

Abschnitt III:

Zukunftsauswirkungen der Künstlichen Intelligenz



Einführung und Übersicht zum Abschnitt III

I. Wachsmuth, M. Wilker

Die theoretische Künstliche-Intelligenz-Forschung hat sich die folgende Aufgabe gestellt: Jeder Aspekt menschlicher Intelligenz soll so präzise beschrieben werden, daß er durch ein künstliches System – einen Computer – simuliert werden kann. Daß dies prinzipiell möglich sei, bezieht sich auf eine Ausgangshypothese, nämlich daß Menschen – zumindest als intelligent Handelnde – informationsverarbeitende Systeme sind. Verband sich mit den Ursprüngen der KI (Dartmouth Conference 1956) in erster Linie der Gedanke, auf diese Weise Erklärungsmodelle für Intelligenz zu erhalten, so wird bei Anhängern der sog. „harten KI“ die Auffassung vertreten, daß es keine prinzipiellen Unterschiede zwischen menschlichen Denkleistungen und ihrer Nachbildung auf Maschinen gibt. Das hieße, daß sich menschliche Intelligenz auf dem Computer reproduzieren läßt und – von der anderen Seite her betrachtet – daß menschliche Intelligenz sich auf Informationsverarbeitung reduzieren läßt.

Die Künstliche-Intelligenz-Forschung befaßt sich aber auch mit der Konstruktion von Systemen, die kognitive Leistungen erbringen, um die theoretisch entwickelten Konzepte und Techniken nutzbringend einzusetzen: Auf informationsverarbeitende Maschinen werden Eigenschaften menschlicher Intelligenz, etwa Schlußfolgerungsfähigkeiten, übertragen; dadurch sollen geistige Tätigkeiten des Menschen unterstützt, verstärkt oder entlastet werden. In dieser häufig als „weiche KI“ charakterisierten Richtung beschränkt sich die Erwartung der erzielbaren Leistungen auf Systeme, die in abgrenzbaren Bereichen als Werkzeuge menschliche Intelligenzfähigkeiten substituieren oder sie möglicherweise auch weit übertreffen, die aber nicht menschengleich sein werden. Selbst wenn die Realisierung einer intelligenten Maschine unerreicht ist und bleibt, beginnt die Forschung über „Künstliche Intelligenz“ zum Verständnis des Denkens und des Menschen beizutragen.

Durch den in der Künstlichen Intelligenz vertretenen Anspruch, *geistige* Tätigkeiten des Menschen zu formalisieren und auf Maschinen zu übertragen, wird eine neue Qualität technischer Mittel berührt. Die mit einer „denkenden Maschine“ verbundenen Möglichkeiten und utopischen Vorstellungen gehen einher mit Umgestaltungen der Arbeits- und Lebenswelt. Sie berühren die Neubewertung beruflicher Qualifikationen, die Veränderung von Werten, ja selbst des *Menschenbildes*. Die Implikationen solcher Vorstellungen werden in diesem Buchabschnitt diskutiert. Nach Ansicht der Autoren haben die beiden als „harte“ bzw. „weiche“ KI bezeichneten Standpunkte völlig verschiedene Menschenbilder zum Gegenstand, die die Art und Weise, wie Künstliche Intelligenz in gesellschaftliche Wechselwirkungen tritt, betreffen und damit die Entwicklung der zukünftigen informationstechnischen Gesellschaft fundamental beeinflussen könnten. So wird gegenwärtig nicht ausschließlich davon gesprochen, inwieweit „intelligente“ Funktionen in technischen Systemen realisiert werden, sondern auch davon, daß maschinelle „Akteure“ als Kooperationspartner in menschliche Arbeitszusammenhänge eintreten.

Die in den Medien dazu diskutierten Fragen schwanken zwischen euphorischen Zukunftsvisionen einerseits und greifen andererseits menschliche Urängste vor künstlichen Kreaturen auf. Der mit den gegenwärtig begrenzten Erfolgen und möglicherweise inhärenten Schwächen solcher Systeme Vertraute vermag sich schwerlich vorzustellen, daß in absehbarer Zeit eine Konkurrenz zum intelligent handelnden Menschen in seiner Gesamtheit erwachsen könnte. Eine denkende Maschine, die sich ihrer selbst bewußt ist, bleibt vielleicht immer dem Bereich der Science Fiction zuzurechnen. Aber selbst dann äußert sich in den geführten Debatten das offensichtliche Bedürfnis, sich mit einer solchen Zukunftsvision zu beschäftigen, bevor sie auch nur ansatzweise Realität wird. Aus den angesprochenen Gründen muß ein solcher Diskurs Fragen des Menschenbildes einschließen.

Zwar gibt es kein konsensfähiges, universell gültiges Menschenbild; aber es gibt Menschenbilder, die in bestimmten Epochen und bestimmten Regionen eindeutig dominieren und darum relevant sind. In den hochentwickelten Industriegesellschaften war in den letzten Jahrzehnten (innerhalb und außerhalb des

Wissenschaftssystem) ein Menschen- und Weltbild erfolgreich, das den Menschen als zwar hochkomplexe, aber durchgängig determinierte und kausalen Gesetzmäßigkeiten unterworfenen Maschine darstellt. Die „Maschine“ Mensch, wie der sie umgebende Kosmos, ist prinzipiell verstehbar und deshalb symbolisch rekonstruierbar. Dieses Axiom ist die Grundvoraussetzung des Postulats einer „harten“ KI. Das menschliche Hirn gilt als Sitz des Bewußtseins, welches als die Summe der im Hirn ablaufenden biophysikalischen Prozesse aufgefaßt wird.

Schon immer ist ein Menschenbild kritisiert worden, das menschliches Handeln auf kausale Prozesse reduziert. Innerhalb des Wissenschaftssystems wird dies in jüngerer Zeit vor allem in der Chaostheorie bestritten: Kausalität ist nur ein Ausschnitt aus einer umfassenderen Wirklichkeit; für die „harte“ KI könnte das bedeuten, daß Bewußtsein sich als ein Phänomen erweist, das zwar von kausalen Prozessen begleitet wird, aber nicht aus ihnen besteht. Außerhalb des Wissenschaftssystems manifestiert sich die Vorstellung einer vollständig determinierten Welt in einer Ideologie der Machbarkeit, die von einer ökologischen Krise begleitet und darum zunehmend angefochten wird.

Inwiefern kann eine *entwickelte* Künstliche Intelligenz, welche (gegenwärtig nicht erreichte) Zielvorstellungen der anfangs angesprochenen Art realisiert, Auswirkungen auf die konkurrierenden Menschenbilder haben? Es könnte sich herausstellen, daß es trotz aller Anstrengungen unmöglich ist, eine bewußtseinsfähige Maschine zu bauen. Sollte sich dabei zeigen, daß die Unmöglichkeit nicht auf einem Komplexitätsgefälle zwischen hirnorganischen Vorgängen und ihren Simulationen beruht, sondern prinzipieller Natur ist, bedeutete das das Ende des naturwissenschaftlichen Instrumentalismus und damit das Ende einer geistesgeschichtlichen Epoche.

Außerwissenschaftliche, „vulgäre“ Menschenbilder sind derzeit geprägt von einem anthropologischen Pessimismus, der weit über eine Fin-de-siècle-Stimmung hinausgeht. In zunehmendem Maße begreifen sich die Menschen in den hochentwickelten Industrieländern als Wesen, die durch ihre technologischen Fähigkeiten ihre ökologischen und sozialen Voraussetzungen untergraben. Her-

kömmliche Technologien werden zunehmend als Beschleuniger ökologischer und sozialer Entropie wahrgenommen. Ein solches pessimistisches Menschenbild könnte korrigiert werden, sofern sich herausstellt, daß eine entwickelte KI einen Beitrag zur Lösung globaler Probleme leisten kann, etwa indem sie die Ressourcenverteilung optimiert oder durch neuartige Informationssysteme den Verkehrskollaps verhindert.

Künstliche Intelligenz stellt eine besondere Herausforderung für die Zukunftsgestaltung dar: Auch wenn die Realisierung „entwickelter“ KI noch nicht absehbar ist, soll man sich daher schon jetzt mit ihren möglichen Auswirkungen befassen. Zukunft sollte keine passive Angelegenheit sein, in die man sich treiben läßt, sondern es ist uns aufgegeben, die Zukunft aktiv zu gestalten. Die Richtungen, in die sich die Zukunft durch Künstliche Intelligenz entwickeln könnte, sind nicht zufällig, sondern sie werden geleitet von den mehr oder weniger expliziten Spekulationen der Beteiligten oder ihrer Auftraggeber.

So spekuliert etwa der amerikanische Computerwissenschaftler Hans Moravec darauf, daß – unter der Voraussetzung, daß Computer zehn Billionen Rechenoperationen pro Sekunde schaffen werden – in etwa 50 Jahren menschliche Intelligenz inklusive ihrer intuitiven Fähigkeiten auf einem Rechner erzeugt werden kann. Seine Zukunftsvision beinhaltet drastische Veränderungen der gesellschaftlichen Verhältnisse und schließliche Ablösung des Menschen durch Maschinen: „Die Maschinen können jede Arbeit verrichten. Sie werden die Ingenieure ersetzen. Die Unternehmen werden von Management-Systemen geleitet, die wiederum von Computern kontrolliert werden.[...] [Die Menschen] werden die Verbraucher sein. Ich glaube, auf lange Sicht sind die Menschen für das Funktionieren und die Entwicklung der Gesellschaft irrelevant. [...] Man kann [die Maschinen] aber auch als unsere Produkte sehen, Erzeugnisse, die eines Tages, wenn sie selbständig sind, unsere Kultur erben sollen.“¹

¹ Auszüge aus einem Interview von Helene Conrady mit Hans Moravec (Abdruck in den VDI-Nachrichten vom 6. Dezember 1991).

Die technische Realisierbarkeit einmal beiseite gelassen, steht hier vor allem die Wünschbarkeit solcher Entwicklungen zur Debatte, die eine wohl denkbar drastische Änderung heutiger Vorstellungen vom Menschen bedeuten würde. Auch wenn die Machbarkeit von Moravecs Spekulationen gegenwärtig kaum ernsthaft wissenschaftliche Diskussion findet, darf doch ihr Einfluß auf mögliche Zukunftsentwicklungen nicht unterschätzt werden. So haben Beispiele aus jüngerer Zeit – zu denken ist etwa an SDI und Robotik – gezeigt, daß Visionen des bisher kaum Vorstellbaren Leitlinie für die Mobilisierung von Forschungsprogrammen und -investitionen sein können, deren Umsetzung auch bei nur teilweisem Erreichen der Ziele kaum zurücknehmbare Auswirkungen nach sich zieht.

Die Beiträge dieses Abschnitts entstanden innerhalb eines Diskurses unter Teilnehmern unterschiedlicher Denktraditionen und Arbeitsrichtungen, der um die möglichen Auswirkungen einer – hypothetisch betrachteten – entwickelten KI geführt wurde. Es konnte erwartungsgemäß nicht gelingen, zu einer auch nur ansatzweise zufriedenstellenden Beantwortung der im Verlauf des Diskurses auftretenden Fragen zu gelangen. Daher liegt der Wert des Erreichten eher in der Entwicklung von Kategoriengefügen, vor deren Hintergrund eine Behandlung der durchweg als drängend empfundenen Zukunftsfragen in reflektierter Weise möglich ist, als in deren letztgültiger Beantwortung – was ein wohl illusorisches Ziel darstellen würde.

Im folgenden Abschnitt setzt sich A. Kremeier mit der Rolle der Informationstechnik im gesellschaftlichen System auseinander und stellt damit den allgemeinen Hintergrund für gesellschaftliche Wechselwirkungen der KI bereit. Der daran anschließende Gruppenbeitrag der Teilnehmer des Zukunftsdiskurses betrachtet mögliche Auswirkungen einer entwickelten KI entlang gestaffelter Stadien, die hypothetisch Merkmale zukünftiger KI-Systeme beschreiben. Fragen der Verantwortung von KI-Produkten greift P. Schreiber auf, die sich auch mit neuen Rollenverteilungen zwischen Mensch und Maschine auseinandersetzen, und H. Röpke behandelt schließlich potentielle Gefahren von KI-Systemen.