

# Einsatz von KI in der Medizin unter philosophischen Aspekten

B. Heinrichs

## Künstliche Intelligenz als Problem der Philosophie

Schaut man auf die Entwicklung der künstlichen Intelligenz (KI), dann stellt man schnell fest, dass von Beginn an – das heißt seit Mitte des 20. Jahrhunderts – philosophische Fragestellungen eine wichtige Rolle gespielt haben. In seinem einflussreichen Beitrag „Computing Machinery and Intelligence“ [1] fragt der KI-Pionier Alan Turing, ob Maschinen denken können, um diese Frage dann allerdings aufgrund mangelnder Klarheit direkt wieder zu verwerfen. Anschließend konzipiert er das, was heute als Turing-Test bekannt ist: ein Verfahren, das auf die Unterscheidbarkeit von menschlichem und maschinellem Kommunikationsverhalten abstellt. Die philosophische Auseinandersetzung mit der Möglichkeit künstlicher Intelligenz reicht indes wesentlich weiter zurück, nämlich mindestens bis ins 17. Jahrhundert. Bereits in dieser Zeit setzt Thomas Hobbes Denken mit Rechnen gleich und machte damit einen wichtigen konzeptionellen Schritt in Richtung KI [2]. Es folgen bald die ersten mechanischen Rechenmaschinen, gebaut beispielsweise von Blaise Pascal (1642) und Gottfried Wilhelm Leibniz (1673). Weitere wichtige Schritte auf dem Weg zur künstlichen Intelligenz im 19. Jahrhundert bilden Arbeiten zur Logik, insbesondere von George Boole und Gottlob Frege. Aus heutiger Sicht mündet all dies schließlich in Turings bahnbrechender Schrift [3], in der er eine universelle Rechenmaschine beschreibt – heute spricht man von Turing-Maschinen –, die die theoretische Grundlage für moderne Computer darstellt. Noch einmal knapp



Was unterscheidet den Menschen von anderen Wesen und Artefakten?

20 Jahre später kommt es zur Wortschöpfung „artificial intelligence“ durch John McCarthy et al. [4].

## Eine Aufforderung zur menschlichen Selbstvergewisserung

Mit der Mechanisierung des Rechnens und der daraus resultierenden Möglichkeit von „Denkmaschinen“ stellt sich eine sehr grundlegende Frage, nämlich

**„KI scheint nicht irgendeine neue Technologie zu sein, sondern stellt vielmehr eine Disruption dar, die eine erneute grundsätzliche Selbstvergewisserung des Menschen erforderlich macht.“**

was den Menschen von anderen Wesen und Artefakten unterscheidet. Die Bestimmung, die Aristoteles (384 – 322 v. Chr.) in seiner *Politik* gibt – der Mensch als *animal rationale* –, hat das

abendländische Denken maßgeblich geprägt. Sie verliert aber erheblich an Überzeugungskraft, wenn Denken nichts anderes ist als Rechnen, das auch durch Maschinen ausgeführt werden kann. Luciano Floridi, einer der derzeit prominentesten Philosophen, die sich mit KI beschäftigen, hat daher nicht Unrecht, wenn er von einer „vierten Revolution“ spricht, der sich die Menschheit nun, das heißt nach Kopernikus, Darwin und Freud, gegenüber sieht [5]. KI scheint nicht irgendeine neue Technologie zu sein, sondern stellt vielmehr eine Disruption dar, die eine erneute grundsätzliche Selbstvergewisserung des Menschen erforderlich macht.

## Was ist Intelligenz?

Ein möglicher Ansatzpunkt für eine solche Selbstvergewisserung ist die Frage, was unter Intelligenz zu verstehen ist. Die einschlägige Fachliteratur zeigt, dass es sich um einen hochkontroversen Begriff in Psychologie, Kognitionswissenschaften, Informatik und Philosophie handelt. Diese Disziplinen sind dabei jeweils von unterschiedlichen Erkenntnisinteressen bestimmt, was

sich in unterschiedlichen Definitionsversuchen niederschlägt. Dimitri Coelho Mollo [6] hat unlängst trotzdem versucht, eine Begriffsbestimmung vorzunehmen, die für viele Disziplinen konsensfähig sein soll. Im Kern bestimmt er Intelligenz als generelles, flexibles, adaptives und zielgerichtetes Verhalten. Offenbar sind nicht nur Menschen, sondern auch höhere Tiere und ebenso einige Artefakte zu solchem Verhalten fähig. Intelligenz ist demnach kein Alleinstellungsmerkmal des Menschen. Man sollte aus diesem Befund jedoch nicht voreilig folgern, dass es keine wichtigen Unterschiede zwischen Menschen und KI mehr gibt. Abgesehen davon, dass die Generalität menschlicher Intelligenz nach wie vor unerreicht ist, es also bislang keine *artificial general intelligence* (AGI) gibt, fehlt in Coelho Mollo's Definition mindestens eine Bestimmung, die aufs Engste mit menschlicher Intelligenz verknüpft ist: Bewusstsein. Es ist einstweilen völlig unklar, ob sich Bewusstsein künstlich erzeugen lässt. Das liegt vor allem daran, dass wir das Phänomen Bewusstsein bislang nur unzureichend verstehen. David Chalmers [7] hat die mittlerweile verbreitete Wendung vom „hard problem“ geprägt, vor das uns Bewusstsein stellt.

Noch ein weiteres Phänomen sperrt sich hartnäckig einem umfassenden Verständnis: Intentionalität. Damit bezeichnet man in der Philosophie den Umstand, dass unser Denken offenbar stets (oder doch zumeist) von etwas handelt oder sich auf etwas bezieht. Franz Brentano vertrat im letzten Viertel des 19. Jahrhundert die These, dass Intentionalität eine charakteristische Eigenschaft des Mentalen sei und prägte damit die philosophische Diskussion bis heute. Neben vielen anderen beteiligt John Searle sich intensiv an dieser Diskussion und arbeitet die Eigenarten von Intentionalität heraus

[8]. Aktuelle Ansätze wie der von Coelho Mollo entkoppeln Intelligenz von Bewusstsein und Intentionalität. Künstliche Intelligenz mag demnach zwar intelligent sein, in anderen wesentlichen Bestimmungen unterscheidet sie sich aber weiterhin von menschlichem Denken. Auf Searle geht ein weiterer wichtiger Beitrag in der Diskussion um KI zurück. In einem Aufsatz aus dem Jahr 1980 [9] führt er die Unterscheidung von „starker“ und „schwacher“ KI ein, die mittlerweile als etabliert gelten kann. Mithilfe eines Gedankenexperiments („Chinese Room“) versucht Searle zu zeigen, dass KI das Verständnis von Informationen lediglich simuliere, statt über ein eigenständiges Verständnis zu verfügen, und insofern lediglich „schwach“ sei. Viele haben dem in der Folge widersprochen. Unterdessen muss man die Frage wohl als offen bezeichnen, wie triftig Searles Argument tatsächlich ist. Klar ist jedenfalls, dass aktuelle KI nicht „stark“ im Sinne von Searle ist und es damit weiterhin erhebliche Unterschiede zwischen Menschen und Maschinen gibt. Dessen ungeachtet hat die KI-Forschung sehr große Fortschritte gemacht, und es gibt inzwischen eine Fülle von KI-Anwendungen, die in den unterschiedlichsten Lebensbereichen kaum noch wegzudenken sind.

### Künstliche Intelligenz in der Medizin

Wie in anderen Bereichen, so hat KI in den vergangenen Jahren auch in der Medizin Einzug gehalten. Vor allem die Radiologie ist mit ihren Bilddaten prädestiniert für KI-basierte Technologien. Der KI-Entwickler Geoffrey Hinton prophezeite entsprechend bereits im Jahr 2016, Radiologinnen und Radiologen seien vom Aussterben bedroht. Diese Prophezeiung hat sich nicht bewahrheitet. Richtig ist indes, dass KI-Systeme heute schon teilweise bessere Ergebnisse liefern als Fachärztinnen

und Fachärzte. Im Jahr 2020 wird etwa berichtet, dass eine KI-Anwendung Radiologinnen und Radiologen beim Mammographie-Screening übertrifft [10]. Es ist absehbar, dass sich dieser Trend fortsetzt und auch Teilgebiete der Medizin erfasst, in denen man es zunächst womöglich weniger erwarten würde, weil die medizinischen Daten in Textform vorliegen. Die rasante Entwicklung im Bereich der *large language models* (LLMs) zeigt, dass KI auch bei der Verarbeitung von Text inzwischen enorm leistungsfähig ist. Entsprechend beginnen LLMs nun zunehmend im klinischen Kontext eingesetzt zu werden [11]. Ihr Einsatz beschränkt sich nicht auf die Unterstützung von Diagnose- und Therapiefindung, sondern betrifft auch alltägliche Arbeitsabläufe. Firmen sind längst dabei, Lösungen anzubieten, die beispielsweise die Erstellung von Arztbriefen aus Patientengesprächen heraus automatisieren. Im Idealfall wird dies dazu führen, dass Ärztinnen und Ärzte wieder mehr Zeit haben, sich im direkten Gespräch mit ihren Patientinnen und Patienten zu befassen.

### Plädoyer für eine „deflationäre“ Sichtweise

Auch wenn es eine interessante und berechtigte philosophische Frage bleibt, ob Maschinen denken können und was das eigentlich genau bedeutet, sollten wir KI nicht hypostasieren. Es gibt nicht „die“ KI. Vielmehr haben wir es mit einer Vielzahl sehr komplexer Werkzeuge zu tun, die wir zur Bewältigung von Problemen und Herausforderungen nutzen können. Darin liegen große Chancen. Wie bei allen Technologien, so gibt es bei der Nutzung von KI jedoch auch Risiken. Diese liegen aber nicht in Weltuntergangsszenarien, wie sie von einigen in düsteren Farben gemalt werden [12]. Sie liegen eher darin, dass wir Werkzeuge leichtfertig einsetzen, ohne zuvor mögliche Auswirkungen sorgfältig bedacht zu haben. Jocelyn Maclure

[13] mahnt daher zurecht eine „deflationären Sichtweise“ an. Nimmt man eine solche ein, dann erkennt man, dass zu realistischen Auswirkungen etwa Fehler und Verzerrungen gehören, die sich aus einseitigen Trainingsdaten ergeben. Dieses als *algorithmic bias* bezeichnete Problem kann dazu führen, dass eine medizinische KI-Anwendung für eine Patientengruppe gut funktioniert, für eine andere, die in den Trainingsdaten unterrepräsentiert war, aber nicht. Zu diesen Auswirkungen gehört auch, dass eine übermäßige Automatisierung die Erkennung von Fehlern unmöglich macht. Dieses Problem ist sicherlich nicht neu, bei KI aber größer als bei herkömmlichen Technologien, weil zumindest einige Formen von KI epistemisch opak sind. Mittlerweile wird unter dem Titel *explainable AI* speziell daran geforscht, dieses Problem besser in den Griff zu bekommen. Mit der Forderung nach Erklärbarkeit sollte allerdings nicht über das Ziel hinausgeschossen werden: Womöglich müssen nicht alle KI-Anwendungen transparent sein. Auch die Wirkweise von vielen klassischen Pharmazeutika ist nicht vollends klar. Wichtig ist, dass sich Risiken abschätzen lassen. Anders als bei Pharmazeutika fehlt im Bereich von KI-Anwendungen bislang aber eine gewissenhafte Prüfung in Form von klinischen Studien [14]. Ohne solche Prüfungen bleibt das tatsächliche Risikoprofil jedoch unklar. Statt also in unkritische Heilserwartungen oder in hyper-skeptische Untergangsstimmung zu verfallen, sollten wir Nutzen und Risiken von KI-Anwendungen in jedem Einzelfall nüchtern abklären. Die Regulierung, die der neue AI Act der Europäischen Kommission vorsieht, versucht dafür einen gesetzlichen Rahmen zu schaffen.

## Fazit

Trotz dieses Plädoyers für eine unaufgeregte, einzelfallbasierte Abwägung von Nutzen und Risiken, wie wir sie aus anderen Bereichen kennen, ist es wichtig, sich immer wieder klarzumachen, was KI – zumindest nach aktuellem Stand der Wissenschaft – ist und was nicht.

**„Wir sind es daher, die darüber entscheiden müssen, wie beziehungsweise auf welche Weise wir die neue und reichhaltige Palette von Werkzeugen nutzen wollen, die uns die Erfolge in der KI-Forschung bescheren.“**

Legt man ein Verständnis von Intelligenz zugrunde, wie Coelho Mollo es vorschlägt, dann sind einige KI-Anwendungen intelligent. Solche Anwendungen sind dazu fähig, flexibel und adaptiv Problemlösungen zu entwickeln. Neuere *foundational models* sind zudem zumindest ansatzweise generell. KI-Anwendungen verfügen hingegen nicht über Bewusstsein oder Intentionalität – Eigenschaften, die beim Menschen eng mit Denken und Problemlösungsverhalten verknüpft sind und die wir daher womöglich voreilig mit Intelligenz in eins setzen. KI-Anwendungen sind vor allem auch keine moralischen Akteure. Menschen hingegen sind Wesen, die für ihr Verhalten Verantwortung übernehmen können und die wir für ihr Verhalten zur Verantwortung ziehen können. Das

unterscheidet uns auf absehbare Zeit von Maschinen (und übrigens auch von Tieren). Wir sind es daher, die darüber entscheiden müssen, wie beziehungsweise auf welche Weise wir die neue und reichhaltige Palette von Werkzeugen nutzen wollen, die uns die Erfolge in der KI-Forschung bescheren. Gerade in der Medizin, in der sich der Mensch als ebenso vielschichtiges wie vulnerables Wesen zeigt, müssen wir den Umgang mit technischen Hilfsmitteln besonders aufmerksam bedenken. Statt in skeptische oder enthusiastische Extreme zu verfallen, sollten wir nüchtern verfahren und die Voraussetzungen für einen verantwortungsvollen Einsatz von KI schaffen. Vor allem aber sollten wir uns davor hüten, Intelligenz mit den vielfältigen Fähigkeiten des Menschen zu verwechseln. Auch intelligente Maschinen bleiben Maschinen, und Menschen sind sehr viel mehr als bloß intelligente Lebewesen. ■

Literatur unter [www.slaek.de](http://www.slaek.de) → Über Uns →  
Presse → Ärzteblatt



Prof. Dr. Bert Heinrichs  
Institut für Wissenschaft und Ethik (IWE)  
Rheinische Friedrich-Wilhelms-Universität Bonn  
[www.iwe.uni-bonn.de](http://www.iwe.uni-bonn.de)

Arbeitsgruppe „Neuroethik und Ethik der KI“  
Institut für Neurowissenschaften und Medizin:  
Gehirn und Verhalten (INM-7)  
Forschungszentrum Jülich  
[www.fz-juelich.de](http://www.fz-juelich.de)