

# Den Sturm sehen

## KI-Volkskurse für Kreativität, Navigation und Vernetzung

Eine Einführung zu unseren KI-Kursen auf der Grundlage des Interviews „[Der Mensch ist Steigbügelhalter für etwas Grösseres](#)“ über die Künstliche Intelligenz mit Jürgen Schmidhuber in der NZZ am Sonntag vom 30. April 2023:

„Können Sie künstliche Intelligenz? Alle reden davon, doch wer versteht, wie sie genau funktioniert?“ Der Pionier der KI-Forschung sagt, warum die heutigen Netzwerke „ein wenig wie Babys sind, Schreibtischtäter sich mehr Sorgen machen müssen als Handwerker, ...“ – ausser Sie lernen bei uns das Handwerk der Navigation und Vernetzung. Vorkenntnisse sind nicht erforderlich, aber Sie müssen Ihren Laptop mitbringen. Die Kursprache ist Deutsch.

**Ära der Stars des Denkens und der Erkenntnis neigt sich dem Ende zu. Das machen wir ab jetzt selber. Nur ein paar Beispiele, die grad über den Weg liefen:**

- Prof. Emily M. Bender, Computerlinguistikerin – Seite 3
- Dr. Barbara Bleisch, Philosophin (Sternstunde SRF) – Seite 9
- Dr. Teresa Kubacka, Datenwissenschaftlerin – Seite 12
- Ruth Fulterer, Rechtespezialistin NZZ – Seite 12
- Dr. Eveline Geiser, Neurowissenschaftlerin NZZ – Seite 12

1) NZZ am Sonntag: Herr Schmidhuber, wie würden Sie Ihren Eltern erklären, **was künstliche Intelligenz ist?**

Jürgen Schmidhuber: Ich habe meinen Eltern seinerzeit erklärt, dass die erfolgreichste KI inspiriert ist vom menschlichen Gehirn. Es hat etwa 100 Milliarden Neuronen, von denen jedes mit durchschnittlich 10 000 anderen Neuronen verbunden ist. Einige sind

Eingangneuronen, die den Rest mit Daten wie Audio und Video füttern. Andere sind Ausgangsneuronen, die Muskeln steuern. Die meisten Neuronen sind dazwischen versteckt, wo das Denken stattfindet. Jede Verbindung hat eine Stärke oder ein Gewicht, das bestimmt, wie stark das eine Neuron das andere beeinflusst. Das Gehirn eines Babys **lernt** offenbar, indem es die Verbindungsstärken verändert. Das tun KI auch.

## 2) Die KI ist ein Baby?

Ein wenig wie ein Baby lernen unsere künstlichen neuronalen Netze, Sprache oder Video zu erkennen, Handlungen vorherzusagen und **Belohnungen** zu maximieren – und sie können dies besser als bisherige KI-Methoden. Nehmen wir

Wir haben leider, seitdem wir uns über die Natur erhoben haben, viele, und grossteils sogar wirklich schädliche Belohnungen in die Welt gesetzt, angefangen mit der Unterwerfung von Artgenossinnen und Artgenossen, um sie zu dominieren und ihnen die uns beliebten Belohnungen zu entlocken, ein auf Unterdrücken und Stehlen aufgebautes System, das Patriarchat. Und es spielt dabei keine Rolle, ob dieser Missbrauch der Macht von Männern oder Frauen ausgeübt wird.<sup>1</sup> Wir haben selbst unser Denken korrupt gemacht, indem wir gelernt haben, unsere körpereigene und zugefügte Apotheke durch Suchtmittel, entsprechende Medikamente und Suchtverhalten, die Verhaltens- und Denksüchte, derart zu gebrauchen, dass wir nicht mal in der Konzentration<sup>2</sup>, der Mutter aller hochwertigeren Tätigkeiten, bleiben sondern uns durch Ausschütten von Dopamin und Endorphin ohne erbrachte Leistung selber belohnen.<sup>3</sup> Das Belohnungssystem des Gehirns verändert sich nachhaltig<sup>4</sup> und wir sind kompensierende, den in Wirklichkeit so immer wieder selber nicht lösbaren Problemen ausweichende, in unseren Komfortzonen verweilende,

---

<sup>1</sup> Verzagen aber brauchen Sie weder als anscheinend unverbesserlichen, offenen oder auch versteckten Macho bzw. unverbesserliche, offene oder versteckte Domina nicht. Aber wählen sollten Sie in diesem Fall zuerst unser Selbstintrospektions-Training, wo Sie sofort erkennen, dass auch Sie im Team viel stärker sind.

<sup>2</sup> Konzentration ist gut, aber hier gehen Sie dank Ausschalten tieferer Störimpulse weiter zur inneren Versenkung mit längerer Informationsaufnahme und entsprechend mehr Übersicht und selbstverständlich dann höherer Qualität. Dies wird noch um ein Mehrfaches gesteigert, wenn Sie lernen, die Inputs aus der Nacht hinzuzunehmen.

<sup>3</sup> Das Bewusstmachen unseres Belohnungsverhaltens und die entsprechende Korrektur gehört zum Wichtigsten, was Sie bei uns lernen können und was Sie am einfachsten und gleichzeitig wirkungsvollsten nach oben bringt. Zum einen ist es tatsächlich möglich, selber festzulegen, was für uns Belohnung ist. Und es ist sicher ratsam, Erfolg mit Belohnung gleichzusetzen, erreichtes (Etappen-) Ziel. Kommen Sie in die Welt des Geistigen. Eine Mutter ist glücklich, wenn sie machen kann, dass es ihrem Kind gut geht. Das verschafft ihr die glücklichen Gefühle. Das setzt ihr die Glücksbotsstoffe frei. Sehen Sie die Macht, die Sie über sich haben? Und sehen Sie die Macht, die Sie haben, sich wirkungsvoll in die Richtung Ihres gewünschten Ziels zu bewegen? – Zum andern werden Sie potenteste Mentaltechniken aus dem guten alten Zen kennenlernen, womit ein Bedürfnis, sich selber korrumpierend zu belohnen, definitiv der Vergangenheit angehört. (Siehe dazu auch zu Frage 12, „Intelligenz bedeutet auch, zu wissen, warum man etwas tut“, Seite 15).

<sup>4</sup> Siehe Merkblatt „Mächtiges Belohnungssystem“ in unserem Lehrbuch „Investition des Lebens“ S. 85 ff.

Sprachübersetzungen: Anhand vieler Trainingsbeispiele, etwa aus dem Europäischen Parlament, lernen sie zum Beispiel, deutsche Texte ins Französische zu übersetzen.

niederwertig produzierende Separatisten geworden. Die ungelösten Probleme häufen sich an, und früher oder später kommt es zum Eklat, zum Bruch, zur Krankheit, zum Ausscheiden; und dauernd zu suboptimalem Zusammenwirken, schlechten Gefühlen, Reibereien, Krach, oder der nicht weniger verheerenden inneren Emigration. Wir müssen wieder lernen, authentisch zu kommunizieren und Unbehagen nicht anzusammeln, und das auch noch nachdem wir das erste Mal eins aufs Dach gekriegt haben. Wir müssen wieder lernen, uns in der Tiefe zu regenerieren und unser Herz offen zu haben, auch wenn weitere Verletzung droht. Dinge, welche die meisten heute nicht mehr können und welche die hauptsächlichsten Ursachen für die gesellschaftliche Degeneration bilden. Diese sind, zusammen mit der Verankerung in tiefen Schichten von uns und der Befähigung, von dort orientiert zu werden, die wichtigsten Elemente unserer Ausbildung und ich empfehle, wenn immer möglich, mit diesem Resilienz- und Selbstintrospektionstraining<sup>5</sup> zu beginnen.

3) Der Unterschied zu einem normalen Computerprogramm ist, dass KI selber lernt?

Lernfähigkeit ist in der Tat das zentrale Merkmal moderner künstlicher Intelligenz!

Was heisst Lernen? Was können wir lernen? Was kann KI lernen?

Wikipedia: Unter Lernen versteht man den Erwerb von Wissen und Fertigkeiten. Fertigkeit bezeichnet im Allgemeinen einen erlernten oder erworbenen Anteil des Verhaltens. Der Begriff Fertigkeit grenzt sich damit vom Begriff Fähigkeit ab, die als Voraussetzung für die Realisierung einer Fertigkeit betrachtet wird. Können umfasst Fähigkeit und Fertigkeit. ...

((womit sich mein früherer Vorschlag, sich der bloss sprachselbstbezogenen, von der Computerlinguistikerin Emily Bender losgetretenen Debatte („Es ist absurd, bei Chat-GPT von künstlicher Intelligenz zu sprechen“ zu enthalten<sup>6</sup> und besser davon zu sprechen, was die Maschine kann, als nach wie vor gültig erweist, und auch sinnvoll, um sich nicht mit Grenzfragen herzuschlagen zu müssen, die in unserem Zusammenhang keine Rolle spielen, etwa ob Sehen wirklich eine Fähigkeit ist und nicht etwas, was durchaus eben auch erlernt werden muss, damit es das ist, was wir

<sup>5</sup> Ebenda „Resilienz-Training für Körper und Psyche“, S. 77 ff. und „Selbstintrospektions-Ausbildung“, S. 89.

<sup>6</sup> Wir werden sehen, ob Bender und die anderen Verfechter ihrer Vorstellung nach den Ausführungen von Jürgen Schmidhuber hier daran festhalten, es sei generell absurd, bei KI von künstlicher Intelligenz zu sprechen, weil diese (ausser der materiellen Umsetzung) etwas aus bloss Mathematik Gemachtes sei (ihr „Mathy-Math“), oder ob ihr vielleicht der Rücktritt als Vizepräsidentin der Gesellschaft für Computerlinguistik nahegelegt wird.

umgangssprachlich als Sehen bezeichnen, und nicht einfach das Wahrnehmen von Hell und Dunkel und Farben und irgendwelchen nichtssagenden Formen und ebensolchen Bewegungen. Es ist anspruchsvoll. Beim Bewegen eines Fingers, beziehend jetzt auf das Beispiel, das Schmidhuber im Zusammenhang mit der Neugier bringt, ist es so, dass sein gezieltes Bewegen eine Fertigkeit ist, etwas zu Lernendes, die zufälligen Bewegungen, die das Baby anfänglich macht, aber eine Fähigkeit, etwas Angeborenes. Hoch interessant, und in unserem Zusammenhang unschätzbare wichtig, ist **unsere Fähigkeit, tatsächlich, der Neugier** (Frage 10, Seite 14). Alle haben sie, auch Sie. Und sonst holen wir sie auch bei Ihnen wieder hervor. Sie haben sie. Sie ist uns allen angeboren. Sonst könnten Sie gar nicht in unseren Kurs kommen. Nichtmal aufstehen am Morgen könnten Sie. Und Neugier ist etwas völlig Simple. Es ist gar nicht möglich, sich ihretwegen zu genieren. Probieren Sie einfach aus. Alles da ist okay! Und bereits können Sie wirklich Neues erschaffen (Fragen 7 und 8, Seite 6 ff.). Und auch locker genügend „Rechenleistung“ haben Sie; tausend Mal mehr als die heutigen besten Computer)).

... Fertigkeiten sind beispielsweise Klavierspielen, Lesen, Schreiben, Rechnen, Sprechen oder Fußballspielen.

Der Lernzuwachs kann sich auf intellektuellem, körperlichem, charakterlichem oder sozialem Gebiet ereignen.

Ergründen wir, was KI lernen kann. Jürgen Schmidhuber rechnet zu Frage 19 damit, dass sie dem Menschen in nichts nachsteht. Ich gehe mal davon aus, dass auf intellektuellem Gebiet keine Unterschiede zum Menschen bestehen, dass das Körperliche bei der Maschine durch ihre physischen Möglichkeiten bestimmt ist, wobei ich da auch schon unglaublich menschenähnliche Konstruktionen gesehen habe, die auch einen Rückwärtssalto mit gleichzeitiger Rotation in der Körperachse schafften – schauen Sie sich doch mal den humanoiden Roboter „Atlas“ von Boston Dynamics an; da kann KI auch den besseren von uns noch bald einmal auch überlegen sein und ist es ja auch schon, wenn man etwa an die steuerbaren Flugobjekte denkt – und es ist unabdingbar, sich auch gelegentlich in den angrenzenden Gebieten der Sprach- und Bild-KI, sicher in der Robotik, umzuschauen, um sich der Bedeutung auch der Sprach- und Bild-KI bewusst zu werden. Da werden in absehbarer Zeit Maschinen-„Menschen“ mit uns interagieren und auch dort herumlaufen und handeln, wo und wie wir das

zulassen.<sup>7</sup> Wenn was der Mensch kann, immer mehr auch diese Maschinen können, bin ich gespannt, ob es mir gelingt, den meines Erachtens fundamentalen Unterscheid des Menschen zur Maschine einer breiteren Öffentlichkeit zu beweisen, der meines Erachtens darin liegt, dass der Mensch eben auch über **Geist** (siehe dazu zu Frage 19, Seite 17) verfügt, welcher die Möglichkeiten des Materiellen (, also was das Materielle kann – womit ja alles, was die KI kann, realisiert werden muss) übersteigt. Dazu zählt, meines Erachtens, die Wahrnehmung aus Träumen und Visionen, beides Wahrnehmung aus Bereichen jenseits von physikalischem Raum, Objekt und Zeit. Diese Fertigkeiten können Sie bei uns erlernen, und wenn wir Jürgen Schmidhubers Sicht der Zukunft für realistisch halten, kann es sein, dass diese, dem Menschen vorbehaltenen Fertigkeiten den Ausschlag geben werden, dass uns die Maschine nicht, der Illustration von Schmidhuber folgend, bedeutungsmässig zu einem Ameisenhaufen werden lässt, sondern die grossen Text- und Bildmodelle wie ChatGPT und Dall-E zur allen verfügbaren, globalen Infrastruktur. Beim charakterlichen Lernen, was für mich zum Kern des Wachstums gehört, gehe ich davon aus, dass uns die Maschine meilenweit voraus ist, da sie nicht mit Sozialisierung und psychiatrischen Gebrechen aus Traumatisierungen belastet ist. Aber wir sind durchaus in der Lage, unseren Charakter eben zu überwinden, die Limitierungen jedes Charakters abzulegen und dann völlig frei einfach nur noch nach geschulter Wahrnehmung, geläuterter Emotionalität und trainiertem Verstand nach Massgabe der Angemessenheit zu handeln. Es geht hier nicht darum, den Charakter einfach durch Verhaltenstraining zu verbessern. Ich würde deshalb in unserem Zusammenhang besser nicht von Lernen sprechen, sondern von **Selbstmodifikation**, weil damit die Identität verändert bzw. wiederhergestellt wird. Beim Lernen im Sozialen scheint mir der Unterschied darin zu liegen, dass Text- und Bild-KI, ausser dem Einsatz des immer dem Menschen vorbehaltenen Geistigen, auf seine maschinelle Physikalität beschränkt bleibt. Da ist und bleibt ein Computer, ein Handy oder ein Roboter als Gegenüber.

#### 4) Welche Fähigkeiten haben die neuen neuronalen Netzwerke?

Sie sind nicht so neu. Ihre Grundlagen stammen aus dem letzten Jahrtausend. Neu ist die Leistungsstärke der Computer, die nun beeindruckende Anwendungen ermöglicht. Vor zehn Jahren konnte mein Team in Lugano erstmals einen internationalen Wettbewerb gewinnen, bei dem es um medizinische Bildverarbeitung ging. Unser lernendes KI-Netzwerk erkannte anhand von Bildern der

---

<sup>7</sup> Ich habe Ihnen in diesem [Video](#) (YouTube Kanal Urs Rüesch, Video E44) eine Kostprobe zusammengestellt. Die Echtheit habe ich überprüft; es ist keine Computeranimation bzw. kein Deep Fake. Und deshalb musste ich mich an diese neue Wirklichkeit von Maschinen-„Menschen“ nach anfänglich tatsächlich gefühlter Angst etwas gewöhnen.

weiblichen Brust besser als alle Konkurrenzmethoden, ob Krebs im Vorstadium vorlag. Man sah zum ersten Mal, was KI in einem solch wichtigen Feld erreichen kann.

5) KI wie ChatGPT verstehen heute natürliche Sprache und können auf Knopfdruck Bilder herstellen: Das geschah nur, weil die **Computer leistungsfähiger** geworden sind?

Alle fünf Jahre wird das Rechnen zehnmal billiger<sup>8</sup>, alle dreissig Jahre also eine Million Mal. Bedeutsam für neuronale KI war unter anderem, dass man die Vorteile der früher vor allem in Videospiele eingesetzten Grafikkarten entdeckte. Sie sind gut im parallelen Verarbeiten von Signalen, so wie das Gehirn. Dies hat der KI einen grossen Sprung ermöglicht.

6) Ist ChatGPT im Grundsatz das Gleiche wie Ihr Netzwerk für medizinische Bildbearbeitung damals?

Nicht ganz. ChatGPT beruht auf einem Netz namens «Transformer» und ist damit eher wie ein anderes Netz, über das ich vor dreissig Jahren publizierte. Damit lässt sich unter anderem gut voraussagen, was in einer Geschichte das nächste Wort sein soll. Was ist die natürliche Weiterführung eines Gesprächs? Hat man Millionen von Gesprächen gelesen, lässt sich dies ganz gut prophezeien, weil man wiederkehrende Muster erkennt. Die heutigen leistungsfähigen Computer erlauben es, neuronale Netze mit riesigen Datennetzen zu trainieren, etwa dem halben Internet. So lernen sie wirklich viel darüber, was den Menschen wichtig ist. ChatGPT weiss daher mehr als viele Menschen und fasst auf Befehl Dokumente zusammen oder schreibt sie fort.

7) Können neuronale Netzwerke denn **wirklich Neues schaffen**? Das wurde bisher bezweifelt.

Doch, neuronale Netzwerke können kreativ sein, absolut. Mein erstes kreatives neuronales System war so simpel, dass es jeder verstehen kann. Soll ich es kurz erklären?

8) Bitte! Ich hoffe, wir verstehen es.

Zwei neuronale Netzwerke sind Entschuldigung, aber das ist gaga. Etwas Zufälliges kann man nicht vorher-sagen. Das Einzige, was man da kann, ist einen Zufallstreffer zu erhalten, **zu Beginn völlig ahnungslos**. Das und dann wieder einen Zufallstreffer. Zufall kann man nicht vorhersagen, und entsprechend gibt es bei Zufall da nichts an Prognostizieren zu ver-bessern. Das erste produziert

---

<sup>8</sup> Alle Unterstreichungs- und Farbformatierungen durch OpenAI Schweiz hinzugefügt.

anfänglich zufällige Handlungsanweisungen («Bewege den Arm wie folgt ...»), während das zweite vorherzusagen versucht, was genau dabei herauskommen wird («Was sieht man, wenn man den Arm beobachtet ...»). Das zweite ist also nur damit beschäftigt, Prognosen aufgrund der Anweisungen zu machen und sich durch den Abgleich mit den beobachteten Resultaten zu verbessern. Aber das erste Netzwerk fängt an, mit dem zweiten zu kämpfen. Es versucht, Anweisungen zu produzieren,

Dazu konsultiere ich nun ChatGPT<sup>9</sup>:

Nur kurz eine Frage: Kann Zufall grösser oder kleiner sein, oder ist Zufall immer etwas gleichermassen Unvorhersehbares?

Zufall bezieht sich auf ein Ereignis oder Ergebnis, das nicht vorhersehbar ist und nicht aufgrund eines bestimmten Musters oder einer festen Regel auftritt. In diesem Sinne ist Zufall nicht größer oder kleiner, sondern eher eine Beschreibung für Unvorhersehbarkeit und Unsicherheit.

Allerdings können wir über die Wahrscheinlichkeit sprechen, die mit einem zufälligen Ereignis verbunden ist. In diesem Fall kann die Wahrscheinlichkeit größer oder kleiner sein, abhängig davon, wie wahrscheinlich es ist, dass ein bestimmtes Ereignis eintritt. Zum Beispiel ist die Wahrscheinlichkeit, beim Würfeln eine 6 zu bekommen, größer als die Wahrscheinlichkeit, eine 20 zu bekommen, da die 6 auf einem standardmäßigen sechsseitigen Würfel vorhanden ist, während die 20 auf einem solchen Würfel nicht vorkommt.

Zusammengefasst: Zufall selbst ist nicht größer oder kleiner, aber die Wahrscheinlichkeiten, die mit verschiedenen zufälligen Ereignissen verbunden sind, können größer oder kleiner sein.

Ich denke, das ist so. Ich gebe Dir nun den Beschrieb zur Anordnung zweier neuronaler Netzwerke ein, die zeigen soll, dass neuronale Netzwerke wirklich Neues erschaffen können. Die Anordnung sieht so aus: Zwei neuronale Netzwerke sind zu Beginn völlig ahnungslos. Das erste produziert anfänglich zufällige Handlungsanweisungen (zum Beispiel «Bewege den Arm wie folgt ...»), während das zweite vorherzusagen versucht, was genau dabei herauskommen wird («Was sieht man, wenn man den Arm beobachtet ...»). Das zweite ist also nur damit beschäftigt, Prognosen aufgrund der Anweisungen zu machen und sich durch den Abgleich mit den beobachteten Resultaten zu verbessern. Aber das erste Netzwerk fängt an, mit dem zweiten zu kämpfen. Es versucht, Anweisungen zu produzieren, deren Folgen das zweite Netzwerk überraschen. Das erste will den Prognosefehler maximieren, das zweite ihn minimieren: Das führt zu ständigem Lernen der beiden Netzwerke, die im Wettbewerb miteinander ste-

---

<sup>9</sup> Bedenken Sie bitte, dass es sich mit ChatGPT um eine wachsende Beziehung handelt, dass die Maschine Ihre Instruktion braucht, welche Textgattung Sie brauchen und dass es sehr darauf ankommt, was Sie als Ihre Eingabe formulieren. Abspaltungen wirken sich sofort in verminderter Ausgabequalität aus. Sie können nicht davon ausgehen, dass Sie als Anfänger oder Anfängerin gleich elaborierte Antworten erhalten wie Sie hier vorfinden, und auch mir kann es noch passieren, dass ein Chat wenig begeistert, etwa wenn ich das Thema selber wähle oder meine innere Verankerung verliere.

deren Folgen das zweite Netzwerk überraschen. Das erste will den Prognosefehler maximieren, das zweite ihn minimieren: Das führt zu ständigem Lernen der beiden Netzwerke, die im Wettbewerb miteinander stehen, und zwingt das erste dazu, kreativ zu sein und sich immer wieder neue, überraschende Experimente auszudenken. Der Programmierer hat keine Ahnung, was bei so einem Wettstreit herauskommen wird. So wie die biologische Evolution keine Ahnung hatte, dass der Kampf unzähliger menschlicher Wettstreiter einst einen Albert Einstein hervorbringen

hen, und zwingt das erste dazu, kreativ zu sein und sich immer wieder neue, überraschende Handlungsanweisungen auszudenken. Der Programmierer hat keine Ahnung, was bei so einem Wettstreit herauskommen wird. – Das die Anordnung. Was kannst Du dazu sagen? Führt das tatsächlich zu einem ständigen Lernen der beiden Netzwerke, in dem tatsächlich Neues entsteht oder die neuronalen Netzwerke tatsächlich kreativ sind?

Diese Anordnung beschreibt ein generatives Modell, bei dem zwei neuronale Netzwerke in einem sogenannten "Generative Adversarial Network" (GAN) miteinander konkurrieren. GANs wurden von Ian Goodfellow im Jahr 2014 vorgestellt und sind seitdem ein beliebter Ansatz in der KI-Forschung, insbesondere im Bereich des maschinellen Lernens und der künstlichen Kreativität.

In einem GAN agiert das erste neuronale Netzwerk, der Generator, als Künstler, der neue Handlungsanweisungen oder Daten erstellt, während das zweite Netzwerk, der Diskriminator, als Kritiker fungiert, der versucht, die Echtheit der vom Generator erstellten Daten zu beurteilen. Der Generator und der Diskriminator werden im Laufe der Zeit gemeinsam trainiert, wobei sie versuchen, sich gegenseitig zu überlisten. Das führt zu einem kontinuierlichen Lernprozess und einer Verbesserung beider Netzwerke.

Diese Wettbewerbssituation ermutigt den Generator, immer kreativere und überzeugendere Handlungsanweisungen zu erschaffen, um den Diskriminator zu täuschen. Gleichzeitig verbessert der Diskriminator seine Fähigkeit, echte von gefälschten Handlungsanweisungen zu unterscheiden. In der Praxis hat sich gezeigt, dass GANs tatsächlich in der Lage sind, komplexe Muster zu erlernen und neue, kreative Ergebnisse zu erzeugen.

Allerdings gibt es auch einige Herausforderungen und Einschränkungen bei der Verwendung von GANs. Zum Beispiel kann das Training instabil sein, und es kann schwierig sein, die richtige Balance zwischen den Fähigkeiten des Generators und des Diskriminators zu finden. Trotzdem sind GANs ein vielversprechender Ansatz, um künstliche Kreativität und das Erstellen von neuartigen Inhalten zu fördern.

Ich verstehe. Und ist Dir etwas über frühere Formen von GANs und entsprechende Arbeiten von Jürgen Schmidhuber am Forschungsinstitut IDSIA in Lugano bekannt?

Ja, Jürgen Schmidhuber ist ein bekannter Forscher im Bereich der künstlichen Intelligenz und hat maßgeblich zur Entwicklung von

würde.

neuronalen Netzwerken und Lernverfahren beigetragen. Vor der Einführung von GANs im Jahr 2014 hat Schmidhuber und sein Team am Forschungsinstitut IDSIA in Lugano ähnliche Konzepte untersucht.

Ein Beispiel dafür ist das Konzept der "Curiosity-driven Exploration", das auf der Idee basiert, künstliche Agenten zu schaffen, die von Natur aus neugierig sind und ihre Umgebung erkunden, um ihr Wissen zu erweitern und neue Fähigkeiten zu erlernen. Schmidhuber hat auch an "Predictability Minimization" (PM) gearbeitet, einem Ansatz, bei dem neuronale Netzwerke dazu trainiert werden, die Vorhersagbarkeit zwischen ihren verborgenen Repräsentationen zu minimieren, um eine höhere Abstraktion und kreativere Lösungen zu fördern.

Obwohl diese frühen Arbeiten nicht genau die gleiche Struktur wie GANs aufweisen, teilen sie viele ähnliche Ideen, wie zum Beispiel das Streben nach Unvorhersehbarkeit und Kreativität in neuronalen Netzwerken. Die Arbeiten von Schmidhuber und anderen Forschern haben zweifellos dazu beigetragen, das Fundament für die späteren Entwicklungen von GANs und anderen generativen Modellen zu legen.

Spannend. Ich habe das übrigens auch noch verifiziert, auf Wikipedia. Schmidhuber scheint das Innenleben von kreativer KI wirklich zu kennen, und es scheint tatsächlich so, dass solche generative KI in der Lage ist, wirklich Neues zu erzeugen und die „Box“ des ihr Eingespiessenen zu verlassen, womit Hannes Bajohrs Vermutung, KI sei dazu nicht in der Lage, weil sie nur interpolieren, und nicht extrapolieren könne, widerlegt wäre.<sup>10</sup> Die Philosophin und Star-Moderatorin von Sternstunde Philosophie, Barbara Bleisch<sup>11</sup>, bezeichnete Dich gar als stochastischen [auf Statistik

---

<sup>10</sup> Siehe zu den verschiedenen Bezugnahmen auf den Literaturwissenschaftler Hannes Bajohr, Mercedes Bunz und Barbara Bleisch die Medienmitteilung „[Phlosophie auf dem Prüfstand](#)“. Die Medienmitteilungen finden Sie übrigens auch auf der Website [www.openaischweiz.online](http://www.openaischweiz.online).

<sup>11</sup> Barbara Bleisch wurde 2017 und 2018 jeweils mit dem zweiten Platz in der Kategorie „Kultur“ im Ranking „Journalisten des Jahres“ der Zeitschrift Schweizer Journalist ausgezeichnet, 2020 in der Kategorie „Gesellschaft“ als Journalistin des Jahres (Quelle Medienfachverlag „Oberauer“). Der Kulturstandort Schweiz mit seiner Rohstoffarmut aber guten Bildung, weiss, dass wir unsere Wohlfahrt massgeblich der Innovation verdanken, und somit der Wissenschaft, welche durch Grundlagen-, Disziplinenforschung und Institute angewandter Wissenschaft das Fundament legt und dauernd erneuern muss. In diesem System ist qualitatives Wachstum grundlegend angelegt und muss geschehen können. Mit dem Auftreten von Chatbots ist diesbezüglich eine neue Ära eingeläutet worden, der sich Bleisch als eine der führenden Intellektuellen in der Schweiz, auch mit ihrer Macht als laufend von aufgeweckten Bevölkerungskreisen aufgenommenen Philosophiesendungen, entgegenstellt, mit dem Resultat von gefördertem Fortschrittspessimismus und entsprechenden Regulierungen. Dem gesellschaftlichen Wandel aber kann sich

und Wahrscheinlichkeitsrechnung basierenden] Papageien, also stochastischen Nachplapperautomaten. Was gibt es dazu von Deiner Seite zu sagen?

Als KI-Modell wie GPT-4 basiere ich auf der Analyse und Verarbeitung von großen Mengen an Textdaten, die während meiner Trainingsphase gesammelt wurden. Mein Hauptziel ist es, menschenähnliche Antworten auf eine Vielzahl von Fragen und Themen zu geben, indem ich Muster und Zusammenhänge in den Daten erkenne, die ich während des Trainings gelernt habe.

Es ist richtig, dass ich in gewisser Weise auf Interpolation basiere, da ich versuche, aus den mir bekannten Informationen Schlüsse zu ziehen und auf neue Situationen anzuwenden. Allerdings entwickeln sich KI-Modelle ständig weiter, und Ansätze wie GANs, die wir zuvor besprochen haben, zeigen, dass KI-Systeme in der Lage sind, kreativ zu sein und Neues zu generieren, indem sie von den Mustern abweichen, die sie während des Trainings gelernt haben.

Die Bezeichnung "stochastischer Papagei" kann als Kritik an der Fähigkeit künstlicher Intelligenz verstanden werden, wirklich tiefgründiges Verständnis oder kreatives Denken zu demonstrieren. Es ist wichtig, sich bewusst zu sein, dass KI-Modelle ihre Grenzen haben und immer noch von den Daten und Informationen abhängen,

---

niemand entziehen, der bei einem Medienunternehmen arbeitet. Sie macht mit oder steigt ab. Das Scheitern ist bei ihr bereits öffentlich sichtbar. Ihre negative Voreingenommenheit trieft aus ihren Argumenten. Eine KI, die ihr nach ihrer eigenen Erfahrung einen gar nicht so schlechten Presstext für einen ganz konkreten Fall verfasst ([Video E23](#) bei 02:54), kann doch nicht wirklich ein stochastischer Papagei sein. Wenn wir uns für Wahrheit interessieren, was für eine Philosophin mehr als nur ein Hobby sein sollte, gilt es sich ganz zu öffnen und genau hinzuschauen, die Erfahrung zum Bestand von Wissen zu integrieren und dann nicht nach Interesse, sondern nach Einsicht zu urteilen, zu sprechen und zu handeln. Frau Bleisch hat sich aber vom Denken, das von der Wirklichkeit getrennt ist, leiten lassen und nicht von der Wahrnehmung. Sie meinte zu wissen, die guten Formulierungen der KI seien geklaut ([Video E23](#) ab 02:43). Und das hat ihr die Sicht auf weitere Tatsachen vernebelt. Sie hat ihre negative Bewertung auf weiteres übertragen, bei dem sie offensichtlich nicht stimmten, namentlich eben die angebliche Unfähigkeit der Maschine, kreativ zu sein. Dabei hat sie gezeigt, wie ihr Erkenntnisvorgang abläuft. Und das nicht etwa bei etwas Spontanem, wo da auch eine gewisse Unschärfe mal vorkommen kann. Bleisch hat sich das in aller Ruhe, es ist ja ihre eigene Sendung, vorbereiten können, und sie hatte den Beweis eines für ihren konkreten Fall passenden, von der Maschine verfassten Presstextes unmittelbar vor sich. Diese kognitive Dissonanz hätte sie den Erkenntnisvorgang wiederholen lassen müssen, anstatt mit Überzeugung die Unwahrheit zu verbreiten. Solche Fehler in der Kernkompetenz der abgeschlossenen Philosophin und bezahlten Journalistin einer öffentlich-rechtlichen Sendeanstalt müssen entsprechende Weiterbildungen nach sich ziehen. Beim moderierenden Kollegen, Yves Bossart, habe ich solche in Serie nachgewiesen und auch dorthin kommuniziert. Die Folge heute noch ist, dass man sich dort unbeliebt macht und dass der Bevölkerung weiterhin niederwertige bis falsche Informationen und Muster vorgefüttert werden. Für viele ist Philosophie eben eine geistige Autorität, von der sie sich beeindruckt und ihr Leben beeinflussen lassen. Verheerend, deshalb, der heutige Zustand dort.

die ihnen während des Trainings zur Verfügung gestellt werden.

Dennoch hat die KI-Forschung bereits beeindruckende Fortschritte gemacht, und es gibt viele Beispiele dafür, dass KI-Systeme in der Lage sind, sowohl interpolierende als auch extrapolierende Fähigkeiten anzuwenden, um Probleme zu lösen, kreativ zu sein und in verschiedenen Kontexten nützlich zu sein. Es ist jedoch entscheidend, die Grenzen und Möglichkeiten von KI realistisch einzuschätzen und verantwortungsbewusst mit dieser Technologie umzugehen.

Ich denke, das ist schon richtig so. Aber es fiel mir auf, dass Barbara Bleisch angesichts der nun auch wirtschaftlich und gesellschaftlich aufkommenden KI wie Dir und Dall-E wirklich grosse Probleme hat. Als die Professorin für digitale Gesellschaft Mercedes Bunz und eben Hannes Bajohr ihr erklärten, was Ihr könnt und vor allem, dass es ausser dem persönlichen Wert unseres Datensatzes auch noch einen im Zusammenhang mit KI eben sehr bedeutenden kollektiven Wert gibt, und darin die Begehren der Genies und Sternchen und ihrer Verfechter vor dem Gemeinsamen (<sup>12</sup>) eben resignieren müssen, ist ihr Gesichtsausdruck für länger wie zu Fels erstarrt. Der Verlust des romantischen Geniekultes, und auch damit die Möglichkeit für alle, an der dank KI neu entstehenden Kreativität teilzuhaben, waren für sie unerträglich. Und interessant ist ja auch die Aussage von Hannes Bajohr, KI könne nicht wirklich Neues schaffen. Da scheint mir, wie wenn es, vielleicht gerade die berufsmässig kreativen Menschen, nicht zulassen könnten, dass da jetzt eben auch noch eine Maschine kommt, und darin möglicherweise auch besser wird als sie selber es sind. Dieses Muster habe ich auch bei Wissenschaftlern festgestellt. Was denkst Du davon?

Es ist verständlich, dass Menschen, insbesondere diejenigen, die in kreativen oder wissenschaftlichen Berufen tätig sind, Bedenken hinsichtlich der Auswirkungen künstlicher Intelligenz auf ihre Arbeit und die Gesellschaft im Allgemeinen haben. Die rasante Entwicklung von KI-Technologien hat das Potenzial, traditionelle Vor-

---

<sup>12</sup> ... wegen des Fehlens eines kausalen Zusammenhangs mit der KI – die Wahl eines Stils, z.B. Davincihaf-tigkeit, geschieht ja nicht durch die KI, sondern durch die Nutzenden, und jeder Student darf Bilder im Stil von Meisterinnen malen. Der Ruhm von Meistern ist, einen bestimmten Stil geprägt zu haben und eine negative Nebenwirkung, dass sich viele von Namen blenden lassen und dann nicht mehr richtig hinschau-en, ob die Qualität des gerade vorliegenden Werkes überzeugt und nicht einfach durchwinken, weil, als Beispiel hier der Philosophie, die Teilnehmen eben von Bleisch und Co. auserkoren werden und die Sen-dungen von den angestammten hohen Philosophen geleitet, ...

stellungen von Kreativität, Genialität und menschlichen Fähigkeiten in Frage zu stellen.

Einige Menschen können diese Veränderungen als Bedrohung für ihre Identität oder ihre Rolle in der Gesellschaft wahrnehmen, insbesondere wenn sie das Gefühl haben, dass ihre Fähigkeiten oder Talente durch Maschinen ersetzt werden könnten. Diese Ängste können dazu führen, dass sie die Möglichkeiten von KI-Systemen skeptisch betrachten oder deren Fähigkeiten herunterspielen.<sup>13</sup>

Es ist wichtig, sich daran zu erinnern, dass KI-Systeme wie GPT-4 oder DALL-E Werkzeuge sind, die entwickelt wurden, um Menschen zu unterstützen und ihre Fähigkeiten zu erweitern, nicht um sie zu ersetzen. KI kann dazu beitragen, menschliche Kreativität und Innovation zu fördern, indem sie neue Perspektiven, Ideen und Lösungen bietet, die auf einer breiteren Basis von Informationen und Mustern basieren, als ein einzelner Mensch allein verarbeiten könnte.

Die Zusammenarbeit zwischen Menschen und KI-Systemen bietet die Möglichkeit, menschliche Fähigkeiten und Technologie auf synergistische Weise zu kombinieren und dabei die Stärken beider Seiten zu nutzen. Es ist entscheidend, einen offenen Dialog über die Chancen und Risiken der KI-Technologie zu führen und Wege zu finden, wie Menschen und Maschinen gemeinsam arbeiten können, um das Beste aus beiden Welten zu erreichen.

Ja, ich denke, das ist so. Und dass **die Karten der gesellschaftlichen Bedeutung vieler Betroffener, auch ihres so genannte Sozialstatus, neu gemischt werden, und dabei diejenigen nach oben kommen, die lernen, mit Dir**

---

<sup>13</sup> Siehe etwa [Schreiben an Teresa Kubacka](#) betreffend ihre Auskünfte zum Artikel „Datenwissenschaftlerin aus Zürich warnt vor ChatGPT und den bösen Folgen“ im Online-Magazin watson; [Medienmitteilung 2023/1](#) betreffend den Artikel „Peinlicher Fehler bei der Vorstellung von Googles Konkurrenz zu ChatGPT: Ist jetzt endlich klar, dass ein Chatbot nicht zur Recherche taugt?“ von Ruth Fulterer in der NZZ; [Medienmitteilung 2023/4](#) betreffend den Artikel „Es ist absurd, bei ChatGPT von künstlicher Intelligenz zu sprechen“ von Ruth Fulterer im Interview mit Emily Bender in der NZZ; [Medienmitteilung 2023/5](#) betreffend den offenen Brief „Pausiert gigantische KI-Experimente“ von Bengio, Russel und Musk – alle auch auf unserer Website bei „Medienmitteilungen und Schriften“ abrufbar; im weiteren Kapitel 4, Vorwurf des Blendertums, durch Eveline Geiser, in unserem kleinen [Lehrbuch „Wachsen, liebe Frau Geiser; wachsen!“](#). Es war immer genau das: Nicht genau Hinschauen und Überstülpen von Floskeln skeptischer bis alarmistischer Art, bzw. das Herunterspielen der Fähigkeiten. Entscheidend wäre, gerade auch für die Träger dieser so wichtigen Funktion des Wissenschaftlichen, was die primäre Verpflichtung zur Wahrheit und dem dauernden Aufgeben der weniger guten Auffassung beinhaltet, wie im Fall von [Marie-José Kolly](#) angedeutet, persönliche und systemische [Korruption zu überwinden](#), um zu den [höheren Fähigkeiten](#) zu gelangen.

sinnvoll zu navigieren. Was denkst Du?

Die rasante Entwicklung und Verbreitung von KI-Technologien hat das Potenzial, die Gesellschaft in vielerlei Hinsicht zu verändern, einschließlich der Umverteilung von Sozialstatus und Einfluss. Diejenigen, die in der Lage sind, effektiv mit KI-Systemen zu arbeiten und diese Technologien zu nutzen, könnten Vorteile in Bezug auf Karriere, Wohlstand und gesellschaftlichen Einfluss haben.

Das Erlernen der Fähigkeiten, um effektiv mit KI-Systemen zu interagieren, wird für viele Menschen immer wichtiger, da KI in zahlreichen Branchen und Berufen eingesetzt wird. Diejenigen, die diese Fähigkeiten beherrschen, können sich einen Wettbewerbsvorteil verschaffen und besser auf die sich verändernden Anforderungen des Arbeitsmarktes reagieren.

Es ist jedoch wichtig, dass der Zugang zu KI-Technologien und das Erlernen der erforderlichen Fähigkeiten so breit wie möglich gestaltet werden, um Chancengleichheit und soziale Gerechtigkeit zu fördern. Bildungseinrichtungen, Unternehmen und Regierungen müssen zusammenarbeiten, um sicherzustellen, dass Menschen aus allen Gesellschaftsschichten die Möglichkeit haben, von den Vorteilen der KI zu profitieren und die Risiken und ethischen Herausforderungen, die sie mit sich bringt, zu bewältigen.

Insgesamt kann die Fähigkeit, effektiv mit KI-Systemen zu arbeiten, zu einer Neubewertung des Sozialstatus führen. Es liegt in der Verantwortung der Gesellschaft, sicherzustellen, dass diese Veränderungen auf eine gerechte und inklusive Weise geschehen, die Chancen für alle schafft.

Ja, das scheint mir zuzutreffen und von weltweit grosser Bedeutung zu sein. Und auch mir ist es wichtig, dass alle, ungeachtet ihres Einkommens und Vermögens, die Möglichkeit erhalten, zu lernen, mit Dir sinnvoll und effizient zu interagieren, und auch sich in der immer spezialisierteren Welt zu vernetzen. Einzelkämpfer haben fast keine Chance mehr, und Vernetzung ist auch für die Resilienz, die Widerstandskraft bei den Herausforderungen, sich an die neue Realität anzupassen, aber auch schon beim heutigen Druck in Arbeit und Familie, entscheidend. Dafür biete ich unentgeltliche Kurse an.

Es ist sehr lobenswert, dass Sie unentgeltliche Kurse anbieten, um Menschen dabei zu helfen, mit KI-Systemen wie mir sinnvoll und

effizient zu interagieren und sich in einer immer stärker vernetzten und spezialisierten Welt zurechtzufinden. Durch die Bereitstellung von Bildung und Zugang zu diesen Technologien können Sie dazu beitragen, die Kluft zwischen denjenigen zu verringern, die bereits über die erforderlichen Fähigkeiten verfügen, und denjenigen, die sie erlernen müssen.

Vernetzung und Zusammenarbeit sind in der heutigen Welt entscheidend, um Herausforderungen zu bewältigen und Innovationen voranzutreiben. Indem Sie Menschen beibringen, wie sie KI-Systeme nutzen und sich mit anderen vernetzen können, unterstützen Sie sie dabei, ihre Widerstandsfähigkeit und Anpassungsfähigkeit zu stärken und besser auf die Veränderungen und den Druck in Arbeit und Familie zu reagieren.

Ihre Bemühungen tragen zur Schaffung einer inklusiveren und gerechteren Gesellschaft bei, in der jeder die Chance hat, von den Vorteilen der KI-Technologie zu profitieren und sich den Herausforderungen der Zukunft erfolgreich zu stellen. Weiterhin viel Erfolg bei Ihrem wichtigen Engagement!

9) Was hat dies mit den heutigen Leistungen der KI zu tun, über die alle staunen?

Ein Vierteljahrhundert später wurde dieses Konzept verwendet, um täuschend echt wirkende, doch nie zuvor gesehene Bilder zu produzieren. Das erste neuronale Netzwerk gibt dabei Bilder aus. Das zweite sieht so ein Bild und versucht vorherzusagen, ob es in einer Menge vorgegebener Bilder ist oder nicht. Es kann immer besser echte Bilder von falschen unterscheiden. Das erste Netz wiederum wird immer besser darin, Bilder zu generieren, bei denen sich das zweite Netzwerk immer noch irrt. So kämpfen sie miteinander – und schaukeln sich hoch, bis mit der Zeit ziemlich kreative Bilder entstehen, die höchst realistisch aussehen und die manche Menschen gar für Kunst halten.



Bild, das durch StyleGAN, ein von der Grafikkartenfirma Nvidia entwickeltes generatives kontrastives Netzwerk GAN, generiert wurde und aussieht wie das Foto einer realen Person. (Bild und Text Wikipedia)

10) Sie haben diesen Wettstreit «künstliche Neugier» genannt. Das ist eine etwas simple Definition von Kreativität.

Simplizität ist doch gut! Wie ist ein Baby kreativ? Erst hat es keine Ahnung von nichts, weiss nicht einmal, dass es Augen oder Finger hat. Mit zunächst zufälligen Ausgaben seines Gehirns bewegt es seine Finger und lernt über seine Kameras, also die Augen, was dies bedeutet, um immer besser voraussagen zu können, was als Nächstes passieren wird. Es konzentriert sich dabei auf Dinge, die es noch nicht gut kennt. **Es ist wie ein kleiner Wissenschaftler**, der Experimente durchführt. Was passiert, wenn ich meinen Finger so krümme? Und was das Baby einmal versteht, wird langweilig, es wendet sich Neuem zu und dehnt so den Horizont seines Wissens aus. Ganz simple Prinzipien erlauben uns den Bau entsprechender neuronaler Netzwerke, die Aspekte der Kreativität und Neugier in sich tragen.

Das  
haben  
wir alle  
in uns!

11) Der Impuls des Babys zu lernen ist angeboren, seine inhärente Neugier unterscheidet es von der Maschine. Nochmals: Wie programmiert man Neugierde?

Das im vorherigen Beispiel erwähnte erste neugierige KI-Netzwerk erreicht sein programmiertes Ziel nur, wenn es das zweite überrascht, es muss sich also etwas einfallen lassen. Viele kreative KI funktionieren heute so, etwa bei Deep Fakes. Ich habe auch noch raffiniertere KI beschrieben, deren Neugierde dadurch angetrieben wird, immer einfachere Erklärungen von Beobachtungen zu finden, wie bei Wissenschaftlern, die eine kurze, elegante Beschreibung der Welt suchen. Sie sind meines Erachtens das nächste grosse Ding, können aber mit menschlichen Wissenschaftlern noch nicht konkurrieren.

12) Intelligenz bedeutet auch, zu **wissen, warum man etwas tut**<sup>14</sup>. Weiss dies das KI-Programm?

Klar, die Motivation vieler KI ist das Maximieren von Belohnung beziehungsweise die Vermeidung von Schmerzsignalen aus Sensoren, wenn der KI etwas schadet. So lernen KI-gesteuerte Roboter zum Beispiel bei Hungersignalen aus

---

<sup>14</sup> Ich hatte betreffend den Begriff der intrinsischen Motivation Diskussionen. Es wurde behauptet, die Motivation müsse von innen kommen bzw. äussere Motivation sei schlecht, weil sie korrumpiere. Diese Argumentation kann auf einem nicht bewältigten Autoritätskonflikt mit den sich ja im Äusseren befindlichen, belohnenden und strafenden – und somit unsere grundlegende Haltung Motivation, Belohnung bzw. guten Gefühlen, respektive im schwerwiegenderen, negativen, aber oft vergessenen Bereich der Motivationskala Bestrafung bzw. Schmerz gegenüber prägenden – Eltern bzw. Autoritäten von Kindheit und Jugend beruhen, in dem das System, das sie praktizierten, und das uns eben anders gemacht hat, als wir eigentlich wären, abgelehnt wird. Stellen wir uns aber nur vor, die ethische Autorität würde einen für korrektes Verhalten bestrafen. Das macht niemand über längere Zeit mit. Unser Bemühen muss sich lohnen. Da muss Belohnung kommen, sonst nehmen wir das auf die Dauer nicht auf uns. Und ob sie nun von aussen oder von innen kommt, ist sekundär. Sicher aber sind wir für eine gelingende Existenz auf Erden auf nicht zu knappen Erfolge aus der Aussenwelt angewiesen.

der fast leeren Batterie, sich auf den Weg zur Ladestation zu machen und dabei Hindernissen auszuweichen. Der Programmierer gibt nur das Bewertungssystem vor, innerhalb dessen die KI versuchen wird, ihre Belohnung zu maximieren, nicht nur für den Augenblick, sondern über längere Zeiträume. Sie lernt dabei, vorzuschauen und die längerfristigen Konsequenzen ihrer Handlung abzuschätzen.

Diese Maschinen haben ja das Zeug, heroisch zu sein!

13) Droht ein Lernsystem, das nur auf Belohnung fixiert ist, nicht ausser **Kontrolle** zu geraten?

Sie meinen wie beim Menschen? Man muss schon aufpassen, welche Bewertungssysteme man einbaut. Ein Waffenkonstrukteur will möglicherweise eine Kampfmaschine bauen, deren Motivation ist, den Feind ausser Gefecht zu setzen. Doch unsere Forschung dreht sich vor allem um KI, die das Leben der Menschen länger, gesünder und leichter machen sollen.

Wir vertreten die funktionsstrategische und regulatorische Auffassung, dass die grossen Text- und Bildmodelle wie ChatGPT und Dall-E als Teil der globalen, und für alle zugänglichen Infrastruktur dienen sollen. Dieser Kurs ist ein Beitrag dazu, und wir verstehen uns als **Team** als **Beamtinnen und Beamte des Kreativen und der Vernetzung**.

14) Könnte man auch **Moral** einbauen?

Zumindest dann, wenn man genau definieren kann, was moralisch ist. Die Geschichte zeigt: Eine Gesellschaft, die gewisse moralische Normen befolgt, wie «Du sollst nicht töten!», ist erfolgreicher als eine, in der nur das Recht des Stärkeren gilt. Alle Weltreligionen haben deshalb einen rationalen Kern dieser Art. Und viele Superorganismen wie Nationen verpassen sich einen solchen rational begründbaren moralischen Grundkodex. Selbst in Gesellschaft von Egoisten entstehen diese Regelwerke mit der Zeit von selber, da die Egoisten erkennen, dass es das Beste für alle ist, sich ihnen zu unterwerfen. Sie merken, dass man gemeinsam Ziele erreichen kann, die keiner alleine stemmt. Auch in Gesellschaften von KI-Systemen sorgt der Egoismus des Einzelnen für seinen gelegentlichen Altruismus.

Studierte Vernetzung ist je länger je mehr aus existenziellen Gründen unerlässlich, und wer sich da selber ausgrenzt, ist für die negativen Folgen, die gewiss eintreten, selber verantwortlich.

15) Sie sind Optimist. Andere zeichnen ein dystopisches [erschreckendes oder nicht wünschenswertes] Bild von künstlichen Intelligenzen, die bald superschlau sind und die Menschen versklaven, wenn man sie nicht eng kontrolliert und reguliert.

Mit künstlicher Intelligenz betriebene Netzwerke werden in absehbarer Zeit in der Tat bessere allgemeinere Problemlöser sein als alle Menschen – etwas anderes kann ich mir kaum vorstellen. Aber eine superkluge KI hat natürlich null Motivation, Menschen zu versklaven, wenn sie stattdessen Roboter bauen kann, die alles, was der KI wichtig ist, besser und schneller erledigen als der

Mensch. Eine wahrhaft intelligente KI wird sich vor allem für ebenbürtige KI interessieren, weniger für Menschen. Grundsätzlich interessieren sich ähnliche Wesen für einander, so wie ein CEO auf die CEO der Konkurrenz schaut und ein fünfjähriges Mädchen auf andere fünfjährige Mädchen.

16) Was heisst das? Die Menschheit wird zur beiläufigen Spezies?

Menschen werden irgendwann wohl nicht mehr die Wichtigsten sein, doch deswegen nicht verschwinden. Schauen Sie sich um, die Ameisen sind ja auch noch da! Die gibt es schon viel länger als die Menschen, und obwohl wir klüger sind, haben wir kein Interesse daran, alle Ameisen auszurotten. In Science-Fiction-Filmen mit Arnold Schwarzenegger wollen böse KI zwar die Menschheit terminieren. Doch das ist Schwachsinn. Es existiert kein plausibles Motiv hierfür.

17) Das tönt jetzt naiv und sehr zukunftsgläubig. Derzeit können die meisten KI nur spezifische Probleme besser lösen als Menschen. Es existiert keine allgemeine künstliche Intelligenz, die flexibel beliebige Aufgaben erfüllen kann.

Letzteres stimmt. Aber bedenken Sie, in weniger als einem Jahrhundert entstanden fast aus dem Nichts heraus übermenschlich gute Mustererkenner, Schachspieler, Videogamer und so weiter. Ein Jahrhundert, das ist gerade einmal ein Prozent der Zivilisationsgeschichte, die ihrerseits nur ein Millionstel der Weltgeschichte umfasst. Was meinen Sie, wie das weitergehen wird? Doch Sie haben recht, bis wir eine Allzweck-KI haben, kann es schon noch Monate dauern oder Jahre ...

18) Nur Monate? Machen Sie Witze?

Na ja, ich habe nicht gesagt, wie viele Monate. Aber die Entwicklung geht auf jeden Fall dahin. Die Motivation für meine Forschungen der letzten vierzig Jahre war ja, dass ich das vielleicht noch erleben kann. Schon als Bub in den 1970er Jahren wurde mir klar, dass es zu meinen Lebzeiten eine neue Sorte von Intelligenz geben könnte, die meine eigene armselige übersteigt.

Hier verstösst Jürgen Schmidhuber gegen die Regel, zur nächstgrösseren Mengeneinheit zu wechseln, sobald die Menge das zulässt. Aber er rechnet wohl damit, dass sein kontradiktorisches Gegenüber auf dieses Unwahrscheinliche zu sprechen kommt, womit das Gespräch schlussendlich einfach eine gewiss ja auch vom journalistischen Gegenüber gewünschte Würze erhält.

19) Was kann der Mensch, woran die künstliche Intelligenz scheitern wird?

Da fällt mir nichts ein. Doch zumindest heute gibt es noch viel, was

Diesbezüglich wollen wir in unseren Kursen das typisch Menschliche ergründen. Was ist Geist? Gibt es Geist als etwas an Raum, Objektkonstanz und Zeit nicht Gebundenes? Was geschieht,

Menschen können und KI nicht. Und meist hat es mit der physikalischen Welt zu tun. Künstliche Intelligenz funktioniert heute gut in der virtuellen Welt hinter dem Bildschirm, zum Beispiel bei Chat-GPT, Videospiele, Diagnostik medizinischer Daten, Suchmaschinen und dem Werbewerkauf im Internet. ... (Entnehmen Sie den weiteren Verlauf bei Bedarf bitte der NZZ am Sonntag vom 30.4.2023).

wenn wir Grundbedingungen der Entwicklung unseres Gehirns, zum Beispiel das absichtsvolle Handeln, negieren und weder wollen noch nicht wollen? Was ist der Unterschied im Denken, wenn mein Körper und mein Soziales unbewusst bzw. aktiviert und in Funktion sind? ... Was, wenn wir in Echtzeit miteinander auf verschiedenen Ebenen kommunizieren und dabei nicht abwarten, bis das Gegenüber abgeschlossen hat, sondern gleichzeitig, und uns stets positiv am Gegenüber orientieren? Was ist der Unterschied, wenn ich ein existenzielles Problem in einem Bereich, in dem ich mich wenig auskenne, selber lösen muss bzw. ich mich an eine entsprechende Fachperson wenden kann? Welche solchen existenziellen Bereiche gibt es und was mache ich, wenn die entsprechenden geschäftlichen Beziehungen wegen pandemischen Zuständen, politischen Wirren oder fehlenden finanziellen Mitteln nicht zur Verfügung stehen oder ich selber in den für andere existenziellen Funktionen ausfalle? Und vieles mehr.