

**Referat Computerethik**  
**Ethische Fragestellungen zur Künstlichen Intelligenz**

**11. Juli 2003**

**von Holger Biehl und Martin Eichhorn**

# Inhaltsverzeichnis

Einleitung.....	3
Joseph Weizenbaum – Die Person.....	4
Was ist Künstliche Intelligenz?.....	5
Umsetzung – Realität oder Utopie?.....	7
Einsatzgebiete der KI.....	11
Ethische Fragestellungen zur „Künstlichen Intelligenz“.....	13
Extreme Ansichten.....	14
Fazit.....	19
Glossar.....	20
Literaturverzeichnis.....	22
Weiterführende Literatur und Weblinks.....	22

## **Einleitung**

Dieser Aufsatz beschäftigt sich mit einem Teilgebiet der Informatik (Computer Science), der „Künstlichen Intelligenz“ (im folgenden auch als KI bezeichnet). Zunächst werden die grundlegenden Begrifflichkeiten und Richtungen der künstlichen Intelligenz beschrieben und auch einige technische Details und Probleme angesprochen, um danach kritische Betrachtungen auf der Basis verschiedener Essays [Weizenbaum01] des emeritierten Informatik-Professors und Philosophen Joseph Weizenbaum anzustellen.

Viele Aspekte in der KI werden noch heute kontrovers diskutiert und die Forschungsgemeinde ist in viele Lager gespalten. Der Leser sollte nach dieser Lektüre die verschiedenen Hauptrichtungen der KI und ihrer Anhänger kennen und zumindest einen kleinen Einblick in die vielen ethischen und moralischen Probleme erlangen, die durch dieses Thema aufgeworfen werden. In dieser wie auch in vielen anderen Wissenschaften stellt sich die Frage, an welchem Punkt man einfach nicht mehr weitergehen sollte, um gewisse ethische und moralische Grundsätze nicht antasten zu müssen.

Letztendlich sollte sich jeder eine eigene Meinung bilden und wir denken, dass diese Ausarbeitung eine gute Basis für Überlegungen darstellt.

Wir haben versucht, technisch so wenig wie möglich ins Detail zu gehen. Um die verschiedenen Überlegungen und Ansichten besser beurteilen zu können ist allerdings ein gewisses Hintergrundwissen unumgänglich. Daher werden an einigen Stellen auch technische und informationstheoretische Punkte diskutiert.

## ***Joseph Weizenbaum – Die Person***

Joseph Weizenbaum wurde am 8. Januar 1923 in Berlin geboren und emigrierte 1936 mit seiner Familie in die USA. Er studierte zunächst Mathematik an der Wayne-Universität in Detroit und arbeitete während des zweiten Weltkrieges als Meteorologe bei der US-Luftwaffe. Danach wurde er wissenschaftlicher Assistent an der Wayne-Universität und war führend an der Entwicklung von Computern und Computersprachen beteiligt bis er dann 1963 einen Ruf an das Massachusetts Institute of Technology (MIT) in Cambridge, Massachusetts erhielt.

Als Professor für Informatik kennt er sich auf dem Fachgebiet bestens aus, entwickelte er doch im Jahre 1966 das noch heute legendäre KI-Programm ELIZA. Hierbei handelt es sich um ein Programm, das einen Psychotherapeuten vortäuscht. Man kommuniziert auf Textbasis mit ELIZA und bekommt nach einem eingegeben Satz eine Antwort oder eine weiterführende Frage gestellt. In einem eingeschränkten Rahmen lieferte die Software verblüffende Ergebnisse, nach einer Weile fällt einem menschlichen Benutzer allerdings doch auf, dass das Gespräch sich im Kreis bewegt.

1976 erschien sein Werk „Die Macht der Computer und die Ohnmacht der Vernunft“ („Computer Power and Human Reason“), das nicht nur als computerkritisches, sondern als gesellschaftskritisches Standardwerk gilt. Weizenbaum ist seit vielen Jahren als Kritiker unreflektierter Computereuphorie bekannt. In vielen seiner Aufsätze zeigt er die Kontexte auf, in denen Computer entwickelt werden und verweist dabei auf die enge Verknüpfung mit dem militärischen Komplex. Er kritisiert die Allmachtsphantasien der Künstliche-Intelligenz-Forschung und fordert eine Kultur der Verantwortung des Wissenschaftlers für seine Arbeit.

Er ist auch Mitbegründer der „Computer Professionals for Social Responsibility“ (USA) und des "Forum der Informatiker für Frieden und gesellschaftliche Verantwortung - FIFF" (Deutschland).

Im Jahre 1988 emeritierte Weizenbaum, ist seitdem aber immer noch in Vorträgen und Reden aktiv. Im Januar feierte er seinen 80. Geburtstag.

## ***Was ist Künstliche Intelligenz?***

Der Begriff „Künstliche Intelligenz“ ist aus dem englischen „Artificial Intelligence“ übersetzt worden. Der englische Begriff wurde dabei in den fünfziger Jahren geprägt. Allerdings spiegelt die deutsche Übersetzung nicht ganz die Bedeutung des Originals wieder: „artificial“ heißt neben „künstlich“ nämlich auch „unecht“ und „Schein-“. Eine künstliche Intelligenz verhält sich also so, „als ob“ sie Intelligenz hätte. Auf welchem Wege dies erreicht wird, ist dabei völlig irrelevant.

Man muss allerdings erwähnen, dass die Forschergemeinde der KI sich hier schon in zwei Lager aufteilt: die Aktivisten und die Ideologen. Die Aktivisten wollen von den Computersystemen nur Antworten, die auch von einem Menschen stammen könnten (wie dies im inneren realisiert wird, ist ihnen egal). Sie sind damit sehr viel pragmatischer veranlagt als die Ideologen. Die Ideologen streben danach, Modelle des menschlichen Denkens zu entwerfen, und die Arbeit des menschlichen Gehirns zu verstehen, um diese auf dem Computer nachzubilden und letztlich den Menschen zu übertreffen.

Beiden gemein ist das Fernziel, den Computer so zu programmieren, dass sein Verhalten nicht mehr von dem eines Menschen unterscheidbar ist. Diese Systeme werden auch „intelligente Agenten“ genannt, sie sollen selbständig operieren und intelligentes Verhalten aufweisen. Im Rahmen dieser Definition wäre der Mensch auch ein intelligenter Agent.

In der Informatik unterscheidet man verschiedene Arten „intelligenter Agenten“:

- Ein Roboter (robot) hat als Kern einen Computer und agiert in der physikalischen Welt. Er hat durch verschiedene Mechanismen die Fähigkeit, seine Umwelt zu beeinflussen. Dabei muss nicht unbedingt an Roboter gedacht werden, die menschenähnlich aussehen (Androide).
- Ein Software Roboter (softbot) hat eine Wissensbasis und gibt Antworten und Ratschläge (agiert aber nicht in der physikalischen Welt). Der menschliche Benutzer formuliert seine Fragen in einer dem Computer verständlichen Weise. Dies muss nicht notwendigerweise über menschliche Sprache geschehen.

Es gibt noch andere „intelligente Agenten“ (z.B. Web-Roboter, Kognitive Simulation, etc.) aber diese genauer zu erläutern würde im Rahmen dieses Aufsatzes zu weit führen. Die beiden erstgenannten reichen aus, um die Problematiken zu erkennen.

### Roboter-Fußball

Ein ganz lustiges Forschungsgebiet, was derzeit von Forschern aus der ganzen Welt betrieben wird, ist der Roboter-Fußball. Es werden regelrechte Weltmeisterschaften<sup>1</sup> ausgetragen. Ein besonders erfolgreiches Roboter-Team wurde von der Universität Freiburg<sup>2</sup> entwickelt. Sie spielen seit Jahren in der Weltspitze mit.



Abbildung 1: Die Roboter-Fußball-Mannschaft der Universität Freiburg

---

1 <http://www.robocup.org/>

2 <http://www.informatik.uni-freiburg.de/~robocup/>

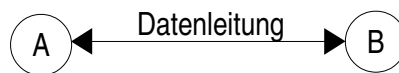
## ***Umsetzung – Realität oder Utopie?***

Dieses Fernziel der „Künstlichen Intelligenz“ wirft bei genauerer Betrachtung einige Fragen auf.

### *Wann hat man das Ziel erreicht, einen intelligenten Agenten zu programmieren?*

Die Beantwortung dieser Frage ist wohl nicht so einfach, wie es im ersten Moment erscheint. Hierzu kann man sich viele Gedankenmodelle überlegen, die testen sollen, ob ein Computer menschlich genug ist.

Die naheliegendste Versuchsumgebung wird vom so genannten Turing-Test beschrieben. Folgende Skizze soll den Versuchsaufbau des Tests erklären:



Der Mensch sitzt an Computer-Terminal A und führt ein ausreichend langes Interview mit B (Turing schlägt 30 Minuten vor). Dabei soll dem Menschen nicht klar sein, ob die Antworten von einem KI-System oder einem anderen Menschen gegeben werden. Wenn die Antworten von einem Computer erzeugt wurden und der Tester nach Ablauf des Interviews der Ansicht ist, dass er mit einem anderen Menschen kommuniziert hat, dann kann man von einem „intelligenten Agenten“ sprechen.

Bei genauerem Hinsehen erweist sich dieser Test allerdings als unzureichend:

- Selbst wenn man die Versuchsdauer deutlich verlängern würde, könnte man in dieser Zeit nur eine sehr begrenzte Anzahl von Fragen stellen, um das System zu testen. Die Antworten könnten z.B. aus einer Datenbank stammen (durch bloßes „Nachschlagen“).
- Der Test-Ausgang hängt stark von den Fähigkeiten des Fragestellenden, sowie von seinem Wissen über das KI-System ab (wenn er die Schwächen des KI-Systems kennt, kann er sie gezielt abfragen).

## Kann ein Computer etwas verstehen?

Hierzu zunächst ein Gedankenexperiment: Der chinesische Raum (nach John R. Searle)

Jemand, der chinesisch weder lesen noch schreiben kann, sitzt in einem Raum. Dort gibt es keine Türen – und keine Fenster. Nur einen kleinen Briefkastenschlitz. Diese Person hat keinerlei Kontakt zur Außenwelt, wird aber ausreichend mit Nahrung versorgt. Ihr steht ausschließlich ein dickes Handbuch in ihrer Muttersprache zur Verfügung, das Regeln beinhaltet, um auf chinesisch gestellte Fragen chinesische Antworten zu formulieren. Durch den Schlitz erhält die Person einen Zettel mit chinesischen Schriftzeichen. Unter Zuhilfenahme des Handbuchs erzeugt er einen neuen beschrifteten Zettel (mit chinesischen Schriftzeichen) und reicht diesen nach außen.

Searle wollte mit diesem Gedankenspiel nachweisen, dass der Computer nichts versteht. Viele Wissenschaftler akzeptieren dieses Experiment als Nachweis aber es besteht noch keine breite Einigkeit über diese Frage. Kritiker behaupten, dass das System als Ganzes verstehen kann. Wenn auch nicht die Einzelkomponenten.

Joseph Weizenbaum hat auf die Frage des „verstehenden Computers“ eine eindeutige Position: „Menschen können einander aufgrund ihrer persönlichen Lebensgeschichte verstehen. Weil keine zwei Menschen die gleiche Lebensgeschichte haben, können sie sich auch nicht perfekt verstehen. Der größte Teil des Sprachverständnisses steht in Kontexten, die wir mit anderen teilen. Ohne diesen Kontext ist keine Verständigung möglich. Eine Botschaft außerhalb jeden Kontexts ist absurd.“ Ein Wissen über die Empfindungen eines Menschen seien notwendig, um diesen Kontext immer herstellen zu können, zumindest um zu verstehen, was ein Mensch sagen will.

Manche KI-Forscher wiederum behaupten, dass man menschliche Empfindungen wie Liebe, Kummer, Freude, Trauer, etc. durchaus im Computer nachbilden kann und widerlegen damit Weizenbaums Argumentation. Eine lückenlose Beweisführung ist an diesem Punkt scheinbar



unmöglich.

Weitere ähnlich kontroverse Fragen sind z.B. „Kann eine Maschine denken?“ oder auch „Kann ein Computer/Roboter ein Bewusstsein haben?“. Aber auch hierauf scheint eine Antwort in naher Zukunft nicht zu finden sein.

### Aktivisten und Ideologen

Weizenbaum ist der Ansicht, dass die Ideologen heute einen größeren Einfluss auf die Meinung über die „Künstliche Intelligenz“ in der breiten Öffentlichkeit haben. Hans Moravec<sup>3</sup> meinte z.B.: „Jede wesentliche Funktion des Menschen, sei sie nun körperlich oder geistig, wird schon bald ein künstliches Gegenstück haben. Die Robotiker werden eine Maschine schaffen, die denkt und handelt wie ein Mensch, so wenig sie ihm auch in physischen oder geistigen Details ähneln mag. Solche Maschinen können unsere zivilisatorische Evolution vorantreiben.“ Dies klingt zunächst sehr beeindruckend, weil man sich die Entwicklung eines solchen Systems als sehr schwierig vorstellt. Es suggeriert weiterhin, dass die Forschung bereits ziemlich weit fortgeschritten ist. Dies zeigt, so Weizenbaum, dass die Ideologen häufig an ihren Visionen gemessen würden, die Aktivisten hingegen an ihren wirklichen Taten und damit etwas im Schatten stünden.

Dabei werden die Ergebnisse der Aktivisten bereits an vielen Stellen erfolgreich eingesetzt. Man denke hierbei z.B. an Schachcomputer, die Schach auf Weltklasse-Niveau spielen, oder Handschrift-Erkennungssysteme. Darauf wird im folgenden Kapitel näher eingegangen.

Hieran sieht man aber auch gleich die Grenzen des heute Möglichen. Die heute eingesetzten KI-Systeme sind immer auf ein sehr kleines Einsatzgebiet spezialisiert und führen dabei bereits die heute Technik ans Limit. Alle Forscher sind sich zwar einig, dass das Fernziel der „Künstlichen Intelligenz“ mit den heute denkbaren Computern nicht erreichbar ist, die Meinungen gehen allerdings sehr stark auseinander, wann unsere Technik den nötigen Stand erreichen wird und ob

---

<sup>3</sup> Hans Moravec, Leiter des Mobile Robot Laboratory an der Carnegie Mellon University (Pittsburgh, PA), Autor des Buches „Mind Children“, in dem er seine Vorstellungen über die Zukunft der künstlichen Intelligenz berichtet. (siehe auch Abschnitt „Extreme Ansichten“)

es überhaupt jemals realisierbar sein wird.

Um eine Vorstellung von der Größenordnung zu bekommen, die nötig wäre, um ein Gehirn im Computer zu simulieren, muss man sich vorstellen, dass das menschliche Gehirn aus bis zu 100 Milliarden Neuronen besteht. Diese arbeiten höchst parallel. Die Basis der heutigen Rechnerarchitektur ist nicht auf wirkliches Parallelrechnen ausgelegt. Bei wachsender Anzahl der Verarbeitungseinheiten steigt der Verwaltungsaufwand zu deren Synchronisierung enorm an. Würde man also jedes Neuron im menschlichen Gehirn durch ein elektronisches Pendant ersetzen, so würde dies eine unvorstellbare Kapazität erfordern.

Alle bis heute erfundenen Rechner basieren auf ein und der selben Grundarchitektur und es ist auch nicht absehbar, dass sich dieser Umstand in näherer Zukunft ändern wird. Die Schranke bei der Parallelisierung ist gerade durch diese Architektur bedingt – daher trägt auch die (scheinbar) so schnelle Weiterentwicklung der Computertechnik nicht wesentlich zur Lösung des obigen Problems bei.

## ***Einsatzgebiete der KI***

Nachdem jetzt klar ist, dass es einige Probleme bei der Entwicklung von KI-Systemen gibt und man vom Fernziel der „Künstlichen Intelligenz“ noch weit entfernt ist, möchten wir einmal die Disziplinen vorstellen, an denen tatsächlich gerade gearbeitet wird und was man sich von ihnen in der nahen Zukunft verspricht.

### *Kognitionswissenschaft*

Dieser Zweig der „Künstlichen Intelligenz“ beschäftigt sich sowohl mit der Untersuchung der Funktionsweise menschlicher kognitiver Fertigkeiten als auch deren Modellierung in KI-Systemen (wobei die Funktionsweise beim Menschen nicht unbedingt nachgebildet werden soll).

Zu den Anwendungen der Kognitionswissenschaft gehören unter anderem folgende Aufgabengebiete:

- Sprachverstehen: menschliche (akustische) Sprache soll zur weiteren Verarbeitung erkannt werden. So soll z.B. ermöglicht werden, dass der Computer über Sprache gesteuert werden kann, oder Diktate entgegengenommen werden können. Gerade auf diesem Gebiet ist man schon heute sehr erfolgreich und durch einiges Training (die Software stellt sich dabei im Moment nur auf die Stimme einer einzelnen Person ein) sind sehr gute Erkennungsquoten erreichbar. Diese Programme sind bereits seit Jahren auch auf dem Massenmarkt (gerade in Textverarbeitungssystemen sehr beliebt) erhältlich und verbessern sich stetig. Ein anderes Anwendungsgebiet für diese Technik würde sich auf dem Geheimdienst-Sektor finden, z.B. beim Abhören von Telefonaten. Der Computer könnte nach bestimmten Schlüsselworten suchen.
- Bildverstehen: Inhalte von Bildern sollen erkannt werden, darunter zählt aber auch das Erkennen von handgeschriebenen und gedruckten Buchstaben (OCR). Natürlich gibt es auch hier einen militärischen Verwendungszweck dieser Technik. So könnten beispielsweise Satellitenbilder automatisch nach Feindaktivität o.ä. durchsucht werden. Auf dem zivilen

Gebiet wäre es für einen Anwender zum Beispiel interessant, eine Anfrage der Form „finde alle Bilder, auf denen ein rotes Haus ist“ an eine Bilddatenbank zu stellen.

### Ingenieurmäßig ausgerichtete „Künstliche Intelligenz“

Ziel dieser Disziplin ist es, Methoden zum Lösen komplexer Anwendungsprobleme, die normalerweise kognitive Fähigkeiten erfordern, zu entwickeln. Nun folgen einige Beispiele für dieses Anwendungsgebiet:

- Deduktionstechniken: KI-Systeme, zur Analyse logischer Zusammenhänge. So können automatisch Widersprüche erkannt und Folgerungen aus bereits bestehenden logischen Formeln erzeugt werden. Diese Systeme könnten Wissenschaftlern unheimlich helfen, man scheint aber schon an die Grenzen des Möglichen gestoßen zu sein.
- Neuronale Netze zur Mustererkennung: Eingabedaten werden auf vorhandene Muster untersucht (auch in der Hoffnung, Muster zu finden, die einem vorher nicht bewusst waren).
- Lernalgorithmen: Der Computer lernt Entscheidungen zu treffen, unterstützt durch „Erfahrungen“ vorheriger Entscheidungen.
- Expertensysteme (wissensbasierte Systeme): Diese stellen den Versuch dar, Wissen hochspezialisierter Fachleute in einem Computersystem nachzubilden. Auf diese Art sollen menschliche Experten bei Routineaufgaben unterstützt und entlastet werden.

## ***Ethische Fragestellungen zur „Künstlichen Intelligenz“***

Vorausgesetzt das Fernziel der „Künstlichen Intelligenz“ ist überhaupt irgendwann einmal zu erreichen, stellt sich natürlich die Frage, ob das überhaupt wünschenswert ist. Sollten die Bemühungen in diese Richtung fortgesetzt werden und wo ist die Grenze, die nicht überschritten werden sollte? Welche Vorteile und Risiken liegen hier verborgen? Joseph Weizenbaum ist durchaus der Ansicht, dass man bestimmte Ideen nicht weiter verfolgen sollte: „Es gibt keine absolute Forschungsfreiheit, und es sollte sie auch nicht geben. Ich will weder eine Diktatur von oben noch eine heimliche Diktatur. Ich will keine gekaufte Diktatur, so wie es heute ist, in deren Licht sich Wissenschaftler und Forscher als Prostituierte des Geistes verhalten.“

Mit Diktatur ist hier gemeint, dass viele große Projekte nur von einigen wenigen Institutionen gefördert und veranlasst werden (können), die dazu finanziell die Macht haben. Diese haben im Allgemeinen nicht im geringsten die Absicht, der Allgemeinheit zu dienen, sondern wollen zunächst ihre eigene Position stärken und ihre Interessen durchsetzen. Mit der größte Etat für diverse Forschungsprojekte wird oft vom Militär und den Geheimdiensten gestellt.

Die an diesen Projekten beteiligten Forscher auf der anderen Seite sehen in solchen Aufgaben oft große Herausforderungen und arbeiten deshalb nur um des Forschens willen mit. Deswegen bezeichnet Weizenbaum diese als „Prostituierte des Geistes“.

Insbesondere verurteilt Weizenbaum die große psychologische Distanz zwischen den Taten und den Auswirkungen des Tuns. Gerade an den großen technischen Hochschulen (so z.B. das MIT) habe das Militär einen sehr großen Einfluss. Sie motivierten und finanzierten sehr viele Forschungsprojekte und die Mehrheit der Wissenschaftler mache sich (so Weizenbaum) nicht bewusst, was mit ihrer Arbeit später geschieht. Im besonderen Maße gelte dies für die künstliche Intelligenz und Robotik, da sich diese Bereiche sehr gut zur Verbesserung von Waffensystemen und zur zunehmenden Ersetzung des Menschen eignen. Menschen haben Moralvorstellungen und Hemmungen, wodurch sie weit vom „perfekten Soldaten“ entfernt sind.

Wo sollte die Forschung aufhören und wo sollte sie fortgesetzt werden? Letztendlich können die

meisten Forschungsprojekte natürlich auch zum Guten eingesetzt werden, aber solange die Menschen immer auch einen militärischen Nutzen aus den Forschungsergebnissen ziehen wollen, wird sich auch dort eine Anwendung finden. Weizenbaum meint, dass sich Ideen prinzipiell nicht verbieten ließen. Diese Ideen kämen von ganz alleine – bevor man diesen Ideen aber schließlich nachgeht, sollte man sich die Konsequenzen bewusst machen.

Weiterhin sei die Öffentlichkeit gegenüber der Naturwissenschaft nicht kritisch genug. Weizenbaum vergleicht den Status der modernen Naturwissenschaft mit dem einer Religion. Die großen technischen Universitäten entsprächen den Kathedralen und die Professoren den Kardinälen. Durch die alltägliche Demonstration unserer Technologien (Flugzeuge, Computer, etc.) würde den Wissenschaftlern fast grenzenlose Autorität und Glaubwürdigkeit verliehen.

## ***Extreme Ansichten***

In den verschiedenen Essays kritisiert Weizenbaum wiederholt einige führende Forscher im Gebiet der „künstlichen Intelligenz“ und Robotik, welche zum Teil sehr extreme Ansichten über die (angestrebte) Zukunft ihrer Disziplin vertreten.

In diesem Abschnitt möchten wir einige dieser Ansichten vorstellen und einfach wirken lassen. Moravec und Minsky gehören sicherlich zu den extremsten Vertretern der Ideologen von daher können ihre Ansichten nicht stellvertretend für alle KI-Forscher betrachtet werden. Die Aktivisten haben in der Regel bodenständigere und mit unserem moralischen und ethischen System verträglichere Vorstellungen. Man sieht aber hieran auch sehr gut Weizenbaums Position, dass die Aussagen der Ideologen zunächst mehr Eindruck auf die Bevölkerung haben.

### *Hans Moravec:*

Moravec arbeitet seit 1980 am Mobile Robot Laboratory an der Carnegie Mellon University (Pittsburgh, PA) und ist inzwischen Leiter dieser Einrichtung. Er beschäftigt sich schon sein gesamtes Leben mit Robotik und ist als ein führender Forscher dieses Gebietes anerkannt. Neben zahlreichen anderen Veröffentlichungen ist er auch Autor des Buches „Mind Children“, in dem er seine Vorstellungen über die Zukunft der künstlichen Intelligenz und Robotik berichtet.

Die folgende Darlegung seines Standpunktes wurde aus [Moravec96] zusammengefasst. Sie beschreibt die Entwicklung der Robotik in den nächsten vierzig Jahren aus seiner Sicht, dabei teilt er diese „Evolution“ in 5 Generationen auf.

Die Roboter auf dem damaligen Stand der Technik nennt er die nullte Generation universeller Roboter. Dies seien bloße „unprogrammierte Reinigungsmaschinen“, die noch nicht universell einsetzbar seien. Im nächsten Schritt werde man Roboter der ersten Generation entwickeln, die über so beschränkte Leistung verfügen, dass die gesamte Rechenkapazität ausschließlich zu Steuerung der Primärfunktionen diene (Bewegung, Analyse der Umgebungsbeschaffenheit, etc.),

sie sollen aber über die Persönlichkeit einer Waschmaschine verfügen. In der zweiten Generation soll die Rechenkapazität dazu ausreichen, dass die Roboter noch etwas dazulernen können – weil sie entweder von einem Menschen trainiert werden oder aus eigenen Erfahrungen lernen. Die dritte Generation wird dann über dermaßen schnelle Computer verfügen, dass die meisten Ereignisse schon vorher simuliert werden können, die Lerngeschwindigkeit wird noch rapider ansteigen. Bevor man diese Roboter allerdings auf wirklich wichtige Aufgaben ansetzen kann, müsse man sie erst eine Weile „spielen“ (besser: „selbständig trainieren“) lassen.

Soweit so gut. Die Roboter der vierten Generation werden dann in der Lage sein, nicht nur die Welt zu simulieren, um dadurch zu lernen, sondern auch gleichzeitig darüber nachzudenken. Diese Roboter können Absichtserklärungen von Menschen verstehen und in detaillierte Programme einbauen. Es könnte sich um Aufgaben wie „Geld verdienen“, „mehr Roboter bauen“ oder „intelligentere Roboter entwerfen“ handeln. Die Maschinen sollen alle Aufgaben übernehmen können, die ein Mensch auch ausführen kann und noch darüber hinaus. Sie könnten ihre Arbeit auch im Weltraum erledigen und gegebenenfalls die Rohstoffe gleich dort verarbeiten, wo sie vorhanden sind. Die Menschen würden dann ohne jegliche Arbeit auf der Erde zurückbleiben und sie würde zu einem Naturreservat werden.

Diese vier Generationen, so Moravec, seien vergleichbar mit vier einhundert Millionen Jahre dauernden Entwicklungsstufen des Gehirns von Wirbeltieren. Man müsse sich darauf einstellen, dass die Grenze zwischen organischen Strukturen und toter Materie in Zukunft zu einer völlig neuen Definition von Leben aufgehoben werden müsse. In 40 Jahren seien die Maschinen uns in körperlicher Hinsicht zumindest ebenbürtig, uns in ihren logischen Denkfähigkeiten aber weit voraus. Sie werden anfangen, sich selbst zu erforschen und ständig zu verbessern.

In [Moravec88] beschreibt er sogar, dass die Maschinen darüber nachdenken, ob es überhaupt noch nötig sei, dass der Mensch existiert und sie würden zu der Ansicht „Richtig, wir brauchen keine Menschen“ gelangen. Er behauptet hier weiter, dass es möglich sein wird, den ganzen Menschen im Computer zu erfassen. „Der Mensch ist eine Menge von Informationen“. Wenn man diese im Computer weiterleben lassen würde, könne der Mensch unsterblich werden.



Diese Meinung erinnert, so Weizenbaum, mehr an einen Science Fiction Roman, als an eine wissenschaftliche Veröffentlichung. Dennoch wurde [Moravec88] von der Harvard University Press herausgegeben, einer der anerkanntesten wissenschaftlichen Verlage in der Welt, die Herausgeber hielten diese Ansichten ganz und gar nicht für Science Fiction. Dies zeigt, wie die Wissenschaft heute ihre Autorität genießt. Die Menschen glauben an alles Mögliche, weil die Wissenschaften an sich für jeden einzelnen viel zu undurchschaubar geworden sind.

Es zeigt aber auch, dass die Wissenschaftler sich über ihre Taten Gedanken machen sollten, bevor sie große Schwellen überschreiten. Ein großes Problem liegt nach Weizenbaums Auffassung in unserer Kultur begraben. Heute gehe die Tendenz ganz klar zur Abgabe von Verantwortung. Niemand sei bereit, heute noch Verantwortung zu übernehmen und dieses Verhalten würde durch unser gesamtes System nur unterstützt werden.

### Marvin Minsky

Minsky ist Professor für Elektrotechnik und Informatik am MIT und ein führender Forscher im Bereich der „künstlichen Intelligenz“, hat aber auch einige Forschungsarbeiten in Richtung Robotik und verwandten Gebieten veröffentlicht. Er ist ein sehr angesehenes Mitglied der internationalen Forschergemeinde und hat zahlreiche Auszeichnungen für seine Arbeiten erhalten.

*Zur Verdeutlichung einiger seiner Gedanken, haben wir einige zentrale Punkte aus seinem Text „Will Robots Inherit the Earth?“ zusammengefasst:*

Minsky sagt, jeder wolle Weisheit und Wohlstand. Das Problem sei, dass unsere Gesundheit uns oft hindere, dies zu erreichen. Um unser Leben zu verlängern und unseren Geist zu verbessern, müssten wir in Zukunft unsere Körper und Gehirne austauschen. Sobald wir von den Einschränkungen, die uns von der Biologie vorgegeben sind, befreit seien, könnten wir selber über die Länge unseres Lebens entscheiden – mit der Option auf Unsterblichkeit – und könnten auch unter anderen, unglaublichen Fähigkeiten wählen.

Auch wenn man alle Krankheiten erfolgreich bekämpfen könnte, würden unsere Körper nur einen sehr begrenzten Zeitraum überleben können. Ein noch größeres Problem sei allerdings die sehr begrenzte Leistungsfähigkeit des menschlichen Gehirns. Es lerne viel zu langsam und habe auch nur begrenzte Kapazitäten. Wenn die Forschung weit genug ist (und Minsky scheint keine Zweifel zu haben, dass es dazu kommen wird), sei man in der Lage mit Hilfe von Nanotechnologie Ersatz-Körper und -Gehirne zu konstruieren. Solche künstlichen Gehirne könnten millionenfach schneller arbeiten, also könnten unsere „Geistes-Kinder“ auch entsprechend schneller denken. Einem solchen Wesen könnte eine halbe Minute erscheinen, wie für uns mehrere Jahre, und jede Stunde wie die komplette Lebensdauer eines Menschen. Auch könnten diese bereits mit viel Wissen auf die Welt kommen, so dass die lange Lernzeit entfalle.

Es sei unnötig zu sagen, dass wir durch den Austausch unserer Körper und Gehirne letztlich zu Maschinen würden.

Auf die Frage, ob die Maschinen uns dann ersetzen würden, meint Minsky, es mache seiner Ansicht nach nicht viel Sinn von „denen“ und „uns“ zu sprechen. Er bevorzuge stark die Haltung Moravecs, welcher vorschlägt, dass wir diese zukünftigen intelligenten Maschinen als unsere „mind-children“, unsere „Geistes-Kinder“, betrachten sollten.

Die Menschen tendierten dazu, sich als die Spitze der Evolution zu sehen. Aber die Evolution sei noch nicht abgeschlossen. Wir würden uns sehr deutlich weiterentwickeln – nur nicht auf die vertraute und langsame Darwin-Methode. Man könne nun neue Systeme „designen“ auf der Basis von „unnatürlicher Selektion“, welche konkreten Plänen und Zielen folgen könne. Wir könnten also selber bestimmen, welche Eigenschaften in die nächste Generation vererbt werden sollten.

Großen Nachholbedarf sieht Minsky bei der Ethik: es gebe kein ethisches System, welches in der Lage wäre, sich den Herausforderungen zu stellen, mit denen wir bereits heute konfrontiert seien. Sicherlich müssten wir unsere Vorstellungen vom Kinderkriegen ändern. Heute würden zufällige Individuen geboren. Eines Tages jedoch würden sie „komponiert“ werden, zugeschnitten auf die

vorhandenen Wünsche und Pläne. Überhaupt seien die heutigen Vorstellungen zu sehr auf das Individuum ausgelegt. Da müsste man einen größeren Maßstab anlegen.

Er geht weiter und fragt sich, wie viele von diesen „Geistes-Kindern“ produziert werden sollten, und wer über ihre Eigenschaften entscheidet. Sollte man die Eigenschaften von zwei Individuen zu einem neuen Individuum zusammensetzen? „Imagine a scheme that could review both your and my mentalities, and then compile a new, merged mind based upon that shared experience.“

Auch diese Fragen müssten von einem zukünftigen ethischen System auf vernünftige Art und Weise geregelt werden.

Viele Science Fiction Autoren und große Kinoproduktionen beschreiben ähnliche Szenarien, die illustrieren, wie eine mögliche Zukunft für das Zusammenleben von Menschen und Robotern aussehen könnte. In vielen dieser Geschichten lehnen sich die Maschinen gegen ihre Erschaffer auf und die Menschen enden in der Sklaverei.

Das erschreckende ist, dass Moravec und Minsky keineswegs Science Fiction Autoren sind, sondern zur wissenschaftlichen Elite gezählt werden. Sicherlich haben beide ein sehr abwertendes Menschenbild und verachten biologisches Leben. Minsky: „The brain is merely a meat machine.“; „Gott war nur ein mittelmäßiger Ingenieur, jetzt können wir das viel besser.“ Der Mensch habe zu viele Fehler: er sei zu schwach, werde krank und es dauere zu lange, bis er intelligent wird.

## **Fazit**

Wir sind für uns zu dem Schluss gekommen, dass die Öffentlichkeit im Schnitt zu leichtgläubig gegenüber den Versprechungen und Ankündigungen der Wissenschaft geworden ist. Es scheint nur wenige Leute zu geben, die diesen kritisch gegenüber stehen, insbesondere gibt es zu wenige Fachleute, die ihr eigenes Forschungsgebiet kritisch beurteilen.

Die zu groß gewordene Distanz zwischen der Forschung an sich und der Verantwortung für die Resultate sollte mehr in das Bewusstsein der Forschergemeinde gerückt werden. Bevor man den vielen spannenden Herausforderungen nachgeht, sollte man sich überlegen, wohin die Forschung führen könnte - und das gilt natürlich nicht nur für den Bereich der „Künstlichen Intelligenz“.

Wenn Minsky sagt, die Wissenschaft müsse sich nicht an die Ethik anpassen, sondern die Ethik sei schlichtweg veraltet, dann ist das auf jeden Fall eine Aussage, die viele Menschen hellhörig werden lassen sollte. Man sollte unserer Meinung nach den Wissenschaftlern, die mit solchen Aussagen um sich werfen (und seien sie auch die Päpste ihres Fachgebiets), kritischer begegnen, damit letztlich der Glaube an die Wissenschaft uns nicht zu etwas antreibt, was wir später noch bereuen würden.

Trotzdem sollte man nicht alle KI-Forscher über einen Kamm scheren. Es gibt sehr viele wirklich nützliche Einsatzgebiete für diesen Forschungszweig und einige von ihnen sind nicht so ehrgeizig, dass sie den Menschen von der Erde vertreiben wollen. Letztlich sollte jede Forschung doch dazu dienen, uns das Leben angenehmer zu machen, und nicht zu viele neue Probleme schaffen.

## OCR

Die OCR (Optical Character Recognition = Optische Zeichenerkennung) ist eine Teildisziplin, in der versucht wird, Texte, die nur als Bilddaten vorliegen, in auf Buchstaben basierende Daten umzuwandeln.

## Nanotechnologie

Die Nanotechnologie versucht, technische Strukturen stark zu miniaturisieren. Auf atomarer Ebene soll es schließlich bei der Konstruktion möglich sein, einzelne Atome und Moleküle punktgenau zu kontrollieren. Damit wäre es auch möglich, bestehende Strukturen identisch zu kopieren.

## Neuronale Netze

Ein neuronales Netz soll den Denkprozess im menschlichen Gehirn in elektronischer Form nachbilden. Man erzeugt dazu viele künstliche Neuronen, die in Verbindung mit allen anderen stehen und kann ihnen eine Menge von Eingabedaten übergeben, die in Ausgabedaten umgewandelt werden. Um ein neuronales Netz auf eine bestimmte Aufgabe anzupassen, muss es trainiert werden. Soll es zum Beispiel eine bestimmte Reihe von Bildern in Kategorien einteilen, so muss man zunächst einige Bilder eingeben und am Ende gegebenenfalls das Ergebnis des Netzes korrigieren. Durch diese Korrektur ändert sich das Verhalten der Neuronen im Netz und mit immer mehr Training werden auch die Resultate immer zuverlässiger. Da in einem echten Gehirn bis zu 100 Milliarden Neuronen miteinander verbunden sind, kann man sich vorstellen, warum es noch nicht möglich ist, die gesamte Gehirnfunktion nachzubilden, mal abgesehen davon, dass die Arbeitsweise des Gehirn noch nicht annähernd volligständig verstanden wurde.

## Robotik

Die Robotik ist eine Teildisziplin der „Künstlichen Intelligenz“, in der Computer Systeme eng mit Maschinen gekoppelt werden, so dass sie in unserer Umwelt agieren/reagieren können. Dafür ist es wichtig, durch verschiedene Sensoren (meist audiovisueller Natur) Informationen über die Umgebung zu extrahieren und sinnvoll zu verarbeiten. Die Einsatzgebiete der entwickelten Roboter sind vielfältig. Schon heute gibt es z.B. Roboter, die selbständig staubsaugen oder andere einfache Arbeiten durchführen. Man kann aber auch an Roboter denken, die an Stellen Reparaturen durchführen, die für Menschen unerreichbar oder zu gefährlich sind, möglicherweise als „Astronauten“.

## **Literaturverzeichnis**

[Minsky94] Marvin L. Minsky, *Will Robots Inherit the Earth?*, Scientific American, Oktober 1994

[Moravec88] Hans Moravec, *Mind Children: the future of robot and human intelligence*, Harvard University Press, Cambridge, Massachusetts, Oktober 1988

[Moravec96] Hans Moravec, *Geisteskinder: Universelle Roboter: In vierzig Jahren haben sie uns überholt*, c't Magazin für Computer Technik, Ausgabe 6/96, Juni 1996.

[Weizenbaum01] Joseph Weizenbaum, *Computermacht und Gesellschaft*, Suhrkamp Taschenbuch, Frankfurt, 2001.

### **Weiterführende Literatur und Weblinks**

Joseph Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft*, Suhrkamp Taschenbuch, Frankfurt, 1976

Einige Publikationen auf der Homepage von Marvin Minsky am MIT:  
<http://web.media.mit.edu/~minsky/>

Einige Publikationen auf der Homepage von Hans Moravec an der Carnegie Mellon University:  
<http://www.frc.ri.cmu.edu/~hpm/>