



Michael Altenburg

Artificial Intelligence als Treiber von Selbstbestimmung, Wohlfahrt und Fortschritt

Als Fazit ist festzuhalten, dass AI (Artificial Intelligence) als Treiber von Selbstbestimmung, Wohlfahrt und Fortschritt durchaus einen wahrhaft immensen Beitrag zu leisten imstande ist. Ob es dazu kommt oder ob uns eine zunehmend instabile Kette von Autoritarismus, Ausbeutung und Zusammenbrüchen bevorsteht, hängt aber nicht von technischen Fragen der Informatik ab, die bereits jetzt menschliche Intelligenz in Teilberei-

die verschiedensten Anwendungen sowohl privater wie kommerzieller Nutzer und Kombinationen der beiden.

Der globale Charakter des Internets wurde von Anfang an als auch kommerziell entscheidendes Merkmal erkannt und damit die Skalierbarkeit digitaler Geschäftsmodelle als wesentliche Erfolgsvoraussetzung, unmittelbar gefolgt von der möglichst schnellen Eroberung eines

Ernüchterung. Das passierte vor 20 Jahren, ist also überholt und damit eigentlich Geschichte.

Aber noch immer existiert die auf eine Beseitigung von Wettbewerb abzielende Tendenz des Gorilla Game: Wesentliche Segmente des globalen Marktes wurden inzwischen durch die sogenannten Big-Tech-Unternehmen²⁾ erfolgreich monopolisiert: The winners took all. Die Gorillas haben Platz genommen und versuchen, ihre beherrschende Stellung zu sichern, indem sie tatsächliche oder auch nur vermutete Wettbewerber strategisch gezielt aus dem Markt drängen oder übernehmen.

„Die Interneteuphorie der 1990er Jahre ist inzwischen starker Skepsis gewichen.“

chen weit hinter sich lässt, sondern vorrangig vom gesellschaftlichen Kontext, in dem mehr Transparenz, Rechenschaftspflicht, Dialogbereitschaft und Anpassungsfähigkeit der Entscheidungsträger an veränderte Realitäten die Voraussetzung dafür sind, dass intelligentere Formen von Vernunft und Verantwortung tatsächlich zum Tragen kommen

Als das Internet in den frühen 1990er Jahren begann sich weltweit auszubreiten, wurde diese Entwicklung von geradezu grenzenlosem Optimismus getragen: Weltweite, unzensurierte, nahezu kostenlose und augenblickliche Kommunikation von jedem mit jeder auf der ganzen Welt erschien damals wie ein Wirklichkeit gewordener Traum. Und das nicht nur für private Nutzer, sondern auch und vor allem für die Entwickler neuer Geschäftsmodelle zum Finden, Sortieren, Übermitteln und Managen von Daten. Der E-Mail-Verkehr begann zu boomen und parallel dazu eine Vielfalt von Suchmaschinen für

vom entsprechenden Geschäftsmodell beherrschbaren Segments des globalen Marktes.

Ein Bestseller jener Zeit hieß treffend „The Gorilla Game“¹⁾: Nicht die beste Geschäftsidee setzt sich immer durch, sondern derjenige, der imstande ist, mit seinem Standard den gesamten Markt möglichst schnell zu besetzen und so gegen Wettbewerb abzuschotten.

Der Trend zum Monopol bleibt

Das wurde zunächst vielfach blind geglaubt, sodass Geschäftsmodelle mit besonders hoher Cash Burn Rate bei Venture-Kapitalisten häufig mehr Erfolg hatten als solche, die lediglich mit einem nachhaltigen Geschäftsansatz zu überzeugen versuchten. Bewertungsexzesse waren die Folge und dann, im Jahr 2000, das Platzen der sogenannten Dotcomblase mit einer sich anschließenden deutlichen

Dazu kommt die ständige Weiterentwicklung nicht nur der Übertragungstechnologien wie gegenwärtig des G5-Standards für mobiles Internet und Mobiltelefonie und der Rechenkapazitäten hin zu höherer Schnelligkeit und größeren Speichern, sondern auch der Datenbearbeitungssoftware in Richtung selbstlernende Programme. Diesbezüglich spricht man von Künstlicher Intelligenz (KI) oder Artificial Intelligence (AI).

Die Interneteuphorie der 1990er Jahre ist allerdings inzwischen starker Skepsis gewichen. Das hat vor allem damit zu tun, wie mit den von den Nutzern in den Netzen hinterlassenen Daten umgegangen wird.

Eine besorgniserregende Zukunftsvision

Intelligente Programme können inzwischen aus den Daten von Nutzern weit-



Michael Altenburg

Luzern, Schweiz

In der Diskussion über Künstliche Intelligenz geht es oft nur um die ungeahnten technischen Möglichkeiten, die sich daraus ergeben. Altenburg geht in seinem Beitrag vielmehr den ethisch-philosophischen Fragestellungen der noch jungen Technologie nach. Die größte Gefahr sieht er jedoch weniger darin, dass die Maschinen dem Menschen überlegen sein werden und ihn dann beherrschen – wie es in anderen Beiträgen oft dystopisch ausgemalt wird. Vielmehr besteht eine nicht zu vernachlässigende Gefahr darin, dass sich dadurch computerbasierte Überwachungs- und Unterdrückungsregime ausbreiten könnten. Der Autor betont am Ende jedoch, dass sich die KI auch als Treiber von Wohlstand und Selbstbestimmung der Menschen erweisen könnte. Viel hänge dabei von Dialogbereitschaft und Anpassungsfähigkeit der Entscheidungsträger an veränderte Realitäten ab. Und der Mensch sollte vor allem eines vermeiden: Hybris. (Red.)

gehend auf deren Konsumgewohnheiten, auf deren politische Ausrichtung wie auf deren sexuelle Präferenzen schließen. Aufgrund dieser Einschätzungen wird anschließend im Wege gezielter Werbung auf das Verhalten bei Kauf- oder politischen Wahlentscheidungen Einfluss zu nehmen versucht in häufig so subtiler Weise, dass dem Adressat diese Beeinflussung nicht einmal zum Bewusstsein kommt. Dazu kommen neue technologische Möglichkeiten der lückenlosen Überwachung und Kontrolle, die, übrigens nicht nur in China, ohne Rücksicht auf die Privatsphäre, die Meinungsfrei-

heit oder sonstige Menschenrechte eingesetz werden.³⁾

Die Dystopie einer AI, die intelligenter ist als ihre Schöpfer, also einer superhuman Intelligence, welche die Menschen beherrscht anstatt umgekehrt, wird damit zu einer besorgniserregenden Zukunftsvision. Ist der in der Überschrift dieses Beitrages über AI zum Ausdruck kommende Optimismus gleichwohl zu rechtfertigen?

Dieser Optimismus orientiert sich zunächst an den bereits gemachten oder absehbaren Fortschritten in zahlreichen Anwendungsbereichen der Technik in allem, was mit Lernen, Analysieren, Vergleichen und Bewerten zu tun hat. Ein Paradebeispiel hierfür ist durch Deep Learning unterstützte Frühdiagnostik und Therapie in der Medizin. Ein anderes Beispiel sind die Effizienzgewinne in industriellen Produktionsprozessen durch Automatisierung.

Digitalisierung bietet viele Chancen ...

Im Finanzbereich erlaubt die Digitalisierung triftigere Risikobewertungen und kontinuierliche Monitoring-Prozesse. Eine Expertengruppe unter Leitung von Huw van Steenis hat der Bank of England (BoE) im vergangenen Juni Strategieempfehlungen für zukünftige von der Digitalisierung zu erwartende oder von ihr getriebene Strukturveränderungen vorgelegt,⁴⁾ denen sich der Mitte 2020 ausscheidende Gouverneur Mark Carney weitgehend angeschlossen hat.⁵⁾ Da ist von neuen digitalen Zahlungssystemen mit gewaltigem Potenzial vor allem in noch weniger entwickelten Ländern die Rede.

Diese Vision wurde womöglich angestoßen durch das Libra-Projekt von Facebook, das wegen Missbrauchsrisiken allerdings nicht ohne Zentralbankverankerung umgesetzt werden soll, also ohne Facebook. Ein weiterer Vorschlag bezieht sich auf den Aufbau und die Entwicklung einer digitalen Plattform zur Schließung der Finanzierungslücke für KMU. Die hohen Fixkosten eines profes-

sionellen Due Diligence für nur ein- bis mittlere zweistellige Millionenbeträge sind bekanntlich das größte Hindernis für KMU-Kredite, die daher ohne persönliche Sicherheiten des Unternehmers oder sonstige Besicherungen oft nicht zustande kommen.

... hat aber auch Nebenwirkungen

Der Aufbau einer umfassenden Unternehmensdatenbank bei der Zentralbank bei entsprechender Weiterentwicklung intelligenter Algorithmen könnte mittelfristig nicht nur eine hocheffiziente, kontinuierliche Kreditüberwachung über SaaS (Software as a Service) anstatt nur periodische Ratings oder Zertifizierungen ermöglichen, sondern auch eine substanzielle Entlastung bei der anfänglichen Kreditwürdigkeitsprüfung. Hierdurch könnten die Wettbewerbsnachteile unternehmensnaher, kleiner Finanzinstitute im Firmenkreditgeschäft entscheidend kompensiert werden.

Weiter beabsichtigt die BoE die explizite Einbeziehung der Klimabelastung von Unternehmensaktivitäten in zukünftige Stresstests. Christine Lagarde hat für die EZB ebenfalls eine stärkere Berücksichtigung von Klimaschutzzielsetzungen angekündigt.

Die negative Kehrseite dieser möglichen, beziehungsweise schon absehbaren Entwicklungen sind Arbeitsplatzverluste in größerem Umfang. Im Finanzbereich treten heikle Missbrauchsrisiken in Bezug auf die Daten hinzu. Wenn etwa eine Unternehmensdatenbank bei der Zentralbank die Finanzierungslücke bei den KMU schließen soll, stellt sich die Frage, wer den Zugang zu den Unternehmensdaten kontrolliert neben der Zentralbank, die sie insgesamt pro bono vorhält?

Sollte nicht das Unternehmen, das einen Kredit aufnehmen möchte, einen oder mehrere Anbieter aussuchen dürfen, die Zugang zu den Unternehmensdaten für den Zweck einer kompetitiven Offerte erhalten? Der mit seiner Offerte erfolgreiche Kreditanbieter würde den Zugang



für die Dauer des Kredites behalten. Und dann?

Hier zeigt sich unmittelbar, dass Zukunftsoptimismus eintrübende Umsetzungsschwierigkeiten nicht aus den Effizienzgewinnen der Programme per se rühren, sondern aus allerlei Nebenfolgen wie Arbeitsplatzverlusten oder einer sich grundlegend ändernden Wettbewerbsstruktur. Der Fortschritt hat also seinen ökonomischen und sozialen Preis und wer diesen zahlen soll, kann nur durch politische Prozesse entschieden werden und nicht durch ein Computerprogramm. Bei AI, also selbstlernenden Programmen, ist diese Problematik noch heikler.

Die Gefahr lauert woanders

Es sind also eher politisch ungelöste Fragen, häufig internationaler, möglichst globaler Abstimmung, die einem erfolgreichen Einsatz intelligenterer Software entgegenstehen und daher eher Anlass zu Sorge sein sollten als die Dystopie einer superhuman Intelligence, welche die Menschen beherrscht anstatt umgekehrt.

„Könnte AI überhaupt demokratische oder human-ethische Grundsätze berücksichtigen?“

Bei einem Zerfall der uns bislang vertrauten, weitgehend regelbasierten, liberalen Weltordnung in drei, vier oder noch mehr miteinander wetteifernden oder gar feindselig einander gegenüberstehenden Machtblöcke bleibt außerdem unabsehbar, ob sich nicht innerhalb dieser Machtblöcke computerbasierte Überwachungs- und Unterdrückungsregime ausbreiten könnten mit intern jeweils reibungslosen und effizienten Abläufen.

Vielleicht ist dies ein Hauptgrund für die gegenwärtige Konjunktur auch wissenschaftlicher Bemühungen, die um die Wechselbeziehungen von Informatik und Demokratie kreisen. An der Harvard Law School etwa wurde das Berkman Klein Center for Internet & Society eingerichtet,⁶⁾ bei dem Bundespräsident Stein-

meier Anfang November an einer Diskussionsrunde zum Thema „Ethics and Governance of AI“ teilnahm,⁷⁾ und am Berliner Weizenbaum Institut⁸⁾ arbeitet eine Forschungsgruppe „Demokratie & Digitalisierung“,⁹⁾ um nur diese zwei Beispiele zu nennen.

AI darf sich nicht verselbstständigen

Könnte selbstlernende AI denn überhaupt demokratische oder human-ethische Grundsätze berücksichtigen? Stuart Russell, ein britischer Informatikprofessor an der Universität Berkeley, hat auf die technische Möglichkeit belegbar emanzipatorischer AI hingewiesen. Wesentlich dafür ist nach Russell, dass AI so programmiert ist, dass sie sich nicht zu verselbstständigen imstande ist. Hierzu muss ein Rest von Unsicherheit in der Optimierungszielsetzung des Algorithmus eingebettet sein.

Wenn diese Unsicherheitsschwelle im Optimierungsprozess erreicht wird, hält das Programm inne und beginnt, nach Alternativen zu suchen, indem es sich nach

empirischen Verhaltensmustern orientiert, ehe es den alsdann modifizierten Optimierungsprozess fortsetzt.

Die Optimierungsalternative wird also abgeleitet vom Verhalten auf veränderte Umweltdaten reagierender Personen und nicht direkt von diesen veränderten Umweltdaten selbst, da das Programm ja menschenbeherrscht bleiben und sich nicht verselbstständigen soll. Russell nennt das „inverse reinforcement learning.“¹⁰⁾

Auf diese Weise wird die Berücksichtigung von Trade-offs möglich, beispielsweise langfristige Renditeoptimierung anstatt kurzfristiger Gewinnmaximierung. Die Trade-offs könnten sich aber auch auf andere Dimensionen beziehen

wie auf soziale Gerechtigkeit, CO₂-Verbrauch oder Minderheitenschutz. In welcher Richtung und Gewichtung die Optimierung letztlich vollzogen wird, ist also nicht im Programm vorgegeben, sondern ergibt sich aus seiner neuronalen Plastizität hinsichtlich der konkreten Präferenzen derer, denen das Programm nutzen soll.

Das klingt gut. Aber entscheidende Fragen bleiben offen: Wenn das Programm sich nach nur einer Person richten soll, wird es sich also in seinen Lerniterationen nach den Nutzenpräferenzen dieser Person richten. Und wenn es sich hierbei um eine kluge, verantwortliche Persönlichkeit handelt, wird man hinsichtlich des Ergebnisses unbesorgt sein können, da nach anfänglichen Fehleinschätzungen auch Korrekturen aufseiten des lernfähigen Programms sofort übernommen würden. Schlechter dürfte es aber um die Plastizität des Programms bestellt sein, wenn nach 10 oder 20 Jahren individuell unveränderter Nutzenpräferenzen die Unsicherheitsschwelle praktisch nicht mehr aktivierbar ist, da das Programm „sich“ inzwischen sicher ist, was und wie genau optimiert werden soll.

Noch schwieriger wird es, wenn sich das Programm nach den Nutzenpräferenzen nicht nur einer, sondern einer Vielzahl von Personen ausrichten soll. Was ist, wenn deren Präferenzen sich stark unterscheiden? Wessen Präferenzen sollen dann den Ausschlag geben? Hierüber diskutieren Philosophen, Ethiker, Juristen, Politologen und Ökonomen seit Menschengedenken.

Es stellt sich die Frage nach dem „Wie“

Kenneth J. Arrow, Begründer der Social-Choice-Theorie und Nobelpreisträger von 1972, lehnte die Vergleichbarkeit interpersonalen Nutzen grundsätzlich strikt ab und leitete daraus sogar ein Unmöglichkeitstheorem für Mehrheitswahlen ab, wenn sie minimalen demokratischen Anforderungen genügen sollen. Stuart Russell und andere¹¹⁾ meinen, eine Vergleichbarkeit sei bei doch nur marginalen

interpersonalen Unterschieden durchaus praktikabel und auch vertretbar, wodurch dann eine gesellschaftliche Nutzenoptimierung durch Addition individueller Nutzenpräferenzen möglich wird.

Politischer Missbrauch ist eine Gefahr

Bei Gesellschaften oder Systemen mit tendenziell vorherrschendem Konformismus, also mit einer irgendwie stark dominierenden „Leitkultur“, sei es hierarchisch, ideologisch, konfessionell, repressiv, korrupt, wie auch immer, werden „reformierende“ Verhaltenssignale aus der Gesellschaft oder dem System an das Optimierungsprogramm teilweise oder ganz ausbleiben, selbst wenn Änderungen in der realen Welt geeignete Anpassungen nahelegen. Der Zerfall der Sowjetunion ist hierfür ein gutes Beispiel. Der Optimierungsprozess über „inverse reinforcement learning“ bleibt also anfällig für diktatorischen Missbrauch.¹²⁾

Weil Russell die logischen Hürden eines interpersonalen Nutzenvergleichs nicht ernst nimmt, übersieht er die real immer irgendwie vorgegebenen gesellschaftlichen Macht- und Kontrollstrukturen. Es

rungsprogramms bei dumpf gleichbleibenden Vorgaben über die zu wählende Zielsetzung rasch abstumpfen und verkümmern wird.

Das System oder die Gesellschaft, der es eigentlich nützen soll, wird dann rasch instabil werden. Denn das Signal an das Programm, sich Änderungen der Realität anzupassen, wird nicht vom Programm selbst kommen, sondern nur durch eine entsprechende Änderung der Verhaltensmuster der Gesellschaft, die doch Herr des Programms bleiben soll. Wenn aber die Verhaltensmuster der Gesellschaft sich trotz Änderungen der Realität nicht wandeln, resultieren früher oder später Torpor, Rigor, Kataklysmus.

AI kann zum Segen werden, aber auch zu einer Gefahr

Als Fazit ist festzuhalten, dass AI als Treiber von Selbstbestimmung, Wohlfahrt und Fortschritt durchaus einen wahrhaft immensen Beitrag zu leisten imstande ist. Ob es dazu kommt oder ob uns eine zunehmend instabile Kette von Autoritarismus, Ausbeutung und Zusammenbrüchen bevorsteht, hängt aber nicht von technischen Fragen der Informatik

Ironie, denn er ist nach Bill Gates und Stephen Ballmer seit 2014 der dritte CEO von Microsoft, dem wegen seiner aggressiven Marktdominanz wohl unbeliebtesten Gorilla der 1990er Jahre mit dem Computerbetriebssystem Windows. Unter Nadella hat sich Microsoft erfolgreich neu erfunden als sympathisch-kooperativer Partner im Cloudcomputing: „Seit dem antiken Griechenland bis zum modernen Silicon Valley steht immer nur eine Sache dem Andauern von Erfolg, Relevanz und Impact entgegen, und das ist Hybris.“¹³⁾

Fußnoten

- 1) Geoffrey A. Moore: The Gorilla Game/How to Pick Winners in High Technology, New York, Harper Business, 1998
- 2) Nach einer Übersicht von PwC vom Juli 2019 gehörten hierzu sieben der nach der Marktkapitalisierung weltweit zehn größten Unternehmen: Microsoft, Apple, Amazon, Alphabet/Google, Facebook aus den USA sowie Alibaba und Tencent aus China: <https://www.pwc.com/gx/en/audit-services/publications/assets/global-top-100-companies-2019.pdf>
- 3) Hierzu hat sich der Autor bereits in einem früheren Aufsatz der „Kreditwesens“ vom 15.2.2018 mit dem Titel „Daidalos und die Gefährdungen der digitalen Transformation“ näher geäußert: <https://www.kreditwesens.de/kreditwesens/themenschwerpunkte/aufsätze/daidalos-tuecken-digitalen-transformation-id46982.html>
- 4) <https://www.bankofengland.co.uk/-/media/boef/files/report/2019/future-of-finance-report.pdf?la=en&hash=59CEFAEF01C71A551E7182262E933A699E952FC>
- 5) <https://www.bankofengland.co.uk/-/media/boef/files/report/2019/response-to-the-future-of-finance-report.pdf?la=en&hash=34D2FA7879CBF3A1296A0BE8DCFA5976E6E26CF0>
- 6) Details unter <https://cyber.harvard.edu/>
- 7) Hier das entsprechende Youtube-Video: <https://cyber.harvard.edu/story/2019-11/german-president-steinmeier-discusses-digital-technology-ethics-bkc-event>
- 8) Das Weizenbaum Institut ist ein Verbundprojekt der Universität der Künste Berlin, der FU Berlin, des Fraunhofer Instituts, des Wissenschaftszentrums Berlin für Sozialforschung, der Technischen Universität Berlin, der Humboldt Universität, der Universität Potsdam und des Bundesministeriums für Bildung und Forschung
- 9) <https://digidem.weizenbaum-institut.de/>
- 10) Stuart Russell: Human Compatible/Artificial Intelligence and the Problem of Control; New York, Viking 2019; Seiten 191 ff.
- 11) So etwa die meisten Autoren der umfangreichen Anthologie „Machine Ethics and Robot Ethics“ mit 35 Beiträgen, herausgegeben von Wendell Wallach und Peter Asaro bei Routledge, London und New York, 2017
- 12) So Arrow schon in seiner Dissertation „A Difficulty in the Concept of Social Welfare“ von 1951: „If we exclude the possibility of interpersonal comparisons of utility, then the only methods of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial.“
- 13) Zitat von Nadella aus dem Financial Times Artikel vom 19.12.2019 „FT Person of the Year: Satya Nadella“; <https://www.ft.com/content/0e8c3002-20c7-11ea-92da-f0c92e957a96>

„Nur eine Sache steht dem Andauern von Erfolg, Relevanz und Impact entgegen, und das ist Hybris.“

resultiert, womöglich von Russell so nicht gesehen oder jedenfalls kaum gewollt, eine Auffassung von Ethik, die den Wert einer Handlung nur noch nach ihren Konsequenzen beurteilt. Polemisch verkürzt passt hierzu die Maxime, der gute Zweck heilige die Mittel, mit der sich repressive Regime von Xi Jinping über Baschar al-Assad bis Recep Erdoğan zu rechtfertigen pflegen.

Zwar können und sollen hier weder ethische noch politische Fragen ausdiskutiert werden, aber rein AI-technisch muss festgehalten werden, dass die neuronale Plastizität eines auf „inverse reinforcement learning“ angelegten Optimie-

ab, die bereits jetzt menschliche Intelligenz in Teilbereichen weit hinter sich lässt, sondern vorrangig vom gesellschaftlichen Kontext, in dem mehr Transparenz, Rechenschaftspflicht, Dialogbereitschaft und Anpassungsfähigkeit der Entscheidungsträger an veränderte Realitäten die Voraussetzung dafür sind, dass intelligentere Formen von Vernunft und Verantwortung tatsächlich zum Tragen kommen können.

Recht prägnant hat ausgerechnet Satya Nadella ausgedrückt, worauf es ankommt, um einer tödlichen Erstarrung nach Jahrzehnten von Effizienz und Erfolg zu entkommen. Das ist nicht ohne