

## **Anmoderation:**

Seit Jahren wird über KI gesprochen, über „Künstliche Intelligenz“. Aber nie waren das Reden und Raunen, das Ahnen und Alarmschlagen so laut wie in den vergangenen Monaten, seit ChatGPT für alle greifbar macht, was KI bedeuten könnte. Unterschätzen wir die Chancen dieser Entwicklung? Oder die Gefahren? Verharmlosen wir Künstliche Intelligenz? Oder dämonisieren wir sie? Beruhigen wir uns mit dem Satz, zuletzt sei das alles doch nur eine Maschine? Oder stellen wir uns die immer weiter entwickelten Spielarten Künstlicher Intelligenz viel zu menschlich vor, menschlicher, als sie jemals werden können? Die Kulturwissenschaftlerin Nathalie Weidenfeld hat 2018 gemeinsam mit ihrem Mann Julian Nida-Rümelin ein viel beachtetes Buch geschrieben: „Digitaler Humanismus. Eine Ethik für das digitale Zeitalter“. Die jüngsten Entwicklungen nimmt sie nun in ihrem Gedanken-zur-Zeit-Essay zum Anlass, Dinge und Begriffe aus ihrer Sicht ein wenig geradezurücken:

## **Beitrag:**

Es war zu schön, um wahr zu sein. Gerade in den letzten Wochen durften wir dank der jüngsten digitalen Neuerungen wie ChatGPT, Music LM oder des Metaversum von Facebook weiter träumen: von Robotern, die Gedichte schreiben und Musik komponieren oder einem neuen gigantischen virtuellen Raum, in dem wir Unglaubliches erleben können und uns neue Freunde machen können. Und jetzt das: In einem offenen Brief haben Forscher und Unternehmer, darunter Figuren wie Elon Musk und Apple Mitgründer Steve Wozniak, zu einem Moratorium für zukünftige Forschungen mit Künstlicher Intelligenz aufgerufen. Schon Nick Bostrom und Steven Hawking hatten uns ja vor Jahren davor gewarnt, dass die bösen Roboter uns irgendwann einmal beherrschen, vielleicht sogar versklaven würden. Dass diese Ängste von Hollywood bereitwillig in vielen Filmen gespiegelt wurden, verstärkte sie noch.

Dabei sollte uns eigentlich stutzig machen, dass immer auch schon zeitgleich ein entgegengesetzter Topos zirkulierte, und zwar der des guten, moralisch und spirituell überlegenen Roboters. Denken Sie nur an den unschuldigen kleinen Roboter-Jungen aus Spielbergs Film *AI* - den Charles Dickens nicht besser hätte erfinden können - inklusive der engelshaften Roboter-Wesen, die am Ende des Films als erlösende Himmelsboten auf die Erde kommen, um dem Jungen seinen sehnlichsten Wunsch nach Liebe zu erfüllen.

Also wie jetzt, möchte man fragen: Ist KI extrem gut oder extrem böse?

Wenn diese Art polarisierender Charakterisierungen auftauchen, sollten wir misstrauisch werden und vermuten, dass hier etwas am Werk ist, das weniger mit Realität als mit Projektion zu tun hat. Erklären lässt sich das in diesem Fall gut mit der jahrtausendalten Praxis der Primitivismus - dem kulturellen Reflex der westlichen Welt, sich ein Spiegelbild zu suchen, auf das alle verehrten oder verhassten Charaktereigenschaften in übertriebener Form projiziert werden können. Dies kann wahlweise zur Erbauung dienen („schaut mal, wie zivilisiert und edel wir im Vergleich zu denen da sind“) oder zum Zurechtstutzen des eigenen Größenwahns („schaut mal, wie schlecht wir im Vergleich zu denen da sind“). Zur Zeit der Griechen fungierten als eine solche Projektionsfläche die „Barbaren“, in der Zeit der Aufklärung waren es die wilden „Indianer“, in den USA waren es die Afroamerikaner – um nur ein paar Beispiele zu nennen. Offensichtlich benötigen unsere Gesellschaften immer wieder die imaginäre Konstruktion eines als radikal anders markierten Anderen, an dem wir uns abarbeiten können, um die eigene Identität zu festigen oder auch in Frage zu stellen.

Heute haben wir allerdings ein Problem: die politische Korrektheit. Woran kann man sich da noch abarbeiten? An den Indian..., pardon, den Native Americans? An den African Americans? Eher nicht. Da trifft es sich gut, dass sich als imaginäre Projektionsfläche die Roboter anbieten. Das Problem dabei ist, dass wir uns unserer fantasmagorischen Überzeichnung und primitivistischen Fantasmen nicht bewusst sind, wir importieren achtlos falsche und irreführende Adjektive in unsere Beschreibungen von KI, erwecken falsche Hoffnungen. Darunter auch die, dass Digitalisierung eine quasi heilsgeschichtliche Funktion erfüllen wird. Die Verbindung von Technologie und Erlösung ist nicht neu. So wurde selbst das Automobil Anfang des 20. Jahrhunderts nicht nur als technisches, sondern auch als spirituelles Wunder gefeiert. Henry Ford glaubte fest daran, dass sein „Kraftwagen“ „Frieden auf Erden“ bringen würde. „Unsere neue Denk- und Handlungsart“, so predigte er, „wird uns eine neue Welt schaffen, einen neuen Himmel und eine neue Erde, wie sie die Propheten seit unvordenklichen Zeiten ersehnt haben.“<sup>1</sup>

Ist es wirklich ein Zufall, dass Apple-Stores weltweit wie protestantische Kirchen aussehen, lichtdurchflutet, reduziert auf das Wesentliche, nur als Fluchtpunkt nicht das Kreuz, sondern eben der Apfel? Die in Blau getauchten Bilder der Digitalisierung, die man im

---

<sup>1</sup>. Ford, Henry. Philosophie der Arbeit, Dresden 1919, S. 47

Internet findet, bestätigen die postmillenaristischen Fantasien eines diesseitigen Paradieses bestehend aus Immanenz, Transparenz und einer all-vernetzten-Gemeinde.

Es lauert aber noch eine andere Gefahr im Diskurs über KI. Das sind die regressiven Fantasien, die der Mensch mit den neuen Technologien zu erfüllen sucht. Von dem Traum eines digitalen Schlaraffenlandes, in dem Roboter uns chauffieren, für uns kochen und saubermachen, über die animistische Fantasie, in einem unbelebten Etwas - einer Maschine - den idealen Freund oder die ideale Partnerin zu finden, bis hin zum erfüllten kindlichen Wunsch nach Unverwundbarkeit. Der letzte Clou im Bereich dieser Wunscherfüllung ist das „live forever program“, das nichts Geringeres verspricht als ein ewiges Leben im Metaversum. Wir hätten dann jeder und jede unseren eigenen digitalen Avatar, der als unheimlicher Doppelgänger auch dann noch „weiterlebt“, wenn wir schon längst gestorben sind. Genau das verspricht die Firma Somnium Space, die auf der Basis so vieler Informationen über einen Menschen wie möglich (Aufnahmen seiner Bewegungen und seiner Sprache etc...) einen Avatar schaffen möchte, mit dem auch nach dem Tod seines Identitätsgebers Familienmitglieder im Metaversum kommunizieren können. Freud sah in dem Bild des Doppelgängers eine „Dementierung der Macht des Todes“<sup>2</sup>. Der Doppelgänger fungiert dabei als eine Art Talisman für das eigene unsterbliche Fortbestehen - wie etwa im alten Ägypten die reichverzierten Mumien. Die digitalen Avatare von Somnium Space sind nichts anderes als die Talismane unserer Zeit. Wir lächeln heute über den Aberglauben der alten Ägypter und sind uns dessen nicht bewusst, dass wir unsere eigene Form des Aberglaubens in Form von KI-Projektionen betreiben. Die digitale Technologie kann sich noch so rationalistisch präsentieren, die Träume, die wir mit ihr verbinden, sind es nicht. Sie sind Ausdruck einer kindlichen Sehnsucht nach Allmacht und Unverwundbarkeit.

Digitale Technologien aber werden an der *conditio humana* nichts ändern. Der Mensch muss mit Verletzungen, Trauer und Tod umgehen, das ist der Kern des Erwachsenwerdens. Und keine Technologie kann uns auf magische Weise davon befreien. Und selbst wenn wir mit einem mehr oder wenigen gelungenen Avatar unserer Eltern im Metaverse zu kommunizieren versuchen - es würde sich nie um eine echte Kommunikation handeln. Echte Kommunikation findet zwischen Personen statt, die darauf vertrauen, dass ihr

---

<sup>2</sup> „Das Unheimliche“ (1919).

Gegenüber jemand ist, der bestimmte Wünsche und Intentionen besitzt. Das aber haben nur Menschen. Dass auch hierzulande Kommunikation schon teilweise an Software delegiert wird, mag Kosten sparen, menschlich ist es nicht. So ist es nur richtig, dass der Deutsche Ethikrat vor kurzem eine Stellungnahme veröffentlicht hat, in der klar gefordert wird, dass Maschinen Menschen nicht ersetzen sollen – in dem Sinn, dass wir von Maschinen nicht erwarten, dass sie die Rolle von intelligenten, verantwortungsvollen und (ethisch) handelnden Akteuren übernehmen. Sofern aber neue digitale Technologien menschliche Handlungsräume erweitern, und sofern wir sie dazu einsetzen können, die Umwelt zu schonen oder Krankheiten besser zu erkennen, können und sollten wir KI nutzen.

Was wir im aktuellen KI Diskurs dringend benötigen ist Klarheit im Kopf und in der Sprache. Wie in Medien – und leider auch durchaus von Seiten mancher Wissenschaftler – über KI gesprochen wird, ist oft allzu imaginär und libidinös aufgeladen. Von diesen diffusen Erwartungen profitieren vor allem die großen Tech-Firmen aus dem Silicon Valley. Denn während die Gesellschaft sich träumerischen Visionen hingibt, können Tech-Giganten in Ruhe unermüdlich weiter daran arbeiten, KI zu optimieren, um uns auszuspionieren und als Werbekonsumenten noch besser verkaufen können.

Bemühen wir uns also, bei den Fakten zu bleiben. Und diese lauten: Künstliche Intelligenzen sind Maschinen. Als solche sind sie nicht in der Lage, Bedeutung zu verstehen und Sinn zu stiften. Sie haben weder Gefühle, noch können sie Urteile treffen oder den Gehalt eines Kunstwerks begreifen. Sie sind weder das radikal Andere, dem wir verdanken, dass wir endlich wieder in unserer Überheblichkeit zurechtgestutzt werden, noch das Allheilmittel, das unsere präödpalen Fantasien von Unverwundbarkeit und Unsterblichkeit erfüllt. Sie können kombinieren, wiedergeben, rechnen und wir können sie so einsetzen, dass sie unser Leben besser, also menschlicher machen. Nicht mehr und nicht weniger. Wenn wir uns diese einfache Tatsache nicht bewusst machen, laufen wir Gefahr, dass wir vor lauter Fantasterei übersehen, worauf es wirklich ankommt, und das ist: unser Leben human zu gestalten. Mit Raum für Uneinheitliches, menschliche-Fehlleistungen, Unzulänglichkeiten und Verletzlichkeiten und einem Menschenrecht auf Privatsphäre und menschliche Kommunikation.

Nathalie Weidenfeld