

« Der Mensch wird stolz sein dürfen,
wenn die Computer sich eines Tages als
seine Nachkommen bezeichnen.»

Hans Moravec

Künstliche Intelligenz

Was «Künstliche Intelligenz» bedeutet, ist bis heute genau so wenig sauber definiert wie der Begriff der «Intelligenz». Als allgemein anerkannt gelten jedoch das Geburtsjahr 1956, das Dartmouth-College in Hanover, New Hampshire, als Ort der Geburt und John McCarthy als Vater des Begriffe. Und sicher ist, dass mit dem Forschungsgebiet der Künstlichen Intelligenz ein neues Zeitalter seinen Anfang nahm, vergleichbar mit dem der Industriellen Revolution. War es bei letzterer um den Ersatz der menschlichen Arbeitskraft durch Maschinen gegangen, so ging es bei der neuen Revolution um den Ersatz des menschlichen Denkens durch «Elektronengehirne» wie die Computer anfangs der 1950er Jahre genannt wurden.

Das Forschungsgebiet der «Künstliche Intelligenz» wurde von zwei Fragestellungen vorangetrieben. Die eine hatte zum Ziel, mittels des Computers Einblicke in das menschliche Denken zu gewinnen, die andere wollte dem Computer Fähigkeiten beibringen, die dem menschlichen Denken möglichst nahe kamen. Oder anders gesagt, eine Richtung der KI beschäftigt sich mit der Funktionsweise des Gehirns, die andere mit menschenähnlichem Verhalten von Maschinen.

Wie jede Richtung innerhalb der Informatik, hat auch die Forschung auf dem Gebiet der Künstlichen Intelligenz ihre Vor- oder Frühgeschichte. Die Vision, dass Maschinen Leistungen erbringen könnten, welche von denen des Menschen nicht mehr zu unterscheiden sind, ist schon alt. Bereits die Alten Griechen kannten den Mythos von Hephaistos, dem kunstfertigen Schmiedegott, welcher selbstbewegende Dreifüsse (αὐτοματοι) und künstliche Dienerinnen aus Gold erschuf. Ebenfalls weit bekannt ist die Gestalt der Pandora, welche derselbe Hephaistos auf Geheiß von Zeus menschengleich erschuf, allerdings mit dem Zweck, die Menschen für ihren Übermut zu strafen.

Im 18. Jahrhundert bauten in Frankreich, Österreich und der Schweiz begabte Handwerker sogenannte Androiden (künstliche Menschen), welche Flöte oder Klavier spielten oder Texte schreiben konnten. Der deutsche Schriftsteller Jean Paul ging sogar soweit, das 18. Jahrhundert als das Zeitalter der Automaten zu bezeichnen.

Im 19. Jahrhundert dachte der englische Mathematiker Charles Babbage, welcher von einem Automaten in Gestalt einer silbernen Dame fasziniert war, während seiner Arbeit an der Analytical Engine darüber nach, dass seine Maschine «could do anything but compose country dances». Dass er mit dieser Vermutung nicht falsch lag, konnte aber zu seinen Lebzeiten nicht bewiesen werden.

Eine Maschine, die immerhin schon Schach spielen konnte – zumindest einfache Endspiele mit drei Figuren – gab es dank dem spanischen Ingenieur Leonardo Torres y Quevedo kurz vor dem Ersten Weltkrieg. Das königliche Spiel als Inbegriff strategischen Denkens hat in der Folge immer wieder Mathematiker und Ingenieure dazu veranlasst, darüber nachzudenken, ob eine Maschine in der Lage sein könnte, einen menschlichen Großmeister zu besiegen. So ist es von Konrad Zuse und Claude Shannon überliefert, und der Vater der Kybernetik, Norbert Wiener, äußert sich in seinem Standardwerk «Kybernetik» ebenfalls ausführlich über das Schachspiel.

Mit Shannon, Turing und Wiener sind die drei wichtigsten Namen aus der Frühzeit der KI-Forschung genannt. Norbert Wiener (1894–1964) war habilitierter Mathematiker, sprach viele Sprachen, rauchte leidenschaftlich Zigarren und interessierte sich vor allem für die Randgebiete der wissenschaftlichen Disziplinen. Schon im Ersten Weltkrieg hatte der kurzsichtige, militäruntaugliche Wiener sich in Maryland mit der Berechnung von Flugbahntabellen für die Armee befasst. Während des Zweiten Weltkrieges wandte er sich zusammen mit Julian Bigelow (1913–2003) der Steuerung von Flakkanonen zu, einem der anspruchsvollsten Probleme der Feuerleittechnik. Dabei stieß er auf das Phänomen der Rückkopplung, das er für sein geplantes Abwehrsystem verwenden wollte, um die Reaktion der Piloten auf das Abwehrfeuer vorwegzunehmen. Das System wurde nicht verwirklicht, aber der Artikel «Behavior, Purpose and Teleology» wurde zum Ausgangspunkt einer neuen Forschungsrichtung: der Kybernetik. Nach dem Krieg veröffentlichte Wiener dann mit «Cybernetics. Control and Communication in the Animal and the Machine», sein Standardwerk zur Systemtheorie. Darin untersucht Wiener die Begriffe «Information» und «Kommunikation». Ein zentrales Kapitel ist dem Begriff der Rückkopplung gewidmet und ein weiteres trägt den Titel «Rechenmaschinen und das Nervensystem». Wiener vertritt darin zwei für die Entwicklung der Künstlichen Intelligenz wesentliche Thesen: Zum einen behauptet er, dass Lebewesen durch Rückkopplung lernten und zum andern stellt er fest, dass das menschliche Nervensystem im Grunde auf binären Schaltungen basiere. Die Idee, der binären Schaltung des menschlichen Nervensystems war schon 1943 vom Neurophysiologen Warren McCulloch (1898–1969) aufgegriffen worden, welcher nach dem Krieg regelmäßige Konferenzen im Beekman Hotel in New York veranstaltete, an denen versucht wurde, die Sozialwissenschaften auf ein mathematisches Fundament zu stellen. In einem, zusammen mit seinem Schüler Walter Pitts (1923–1963) veröffentlichten Aufsatz «A logical calculus of the ideas in nervous activity» verglich er die Funktionsweise des Gehirns mit der eines Computers auf der Basis von Relais'. Letztere nannte er «Neurons». Und er sprach stolz von der Geburt einer «Kommunikationsphysik».

Diese kybernetische Sichtweise der Künstlichen Intelligenz wird in der Literatur gemeinhin als «bottom up»-Ansatz bezeichnet. Seine Vertreter, zu denen auch Alan Turing, Claude Shannon, Frank Rosenblatt und der junge Marvin Minsky zu zählen sind, sehen Künstliche Intelligenz als elektronische Kopie von Neuronennetzen.

Daneben entstand eine gegenläufige Richtung, «top down»-Ansatz genannt, der der ältere Marvin Minsky (1927–2016) und Seymour Papert (1928–) oder Hans Moravec (1948–) und Ray Kurzweil (1948–) zuzurechnen sind. Diese Richtung betrachtet Künstliche Intelligenz als Nachvollzug des Gehirns durch Software. Die ersten Vertreter dieser Richtung waren für das Highlight der Dartmouth-Tagung verantwortlich, der Vorstellung des «Logic Theorist». Dieses Computerprogramm, das die Professoren Herbert A. Simon (1916–2001) und Allen Newell (1927–1992) zwischen 1952 und 1956 entwickelten, war in der Lage, formallogisch Theoreme zu beweisen. Damit war gezeigt, dass Computer Probleme lösen konnten, die Intelligenz und Vorstellungsvermögen erforderten. Ja, es hatte sogar Koryphaen wie Russel und Whitehead übertroffen, indem es in einem Fall einen kürzeren Beweis fand als die menschlichen Mathematiker.

Begreiflicherweise war nach dieser Präsentation die Begeisterung groß, und der Künstlichen Intelligenz wurde eine glänzende Zukunft vorausgesagt. Ein Jahr nach Dartmouth prophezeite Herbert Simon zum Beispiel, dass in zehn Jahren ein Computer Schachweltmeister sein werde. Und Edward Fredkin (1934–) postulierte die Künstliche Intelligenz gar als nächste Stufe der Evolution. Aber so schnell ging es dann doch nicht. Zwar flossen in der Folge große Mengen Geldes in die drei wichtigsten Zentren der KI-Forschung in Stanford, an der Carnegie Mellon Universität in Pittsburgh und ans MIT in Cambridge MA, aber die Resultate ließen auf sich warten. Ja es kam sogar zu Bremsmanövern innerhalb der

Forschercommunity. Ein interessantes Beispiel dafür ist der 1969 erschienene Artikel «Perceptrons» von Marvin Minsky und Seymour Papert, in dem sie zeigen wollten, dass die von Frank Rosenblatt (1928–1971) postulierten Möglichkeiten des Perzeptrons utopisch und mathematisch nicht haltbar seien. Rosenblatts 1960 präsentiertes Perzeptron bestand aus einem IBM 704, Photozellen und einem gigantischen Programm, welches nach zweijähriger Konstruktionsphase in der Lage war, Buchstaben zu erkennen. Dies allerdings nur, wenn die Buchstaben im gleichen Font und der gleichen Größe auftauchten. Nach dem Artikel von Minsky und Papert begannen die Mittel für die *bottom up* Richtung zu versiegen, und *top down* hatte Oberwasser. Allerdings stellte sich 1985 heraus, dass Minsky mit seiner Kritik am Perzeptron teilweise falsch gelegen war.

Dann gab es auch Kritiker von außerhalb der Informatik, wie etwa den Philosophen Hubert Dreyfus (1929–2017) von der Universität Berkeley, welcher die Künstliche Intelligenz 1965 mit der Alchemie verglich. Seine Kritik an einem Dame-Programm verhalf der Künstlichen Intelligenz allerdings zu einem Prestige-Erfolg; denn Dreyfus verlor einen Match gegen das Schachprogramm MacHack von Richard Greenblatt (1944–). MacHack hatte immerhin schon die Spielstärke von 1400 Elo-Punkten. Der damals amtierende Schachweltmeister, Tigran Petrosjan, zum Vergleich, wies deren 2362 auf. Dreyfus stand zu seiner Niederlage. Er räumte später auch ein, dass er der Arbeit von Newell und Simon nicht gerecht geworden sei, aber er blieb bei seiner Meinung, dass eine kontextfreie Psychologie, wie sie Vertreter der Künstlichen Intelligenz postulierten, ein Widerspruch in sich sei, und die KI-Forschung an der formalen Darstellung des menschlichen Alltagswissens und am Verstehen natürlicher Sprache scheitern werde.

Ein weiterer Einwand kam vom Philosophieprofessor John Searle (1932–) aus Berkeley. Er stellte 1980 die Frage, ob es möglich sei, dass ein Computer ein kognitives Bewusstsein entwickeln könne, indem auf ihm ein bestimmtes Programm laufe, wie das die Vertreter der starken KI behaupteten. Zur Beantwortung dieser Frage machte er ein Gedankenexperiment. Sein sogenanntes Chinese-Room-Argument lehnt sich an den Aufbau des Turing-Tests an: Wenn eine Person, dank einer Anleitung, auf Chinesisch gestellte Fragen chinesische Antworten generieren kann, dann versteht sie deswegen noch kein Chinesisch. Sie verhält sich gleich wie ein Computer, welcher ein Programm ausführt, das chinesische Sätze generiert: Der Computer versteht deswegen kein Chinesisch. Damit hat Searle die Theorie der starken KI mit den Mittel der Logik widerlegt.

Ein weiteres Einsatzgebiet der Künstlichen Intelligenz waren ab Mitte der 1960er Jahre die sogenannten Expertensysteme. Die Idee hinter solchen Systemen war, Expertenwissen in den Speicher eines Computers abzufüllen und diesen Wissensbestand dann zielgerichtet mittels heuristischer Suche und Wenn-dann-Regeln durchsuchen zu können, um damit Probleme zu lösen. Das erste Beispiel eines solchen Systems hieß DENDRAL und wurde von Edward Feigenbaum (1936–) in Stanford konzipiert. Es war in der Lage, praktisch gleich gut vorausszusagen, welche Moleküle sich mittels bestimmter Atome herstellen lassen, wie ein ausgebildeter Chemiker. In der Folge wurden unter anderem in der Medizin, der Ölförderung, der Schaltkreisanalyse ähnliche Systeme entwickelt und auch kommerziell nutzbar gemacht.

Computer können aber nicht nur eingesetzt werden, um Wissen zu archivieren und Entscheide zu fällen. Sie können auch der Kommunikation dienen. Die Sprache ist ein zentrales Element des Wissenserwerbs und der Wissensrepräsentation. Deshalb begann sich die KI-Forschung auch für die Frage zu interessieren, was denn eigentlich Sprache ausmache. Im Zusammenhang mit der Sprache muss neben Noam Chomsky schließlich der Name desjenigen Mannes genannt werden, welcher den Begriff «Künstliche Intelligenz» prägte: John McCarthy (1927–2011). Der Sohn eines Arbeiterführers gehörte

zu den Organisatoren der Dartmouth-Konferenz und wirkte ab 1957 am MIT, wo er sich mit dem Problem des Time-Sharing befasste und die Programmiersprache LISP entwickelte, deren Charakteristikum es war, dass Programme wie Daten verarbeitet werden konnten. 1963 wechselte McCarthy von Boston nach Stanford, wo er maßgeblich am Aufbau des AILabs beteiligt war. 1971 erhielt McCarthy für seine Verdienste den Turing Award.

Die Möglichkeit des Time-Sharing und von LISP inspirierten dazu, Programme zu entwickeln, die mit dem Computer einen Dialog führen konnten. Das wohl bekannteste war ELIZA von Joseph Weizenbaum (1923–2008). ELIZA simulierte einen Gesprächstherapeuten und kann als frühe Implementierung des Turing-Tests betrachtet werden. Allerdings zeigte sich gerade an diesem Projekt auch eine fragwürdige Seite des Einsatzes von Computern; denn den Computern fehlt das Wissen um den Kontext, wie neben Hubert Dreyfus auch der Computerpionier Joseph Weizenbaum selbstkritisch erkannte. Dieses Wissen um den Kontext, hat einer der bedeutendsten Science Fiction Autoren, Stanislaw Lem (1921–2006), 1964 in seinem Robotermärchen «Der Freund des Automatthias» veranschaulicht. Dort rät ein intelligentes System einem Roboter nach einem Schiffbruch zum Selbstmord, weil das in Anbetracht seiner aussichtslosen Lage, der schmerzloseste Tod sei. Das System weiß aber nicht, dass der Funker des Schiffes vor dem Untergang noch einen Hilferuf absetzen konnte, welcher eine Suchaktion nach Überlebenden auslöste. Würde Automatthias der Vernunft seines Ratgebers Rimohr folgen, wäre er tot, als die Retter auftauchen.

Ein weiterer Kritikpunkt an ELIZA war, dass einige glaubten, ELIZA sei der Durchbruch zum Verständnis natürlicher Sprachen, andere begannen, sich von Computer verstanden zu fühlen und eine Gefühlsbeziehung zu ihm aufzubauen, obwohl es sich bei ihm um eine geistlose Maschine handelte. Und schließlich überraschte Weizenbaum auch die Reaktion der Psychologen, welche in seinem Programm eine Möglichkeit sahen, den Mangel an Therapeuten zu beheben und sich allen Ernstes an die Verbesserung von ELIZA machten. Dies, seine zunehmenden moralischen Bedenken gegen die Delegation von Verantwortung an Maschinen und der wachsende Einfluss des Militärs auf die KI-Forschung und die Informatik ließen aus Weizenbaum einen der führenden Kritiker der Künstlichen Intelligenz werden.

Neben die Arbeit an humanoiden Robotern und an Expertensystemen traten in der jüngsten Vergangenheit neue Forschungsgebiete wie beispielsweise *human enhancement* und *Synthetische Biologie*. Diese haben die «Verbesserung» oder Neuschöpfung biologischer Systeme mittels Informatik zum Ziel. Ray Kurzweil (1948–) hat 2001 prophezeit, dass «wir im Jahr 2030 über Nanoroboter verfügen, die uns ermöglichen, unser Gehirn zu verstehen, nachzubauen und schließlich zu übertreffen». Der Weg in eine posthumane Zukunft ist für ihn und seine Apologeten bereits vorgezeichnet.

Literatur

Conlan, Roberta; Hrsg. (1988): Künstliche Intelligenz. Amsterdam: Time-Life Books

Diemers, Daniel (2002): Die virtuelle Triade. Bern: Haupt

Dietz, Peter (2003): Menschengleiche Maschinen. Berlin: Bühler & Heckel

Dreyfus, Huber L. (1985): Was Computer nicht können. Die Grenzen künstlicher Intelligenz. Königstein/Ts.: Athenäum

Gleick, James (2011): Die Information. München: Redline

- Görz, Günter/Rollinger, Claus-Rainer/Schneeberger, Josef (2003): Handbuch der Künstlichen Intelligenz. München: Oldenbourg
- Lem, Stanislaw (1973): Robotermärchen. Frankfurt am Main: Suhrkamp
- Minsky, Marvin (1990): Mentopolis. Stuttgart: Klett
- Moravec, Hans (1990): Mind Children. Hamburg: Hoffmann und Campe,
- O'Reagan, Gerard (2008): A Brief History of Computing. London: Springer
- Turkle, Sherry (1984): Die Wundermaschine. Reinbek bei Hamburg: Rowohlt
- Weizenbaum, Joseph (1994): Die Macht der Computer und die Ohnmacht der Vernunft. Frankfurt am Main: Suhrkamp
- Wiener, Norbert (1963): Kybernetik. Regelung und Nachrichtenübertragung im Lebewesen und in der Maschine. Düsseldorf: Econ