

Nachwort

Maschinen übernehmen immer mehr Aufgaben, die bisher nur Menschen aufgrund ihrer Intelligenz erledigen konnten. Es gibt enorme Fortschritte in der Anwendung von Künstlicher Intelligenz. Anders als bei Wolfgang von Kempelens berühmtem Schachautomaten aus dem 18. Jahrhundert versteckt sich heute kein Mensch mehr in der Maschine. Aber es findet sich auch kein Geist in der Maschine.

Algorithmen finden Problemlösungen, indem sie systematisch oder zufällig viele verschiedene Lösungen ausprobieren. Sprachmodelle nutzen die statistischen Regelmäßigkeiten in Sprache aus, um Texte zu verarbeiten. Lernalgorithmen passen sich an statistische Regelmäßigkeiten in Daten an, brauchen dafür aber riesige Datenmengen. Deshalb tauchen die Schlagwörter maschinelles Lernen und Big Data so häufig zusammen auf. Zwar sind manche dieser Lernalgorithmen von frühen psychologischen Lerntheorien inspiriert, sie lernen jedoch ganz anders als Menschen. In lernender Künstlicher Intelligenz spiegelt sich bislang vor allem die Intelligenz ihrer Entwickler, die die Systeme so lange trainiert haben, bis die Programme eng bestimmte Probleme selbständig lösen konnten. KI-Forscher träumen zwar davon, Programmierer in der Zukunft durch KI-Systeme zu ersetzen, doch die bisher existierenden KI-Systeme zum Programmieren gehen ganz anders vor als menschliche Problemlöser. Weil wir Menschen oft in Worten denken, ist die aktuelle Aufregung um Sprachmodelle durchaus gerechtfertigt. Von Allgemeiner Künstlicher Intelligenz (AKI) sind wir aber noch weit entfernt.

Das Mantra dieses Buches war: Wir sind es, die den Maschinen vor-schnell eine menschenähnliche Intelligenz zuschreiben. Dabei haben wir weder eine genaue Vorstellung davon, was menschliche Intelligenz eigentlich auszeichnet, noch verstehen wir im Detail, wie genau KI funktioniert. Diese Täuschung über die wahre Natur der Maschinen verstellt uns den Blick auf die Zukunft.

Aller Wahrscheinlichkeit nach wird es wohl keine große Revolte geben und KI-Technologie wird nicht grundsätzlich verboten werden (anders als in Frank Herberts Roman *Dune*). Daher ist zu erwarten, dass unser Alltag künftig noch mehr von Computern bestimmt sein wird als er das heute ohnehin schon ist. Wenn wir uns nicht machtlos gegenüber dieser neuen Automatisierungswelle fühlen wollen, müssen wir als Gesellschaft verstehen, was sich hinter KI wirklich verbirgt. Nur dann können wir zukünftige KI-Systeme in unserem Sinne gestalten und regulieren. Einer breiten gesellschaftlichen Diskussion über KI steht aber im Weg, dass es zwei verschiedene Auffassungen von KI gibt, die nicht immer klar auseinandergehalten werden:¹

1. KI als Ingenieursdisziplin
2. KI als Kognitionswissenschaft

Der größte Teil der aktuellen KI-Forschung fällt in die erste Kategorie und ist eine konsequente Fortsetzung früherer Bestrebungen zur Automatisierung. Dabei werden großartige Ingenieursleistungen vollbracht, die echte Probleme lösen. KI-Automaten übernehmen zwar Aufgaben, die sonst Menschen erledigen, der Anspruch ist aber nicht, dass sie das genauso tun wie Menschen. Im Gegenteil, oft ist der Anspruch, dass sie die Aufgaben besser und schneller und deshalb anders als Menschen erledigen. Es wäre besser gewesen, hätte sich statt KI der Begriff »Automatenbau« für diese Ingenieursdisziplin durchgesetzt. Wenn wir von »künstlicher« Intelligenz sprechen, suggeriert das nämlich, dass es darum geht, »natürliche« Intelligenz zu imitieren, was aber in den allermeisten Fällen gar nicht das Ziel ist.

Computer zu nutzen, um natürliche Intelligenz zu verstehen, ist aber durchaus ein Ziel, das die frühe KI-Forschung in der kognitionswissenschaftlichen Tradition verfolgte. Indem man Computermodelle von psychologischen Prozessen entwickelt und das Verhalten dieser

1 Hier übernehme ich grob die Unterscheidung zwischen »AI-as-engineering« und »AI-as-psychology«, die Rooij et al. (2024) machen. In diesem Artikel werden noch weitere Differenzierungen vorgenommen, die ich hier auslasse. Eine ähnliche Einteilung findet sich auch bei Lighthill (1973), der zwischen KI Typ A (für »Advanced Automation«) und Typ C (für »Computer-based Central Nervous System«) unterscheidet. Beiden Arten von KI bescheinigt er, dass sie Fortschritte seit dem Beginn der KI-Forschung gemacht haben. Skeptisch ist er aber beim Typ B (für »Bridge«), der versucht, die beiden anderen Arten zu verbinden und damit viel Verwirrung stiftet.

Modelle mit dem Verhalten von Menschen (oder Tieren) vergleicht, kann man die Grundlagen von Intelligenz untersuchen. Statt von KI spricht man in diesem Fall heute von »kognitiver Modellierung«, die ein Teilgebiet der Kognitionswissenschaft ist. Im Gegensatz zu der Befürchtung, dass die Maschinen schon bald die Intelligenz des Menschen übertrumpfen werden, ist der Fortschritt bei kognitiven Modellen erschreckend langsam. Menschen sind verdammt komplizierte Wesen und Psychologinnen und Psychologen müssen ihre Modelle erst in vielen Experimenten überprüfen, bevor sie ernsthaft behaupten können, dass ein Computerprogramm ein Problem ähnlich wie ein Mensch löst. Dabei sind sie extrem vorsichtig, dass sie den Maschinen nicht leichtfertig menschliche Intelligenz zuschreiben – schon alleine, weil niemand genau versteht, was menschliche Intelligenz eigentlich ist. Der andere Teil der KI-Forschung, der keine kognitiven Modelle entwickelt, ist oftmals mit seinen Zuschreibungen von Intelligenz nicht ganz so vorsichtig.²

Genauso wie Flugzeuge und Vögel beide das physikalische Prinzip des Auftriebs zum Fliegen nutzen, nutzen Computer und Menschen beide die Prinzipien der Informationsverarbeitung, um sich intelligent zu verhalten. Deshalb sind die zwei Arten von KI, der Automatenbau und die kognitive Modellierung, eng miteinander verwandt. Aber sie haben andere Ziele. Manche Flugzeugentwürfe sind zwar vom Vogelflug inspiriert, aber es erscheint lächerlich, dass jemand versuchen könnte, ein Passagierflugzeug zu bauen, das genauso mit den Flügeln schlägt wie ein Vogel. Natürlich könnte man künstliche Federn herstellen und auf die Flügel kleben, die genaue Form des Skelettes nachbauen und künstliche Muskeln entwickeln, die das Skelett bewegen. Möchte man im Detail verstehen, wie Vögel fliegen, dann ist so ein Modell eines Vogels extrem hilfreich. Für Passagierflugzeuge, die viel größer als jeder Vogel sind und schneller fliegen, taugen Vögel allerdings nur bedingt als Vorbild.³

Im Fall von KI spukt trotzdem ständig die Science-Fiction-Vorstellung von menschenähnlichen Robotern durch die Diskussion. Dabei wird ganz selbstverständlich angenommen, dass die Maschinen eine

2 Mitchell (2023) und Odouard & Mitchell (2022) kritisieren das.

3 Die Analogie zwischen Kognition und KI auf der einen Seite und Vogelflug und Flugzeug auf der anderen ist in der Kognitionswissenschaft sehr beliebt und geht auf Marr (1982) zurück.

menschenähnliche Intelligenz entwickeln werden. Ich verstehe nur zu gut, welche Faszination die Idee ausübt, dass wir menschliche Intelligenz im Computer nachbilden können, schließlich habe ich selbst deswegen die kognitive Modellierung zu meinem Beruf gemacht. Ich kann mir gut vorstellen, dass wir theoretisch immer kompliziertere Computermodelle bauen könnten, die das menschliche Denken immer besser in seiner Gesamtheit simulieren. Aber nur weil ich es theoretisch für möglich halte, heißt das nicht, dass ich es praktisch für machbar halte.⁴

Außerdem halte ich es nicht für sinnvoll. Warum sollte das Ziel von KI als Ingenieursdisziplin das Nachahmen von menschlicher Intelligenz sein? Wenn KI-Systeme in manchen Anwendungen Menschen ersetzen sollen, sollten sie die Aufgaben besser erledigen, als der Mensch es tut. Das wird in den meisten Fällen bedeuten, dass sie es anders machen. Statt einer menschenähnlichen KI brauchen wir dann eine menschenunähnliche KI. In anderen Anwendungen sollen KI-Systeme Menschen nicht ersetzen, sondern unterstützen. Dafür müssen sie die Fähigkeiten des Menschen ergänzen, statt sie nachzubilden. Statt einer menschenähnlichen KI brauchen wir in diesen Fällen eine menschenzentrierte KI.⁵ Wie Menschen und KI-Systeme in Zukunft zusammenarbeiten werden, ist dabei noch nicht ausgemacht. Wir müssen uns entscheiden, welche Tätigkeiten wir an KI-Systeme abgeben und bei welchen sie uns lediglich unterstützen sollen. Dabei müssen wir in jedem Einzelfall sorgfältig darauf achten, ob wir die mit einer Aufgabe einhergehende Verantwortung wirklich an eine Maschine abgeben können.

KI ist keine Naturgewalt, die plötzlich und unerwartet über uns hereinbricht. Natürlich gibt es wirtschaftliche und geopolitische Interessen, die wie bei vergangenen Automatisierungswellen die Entwicklung von neuen Technologien antreiben und die international nur schwer zu kontrollieren sind. Aber wir sind als Gesellschaft diesen Entwicklungen nur dann machtlos ausgeliefert, solange wir sie nicht verstehen und nicht politisch mitgestalten. Mit der KI-Verordnung der Europäischen Union ist ein Anfang gemacht.

4 Siehe nochmal Rooij et al. (2024), die auch darauf hinweisen, dass aus der theoretischen Möglichkeit nicht die praktische Machbarkeit folgt. Sie haben sogar gute theoretische Argumente, warum es praktisch nicht geht.

5 Siehe Shneiderman (2022) für eine Einführung in menschenzentrierte KI.

Diskussionen über KI drehen sich aber leider immer noch zu häufig darum, ob KI-Systeme denken können, inwiefern Menschen nur biologische Maschinen sind oder superintelligente Killerroboter die Welt-herrschaft übernehmen werden. Das sind spannende philosophische Fragen. Es ist auch aufregend, sich die ferne Zukunft vorzustellen. Wir sollten uns aber deshalb nicht von den eigentlichen Fragen ablenken lassen, die sich heute stellen: Welche gesellschaftlichen Auswirkungen haben Digitalisierung, Big Data und KI? Und wie genau sollen Menschen und Maschinen zusammenarbeiten, damit wir die Kontrolle und die Verantwortung behalten? Wenn wir diese Fragen beantworten wollen, dürfen wir uns nicht über die Fähigkeiten von KI täuschen und künstliche mit menschlicher Intelligenz verwechseln.

