man sollte glauben, dass sie ihr Ziel, die Klickrate zu erhöhen, um dadurch mehr Werbung zu verkaufen, lediglich dadurch verfolgen, dass sie dem Nutzer Informationen anzeigen, die er mag, damit er daraufklickt und das Produkt erwirbt. Das ist aber nur die Hälfte der Wahrheit. Im Online-Marketing werden sogenannte *Bandit*-(sic!)-Algorithmen, sowie Algorithmen für *Reinforcement Learning* (selbstverstärkendes Lernen) eingesetzt. Diese Algorithmen verfolgen nicht nur das Ziel, durch ihre Nachrichtenauswahl den Nutzer mit dem passenden Angebot zufriedenzustellen, sie versuchen vielmehr die Vorlieben der Nutzer so zu ändern, dass sie für den Algorithmus transparenter werden. Einem durchschaubaren Nutzer lassen sich die passenden Waren leichter verkaufen und am einfachsten durchschauen lassen sich Menschen mit radikalen politischen Ansichten. »Wie ein rationales Wesen lernt der Algorithmus, wie er den Zustand seiner Umgebung verändern kann, um die eigene Belohnung zu maximieren« so einer der weltweit führenden KI-Forscher über dieses perfide Modell.[4] Demnach wäre die politische Radikalisierung der Bevölkerung nicht nur eine unbeabsichtigte Nebenfolge ungehemmten Online-Marketings – sie wäre ihr Ziel. Die Verdummung und Spaltung der Gesellschaft als Geschäftsmodell, das die reichsten Unternehmen der Welt immer mächtiger macht und die demokratische Ordnung dabei unterminiert.

Noch erfolgt das eigentliche »Strippenziehen« in den sozialen Netzwerken durch recht dumme Algorithmen. Doch auch sie sind Teil der »Künstlichen Intelligenz«, also von Programmen, die sich bei dem, was sie tun, »selbst« »verbessern«, jedenfalls aber selbsttätig verändern. Und die dabei immer besser werden, wenn es darum geht, ihre Ziele zu erreichen. Wenn schon die dummen Bandit-Algorithmen von Facebook die Demokratie an den Rand treiben, was haben wir dann von wirklich intelligenten Algorithmen zu erwarten? Bei den am weitesten entwickelten Künstlichen Intelligenzen ist ein Programmierer nicht mehr er-