

Essay: KI, Transhumanismus und Verkörperung

Josua Kugler, 4119892, Mathematik

27. August 2023

Kürzliche Entwicklungen im Bereich KI durch die „Transformator“-Architektur und Large Language Models haben großes gesellschaftliches Interesse am Thema KI geweckt. Durch die große Bandbreite an Themen, auf denen Sprachmodelle wie ChatGPT oft sehr hilfreiche und menschenähnliche Antworten aus einer riesigen Datenbasis erzeugen können, stellt sich fast von alleine die Frage, was Menschen jetzt eigentlich noch besser können als diese Sprachmodelle. Die scheinbare Abstraktionsfähigkeit, Adaptabilität und Lernfähigkeit dieser Modelle machen viele Fähigkeiten für Maschinen zugänglich, die vorher Menschen vorbehalten waren und teils sogar als Unterscheidungsmerkmale angeführt wurden. Hinzu kommt oft ein radikaler Materialismus, der Leben einfach als biochemischen Algorithmus auffasst und damit zu weiteren Kontroversen führt.

Kurz könnte man sagen: Zwischen der Reduktion auf eine biologische Maschine und der Hybris, sich selbst kopieren zu können und als Maschine wiederzuerschaffen, befindet sich die Vorstellung davon, wer der Mensch ist, in einer Krise. Das Menschenbild schwankt also zwischen Großartigkeit und Selbstverachtung. Den eigenen Wert auf das von sich gezeichnete übertrieben selbstverherrlichende Bild setzend, leidet der Mensch letztlich an der Spannung zwischen dem Glauben an dieses Selbstverständnis und dem gleichzeitigen Anzweifeln jeglichen Selbstwertes. Der Lösungsversuch basiert auf der Rettung des Selbstbildes, was aber von vornherein zum Scheitern verurteilt ist. Insbesondere kann das Menschenbild nicht einfach durch Konstruktion immer besserer Maschinen gelöst werden. Glaubt man den Dystopien einiger Futuristen und Transhumanisten, so führt die beschriebene Entwicklung sogar

zur Ablösung des Menschen durch eine neue Generation künstlich geschaffener Wesen. Das Menschenbild des Tech-Enthusiasmus und der Transhumanisten trägt also konsequent durchgedacht bereits die Abschaffung des Menschen in sich und ist damit in gewissem Sinne kein Menschenbild, sondern ein Anti-Menschenbild.

Daher ist es unabdingbar, dieses Menschenbild kritisch zu diskutieren.

Grundsätzlich wird oft, ohne dies zu hinterfragen, davon ausgegangen, dass der „Geist“ eines Menschen vom Körper getrennt betrachtet und durch Simulation dieses Geistes bzw. des Gehirns ein Mensch hinreichend abgebildet werden kann. Dabei werden jedoch zwei Aspekte vernachlässigt. Zum einen handelt es sich bei dieser Vorgehensweise um Zerebrozentrismus, also den ungerechtfertigt starken Fokus auf das Gehirn. Um den Geist eines Menschen zu verstehen, müssen viele weitere Prozesse berücksichtigt werden, die nicht neuronaler Natur sind (z.B. gewisse Hormone etc.). Diese können nicht durch Simulation des Gehirns und nicht einmal durch Simulation des gesamten Körpersystems adäquat abgebildet werden. Der daran anknüpfende zweite vernachlässigte Aspekt besteht darin, dass der menschliche Geist nicht notwendigerweise allein auf biologische Prozesse reduziert werden kann.

Die unzulängliche Berücksichtigung dieser beiden Aspekte führt in der Entwicklung zu folgendem Vorgehen. Basierend auf der Annahme, der Mensch sei schlicht die Summe der biologischen Prozesse, die in seinem Körper ablaufen, wird versucht, Subprozesse oder Prozesszusammenhänge näherungsweise nachzubilden. Beginnen wir beispielsweise mit dem Gehirn, so sind neuronale Netze ein erster Versuch, sich der Komplexität des Gehirns zu nähern. In der aktuellen Forschung wird immer wieder versucht, die Funktionsweise des Gehirns durch die Architektur der Algorithmen noch genauer abzubilden. Ein Beispiel: Der „Lernprozess“ bei klassischen Neuronen Netzen wird dadurch simuliert, dass Gewichtungen verschiedener Parameter durch Feedback aus den gegebenen Daten verändert und den korrekten Gewichtungen angenähert werden. Das geschieht durch back-propagation: Wird dem Algorithmus ein Datenobjekt zur Klassifizierung übergeben, so wird die Differenz zur korrekten Klassifizierung berechnet und dann mithilfe der Kettenregel zurückgerechnet, welchen Einfluss die einzelnen Gewichte auf diese Abweichung hatten. Stück für Stück rückwärts durch den Algorithmus

gehend, werden dann die Gewichtungen angepasst. Im menschlichen Gehirn allerdings läuft dieser Prozess andersherum ab. Darum gibt es erste Versuche von KI-Architekturen, die in ebendieser Reihenfolge Gewichte abändern (forward-forward). Die Resultate sind bisher nicht überzeugend, sicherlich auch deswegen, weil in diesem Bereich noch vergleichsweise wenig Forschung geschehen ist. Generell ist es aber gut vorstellbar, dass die Algorithmen zukünftig immer mehr Details der Funktionsweise des Gehirns übernehmen und damit seiner Simulation graduell näherkommen (auch wenn das natürlich aufgrund der Komplexität des Gehirns sehr spekulativ bleiben muss). Allerdings muss an dieser Stelle ja nicht aufgehört werden. Stattdessen könnten weitere für das menschliche Denken und Handeln relevante Prozesse integriert werden. Mit Hormonkreisläufen beginnend, könnte man versuchen, sich graduell der Komplexität der im Körper ablaufenden Prozesse anzunähern.

All dem wird man natürlich und auch zu Recht entgegen, dass derartige Vorstellungen völlig unrealistisch sind. Trotz zahlreicher Versuche (z.B. Human Brain Project) hat die Forschung keine signifikanten Resultate in dieser Richtung erzielt. Tatsächlich ist es sogar so, dass nicht nur keine Fortschritte erzielt werden, sondern im Verlauf der Jahre immer deutlicher wird, wie sehr die Komplexität des menschlichen Gehirns im Besonderen und des gesamten menschlichen Körpers im Allgemeinen jegliche simulierbare Komplexität weit übersteigt. Allein schon, dass ein Gehirn analog funktioniert, genügt, um jede digitale Simulation immer nur Näherung bleiben zu lassen. Als kontinuierliche Prozesse entziehen sich die Abläufe in unserem Körper a priori der digitalen Berechenbarkeit als solcher.

Ungeachtet der Frage, ob es jemals technisch möglich sein wird, einen gesamten Menschen auch nur annähernd zu simulieren, muss dennoch untersucht werden, ob der Unterschied zwischen Mensch und Maschine einzig in der ungleich höheren Komplexität des Menschen begründet ist. Ist der Mensch am Ende doch einfach eine unglaublich vielschichtige (biologische) Maschine? Der fundamentale Komplexitätsunterschied mag als Argument genügen, um das Konzept des „mind-uploading“ zu dekonstruieren, aus humanistischer Sicht bleibt diese Diskussion jedoch unbefriedigend. Gibt es keinen prinzipiellen Unterschied zwischen Mensch und KI, so wäre doch zumindest die philosophische Grundlage der Transhumanisten gesichert

(auch wenn man an der technischen Umsetzbarkeit ihrer futuristischen Ideen große Zweifel äußern kann).

Aus allen bisher vorgebrachten Argumenten lässt sich folgern, dass menschliche Forschung und Entwicklung sehr weit davon entfernt ist, irgendetwas in Richtung „mind uploading“ oder Ähnliches zu erreichen. Vermutlich wird es sogar nie vollständig möglich sein, weil der Komplexitätsunterschied so groß ist. Trotzdem ist die Frage noch nicht befriedigend geklärt, ob es überhaupt einen prinzipiellen, nicht nur graduellen Unterschied zwischen Mensch und Maschine geben muss. Kann Bewusstsein künstlich erzeugt werden? Könnten wir theoretisch Maschinen bauen, die bewusst sind? Wäre Bewusstsein nämlich einfach eine Folge biochemischer Prozesse, dann könnte nach den bisher vorgebrachten Argumenten noch nicht ausgeschlossen werden, dass auch Maschinen Bewusstsein entwickeln. Ist Bewusstsein aber transzendent, so würden sich Menschen an dieser Stelle fundamental von Maschinen unterscheiden. Das kann aber nicht allein aus der Außensicht geklärt werden. Insbesondere dann, wenn Maschinen eine perfekte Simulation erreichen sollten, wird deutlich, dass diese Frage nicht empirisch klärbar ist.

Daher ist an dieser Stelle eine gesellschaftliche Debatte dringend vonnöten. Die oft als selbstverständlich angenommenen Prämissen von z.B. Transhumanisten müssen kritisch hinterfragt und diskutiert werden. Sollten die aktuellen Entwicklungen tatsächlich dazu führen, dass zwischen Menschen und guten Simulationen nicht mehr unterschieden wird, so sollte dies zumindest bewusst geschehen und nicht einfach unreflektiert als Folge eines blinden Fortschrittsglaubens.