



Digitaler Humanismus

Warum wir ihn dringender benötigen denn je

**Julian Nida-Rümelin |
Nathalie Weidenfeld**

Prof. Dr. Dr. h. c. Julian Nida-Rümelin, Staatsminister a. D., ist Professor emeritus für Philosophie und politische Theorie an der Ludwig-Maximilians-Universität München und stellvertretender Vorsitzender des Deutschen Ethikrats.

Dr. Nathalie Weidenfeld ist Kulturwissenschaftlerin und Schriftstellerin von Sachbüchern und belletristischen Werken.

Alle paar Monate gibt es einen neuen Aufschrei. Erst, weil Künstliche Intelligenz (KI) so tun kann, als wäre sie ein Restaurantangestellter, der Reservierungen aufnehmen kann; dann, weil KI angeblich kreativ ist und Beethovens nie vollendete 10. Symphonie fertig schreibt; und jetzt, weil ein neues Textproduktionssystem namens ChatGPT in beeindruckender Weise Texte zusammenbauen kann.

Im Frühjahr 2023 schlugen Elon Musk und andere IT-Größen wie Apple-Mitgründer Steve Wozniak ein KI-Moratorium vor, während der »Futurist« Ray Kurzweil vorhersagt, dass in spätestens 20 Jahren Künstliche Intelligenzen den sagenumwitterten Sprung in die »Singulartät« schaffen werden (»singularity is nearer«) – will heißen, dass KI auf wundersame Weise zu beseelten Wesen mutieren werden, die sich mit Menschen auf das Innigste verbinden. Im Gegensatz zu Elon Musk ist er davon überzeugt, dass die neuen digitalen Technologien dem Menschen Unsterblichkeit und Unverwundbarkeit bescheren werden.

An dieser Stelle kreuzt sich Hyperrationalität mit Animismus – ein auf kindlichem Narzissmus beruhender Glaube an die Allmacht des eigenen Wunschdenkens, der sich auch in der Beseelung von Unbeseeltem äußern kann, wie es Sigmund Freud in »Totem und Tabu« beschreibt. Die Diskussion über KI scheint sich zunehmend auf die Frage zu fokussieren,

ob uns die KI erlösen oder vernichten wird, eine in den abrahamitischen Religionen tief verwurzelte Alternative.

Aus Sicht der IT-Firmen ist es zwar wünschenswert, dass die neuen Technologien faszinieren, aber weniger wünschenswert sind apokalyptische Ängste, die sich am Ende geschäftsschädigend auswirken könnten. Wie als Firma also umgehen mit dem berühmten »uncanny valley«, also der in der Psychologie bekannten Tatsache, dass ein allzu menschenähnliches KI-System bei Menschen eher Befremdung und Ablehnung auslöst? Vor allem wenn man weiß, dass positiv besetzte (Wunsch-)Träume inhärenter Teil jeder erfolgreichen Marketingkampagne sind. Die Lösung: Man oszilliert eben zwischen diesen beiden Polen: Ja, die neue KI-Technik wird und kann uns erlösen, weil sie uns menschenähnliche Partner an die Seite stellt, die uns nie enttäuschen werden. Außerdem: keine Sorge, KI ist kein Mensch.

Dieser paradoxe Diskurs wird durch die Kommunikationsstrategie von ChatGPT belegt. Einerseits lassen die Programmierer ChatGPT »sagen«, dass es nur ein Softwaresystem sei und kein Mensch. Andererseits »spricht« das System nicht nur von einem »Ich«, wenn es mit den Usern »kommuniziert«, sondern es »täuscht« den Usern auch Gefühle vor, etwa wenn ChatGPT »behauptet«, es »täte ihm leid«, wenn es keine Antwort produzieren kann. Doch es gibt dringlichere Fragen, denen wir uns stellen müssen, zum Beispiel:

- Wie können wir den Datenschutz besser organisieren, damit Datenopolenz zugunsten des Gemeinwohls, also der Verfügbarkeit umfangreicher Datenbestände für Forschung, Smart-City-Entwicklung und zivilgesellschaftliches Engagement, mit dem Schutz der informationellen Selbstbestimmung kombiniert wird und diese nicht gegeneinander ausgespielt werden?
- Wie kann Europa der Monopolstellung der US-Firmen etwas entgegensetzen und zwi-

schen dem kommerzorientierten US-amerikanischen Weg der Digitalisierung und dem staatszentrierten Weg Chinas eine eigene europäische Strategie einer am Menschen orientierten KI-Entwicklung (»human-centered AI«) einschlagen?

- Wie kann sichergestellt werden, dass menschliche Ansprechpartner in der Kommunikation mit Behörden und Firmen verfügbar und digitale Texte und Stimmen von menschlichen unterscheidbar bleiben?

Der Digitale Humanismus kann als philosophische Grundlage bei der Klärung dieser Fragen helfen. Er macht klar, dass Softwaresysteme zwar rechnen, aber nicht denken können, dass sie weder zu Gefühlen noch zu moralischen Entscheidungen fähig sind und dass sie menschliche Kommunikation nicht ersetzen können. Der Mensch ist keine Maschine, auch keine algorithmisch organisierte digitale – und eine Maschine wird nicht zu einem Menschen, auch wenn sie menschliches Verhalten perfekt simuliert. Nichts liegt dem Digitalen Humanismus ferner, als den technologischen Fortschritt zu behindern. Im Gegenteil: Wenn Softwaresysteme als Akteure, und das heißt als

Der Digitale Humanismus gefährdet nicht den technologischen Fortschritt, sondern fordert ihn.

Personen, interpretiert werden, müsste in der Tat die weitere Entwicklung sofort gestoppt werden, schon um Menschenrechte nicht zu verletzen. Ein empathie- und vernunftfähiger Roboter, der für uns Entscheidungen trifft und Mitgefühl hat, dürfte nicht wie eine bloße Sache, ein technisches Hilfsmittel behandelt werden. Der technologische Fortschritt wird durch seine humanistische Rahmung nicht gefährdet, sondern gefordert.