



**RWTH AACHEN  
UNIVERSITY**

**BERICHTE  
AUS DER  
RHEINISCH-  
WESTFÄLISCHEN  
TECHNISCHEN  
HOCHSCHULE  
AACHEN**

**1/92**

**ISSN-NR.  
0179-079X**

# **TECHNIK UND GESELLSCHAFT**

- Technik und Gesellschaft
- Technikakzeptanz und Technologischer Wandel
- Verantwortung der Wissenschaft
- Evolution ethischer Werte
- Risikokommunikation
- Kommunikative Revolution
- Sicherheit in der Energietechnik

## Aus dem Inhalt:

Das Forum Technik und Gesellschaft	5
Ein spannungsreiches Verhältnis: Technik und Gesellschaft	9
Ethische Werte und ihre mögliche Evolution	11
Altruismus oder Egoismus	13
Der Ingenieur zwischen Moral und Problemlösung	16
Technischer Wandel im Rahmen des ökonomischen Systems	19
Kann Technologiepolitik gesellschaftliche Probleme lösen?	20
Zur Wechselwirkung von Wissenschaft und Technik	21
Ist wertfreie Forschung möglich?	23
Der Turmbau zu Babel	25
Der Snowball-Effect	27
Metropolis — Architektur zwischen Kunst und Technik	28
Modernisierung und Geschichte	33
Kommunikative Revolution und gesellschaftliche Folgen	35
Telegraph und Börse	37
Das französische Technik- und Kommunikationsverständnis	42
Prozeßleittechnik: Informatik und Automatisierung	44
Technikakzeptanz und Risikokommunikation	46
Risiko-Kontroversen im europäischen Vergleich	50
Zu den Wirkungen technisch vermittelter Kommunikation	52
Informatisierung des Straßenverkehrs	54
Hat Kernenergie Zukunft?	58
Physiker lösen Teilchenlawinen aus	63
Gesellschaftliche Auswirkungen der neuen Kommunikationsnetze	65
Informationstechnologische Netzwerke	66
Paralleles Rechnen mit künstlichen neuronalen Netzen	68
Philosophische Aspekte des Projekts Künstlicher Intelligenz	72
Human-ethologische Untersuchungen für eine menschengerechte Informationstechnik	74
Schreiben mit dem Computer	76
Giftstoffen auf der Spur	80
Hilfe für Querschnittsgelähmte: LinguControl	82
Ein neuer Bildschirmarbeitsplatz in der Radiologie	86
Hilfe für den Narkosearzt	88
Namen und Nachrichten	90
Bücher	92
Chronik	93

### Impressum

Herausgeber  
im Auftrage des Rektors:  
Pressestelle der RWTH  
Templergraben 55  
5100 Aachen  
Tel.: 02 41/80-43 22  
Fax: 02 41/80-36 97

in Zusammenarbeit  
mit dem Forum „Technik  
und Gesellschaft“;  
Sprecher des Forums:  
Univ.-Prof. Dr. phil. Max Kerner

Redaktion:  
Carsten Berg  
Angelika Hamacher  
Toni Wimmer  
(verantwortlich)

Anzeigenredaktion:  
Pressehaus Fatter  
Inge-Ulrike Abels  
Talbotstraße 25  
5100 Aachen  
Telefon: 02 41/166 06 45  
Telefax: 02 41/166 06 69

Gestaltung/Layout:  
Klaus Endrikat

Satz und Druck:  
Druckerei Himmels, Heinsberg

Fotos:  
Peter Winandy (Titel)  
Martin Lux  
W. Jansen  
RWTH  
Wissenschaftszentrum  
NRW, Jutta Fedrowitz

Karikaturen:  
Petra Trübenbach

Das Wissenschaftsmagazin  
„RWTH-Themen“  
erscheint einmal  
pro Semester.  
Nachdruck einzelner Artikel  
und Beiträge,  
auch auszugsweise,  
nur nach Rücksprache  
mit der Redaktion  
und mit Quellenangabe,  
Belegexemplare erbeten.

Sommersemester 1992

# Philosophische Aspekte des Projekts Künstlicher Intelligenz

**D**as Projekt Künstlicher Intelligenz (KI) erscheint im Zuge der modernen Computerentwicklung und zielt darauf ab, Intelligenzleistungen, die dem Menschen vorbehalten schienen, technisch zu rekonstruieren. Es versteht sich von selbst, daß es im vorliegenden Zusammenhang nicht um technikimmanente Fragen der KI-Forschung geht, sondern um metatheoretische Probleme, etwa logische Fragen, Probleme des Intelligenz- und Maschinenbegriffs, der technik- und sozialphilosophischen oder auch ethischen Beurteilung der KI-Technologie und so weiter. In dieser Kurzdarstellung werden in einem ersten Teil sehr knapp einige der in diesem Zusammenhang diskutierten philosophischen Fragestellungen angedeutet und in einem zweiten Teil etwas ausführlicher die Bedeutung der Gödeltheoreme für das Problem Künstlicher Intelligenz ins Auge gefaßt.

## Philosophische Aspekte des KI-Programms

1. Einige logische Fragen: Gibt es prinzipielle logische Grenzen für die Möglichkeit maschineller Operationen (siehe hierzu den zweiten Teil)? Gibt es eine Logik der Reflexion? Sind semantische Gehalte algorithmisch faßbar? Sind dafür Formen einer ‚fuzzy logic‘ oder auch ‚para-konsistenter‘ Logiken von Bedeutung? Wie wäre der Informationsbegriff adäquat zu fassen?

2. Intelligenzbegriff: Was ist überhaupt menschliche Intelligenz? Kritiker des KI-Programms argumentieren, daß zentrale Leistungen derselben nicht technisch reproduzierbar seien, zum Beispiel: das Verständnis semantischer Gehalte; intuitive, ganzheitliche Einsicht; Bewußtsein; Selbstbewußtsein; Emotionalität; Intentionalität; Kreativität; Intersubjektivität; Weltverhältnis und andere. Freilich: Kategorische Unmöglichkeitssätze sind auch eine Form des Dogmatismus.

3. Leib-Seele-Problem: Diese uralte Frage der Philosophie hat

mit der Computer- und KI-Entwicklung neue Aktualität gewonnen. Ungelöst ist nach wie vor das Problem einer der äußeren Beobachtung unzugänglichen ‚Innerlichkeit‘: also die Qualität des ‚Seelischen‘, die schon in der animalischen Empfindung beginnt und sich im Emotionalen und Geistigen fortsetzt und vollendet. (Daß Computern diese Dimension fehlt, hat möglicherweise damit zu tun, daß ihr Verhalten – da ohne biotische Basis – nicht unter dem Prinzip der Selbsterhaltung steht. Computer haben daher kein Weltverhältnis und umgekehrt auch kein um sich selbst besorgtes Selbstverhältnis ausgebildet.) Dennoch werden in systemtheoretischer Perspektive ganz neue Einsichten zum Leib-Seele-Problem möglich,<sup>1</sup> so daß wir einer Lösung vielleicht noch nie so nahe wie heute waren – womit freilich nicht gesagt ist, daß die Lösung nahe ist.

4. Technikphilosophisch-anthropologische Aspekte: In der ‚Intelligenzierungstendenz‘ der KI-Technologie wird gleichsam ein neues Technikparadigma sichtbar: Das in der bisherigen Technikentwicklung wirksame Leitmotiv der Organverstärkung und Organüberbietung wird hier zur Idee der Selbstreproduktion des intelligenten Wesens selbst. Es liegt auf der Hand, daß dies Folgen für das Selbstverständnis des Menschen haben muß, insbesondere dann, wenn sich die hochautomatisierten KI-Systeme gleichsam verselbständigen und ihm als allgegenwärtige, quasi autonome Strukturen gegenüber treten.

5. Sozialphilosophisch-ethische Perspektive: Als eine Form gesellschaftlichen Handelns hat die Technisierung immer auch ethische Implikationen. Nun bedeutet die in der KI-Technik wirksame ‚Intelligenzierungstendenz‘, wie gesagt, eine neue Qualität technischen Fortschritts: Dieser erfaßt zunehmend auch Bereiche, die früher geistigen Tätigkeiten des Menschen vorbehalten waren. Damit erscheinen zugleich völlig neuartige Chancen und Risiken am Horizont. Die damit auch

involvierten Gefahren sind nicht mehr, wie bisher, primär physisch-physiologischer Natur (zum Beispiel Luftverschmutzung, Gesundheitsschädigungen), sondern – das macht die eigentümliche Ungreifbarkeit solcher Risiken aus – wesentlich geistiger Natur, und daraus ergeben sich auch neuartige ethische Probleme. Ein Beispiel für viele: KI-Systeme werden unter anderem – als sogenannte ‚Expertensysteme‘ – zur ‚wissensbasierten‘ Entscheidungsfindung eingesetzt. Grundbedingung für den verantwortlichen Umgang mit solchen Systemen ist sicherlich Systemtransparenz. Diese Bedingung ist für hochkomplexe Expertensysteme aber gar nicht mehr ohne weiteres erfüllbar: Zum einen sind Mammutprogramme dieser Art notwendig ein Gemeinschaftswerk vieler Programmierer oder gar Programmerteams; zum anderen liefert das Expertensystem nur Wissen über einen beschränkten Sachbereich, nicht auch Wissen über dieses Wissen selbst: über seine eigenen Bedingungen, seine Grenzen, in es eingehende Wertungen und so weiter, das heißt, was ihm fehlt, ist Reflexivität. Die Folge ist eine konstitutive Intransparenz solcher Systeme. Sind auf dieser Grundlage überhaupt verantwortliche Entscheidungen zu treffen? Wer zum Beispiel trägt die Verantwortung für systembedingte Fehlentscheidungen, mehr noch: Wer kann diese Verantwortung überhaupt noch tragen: die Programmierer, die Softwarefirma, der Händler, der Benutzer, der Dienstherr? Allgemein stellt sich hier die Frage: Wie ist das Verantwortungspostulat menschlichen Handelns im Umgang mit KI-Systemen erfüllbar? Oder droht hier die Gefahr einer Selbstentmündigung des Menschen? Dies, wie gesagt, nur als ein Beispiel für KI-spezifische ethische Probleme.

## Die Gödeltheoreme und das Problem Künstlicher Intelligenz

Gödels berühmte Theoreme über die ‚Unvollständigkeit‘ formaler Systeme sind immer wie-

der als Argument gegen die Möglichkeit Künstlicher Intelligenz ins Feld geführt worden. Ich werde die sehr komplexen Zusammenhänge auf den hier wesentlichen Sachverhalt reduzieren und mich zudem – was ohne sachliche Verkürzung möglich ist – auf Gödels 1. Theorem beschränken.

Wesentlich ist zunächst einmal, daß sich in hinreichend ausdrucksreichen formalen Systemen ein Satz – nennen wir ihn G – konstruieren läßt, der seine eigene Unbeweisbarkeit aussagt. Gödel beweist nun, daß G im Rahmen des zugrundegelegten formalen Systems prinzipiell unbeweisbar ist. Da der Satz aber ebendies selbst besagt, ist er damit zugleich als wahr erwiesen. Das System enthält mit G also einen wahren, aber gleichwohl strikt unbeweisbaren Satz und wird in diesem Sinn ‚unvollständig‘ genannt.

Gödels Theorem scheint so in der Tat auch etwas zum Problem Künstlicher Intelligenz beizutragen: Denn der menschliche Logiker beweist genau den Sachverhalt – nämlich die Unbeweisbarkeit von G –, der in der formalen Gestalt des Ausdrucks G selbst unbeweisbar ist, das heißt der Logiker kann sich offenbar über die Beschränkungen formaler Systeme hinwegsetzen. Betrachtet man Computerprogramme ebenfalls als formale Systeme, so scheint das Denken des Logikers damit auch die Möglichkeiten des Computers in einem grundsätzlichen Sinn zu übertreffen.

Das ist die Pointe bei einem der frühesten und vehementesten Kritiker einer Maschinentheorie des Geistes, J. R. Lucas.<sup>2</sup> Ähnliche Argumentationen finden sich bei vielen anderen Autoren bis hin zu dem kürzlich erschienenen Buch von R. Penrose, Computerdenken (Heidelberg 1991). Wesentlich ist Penrose zufolge, daß der Logiker die Existenz Gödelscher Beweisbarkeitsgrenzen einzusehen vermag. Das menschliche Bewußtsein hat nach seiner Interpretation gleichsam die Möglichkeit, logische Strukturen in einer Art ‚Wesensschau‘ zu erfassen, ohne

dabei formal-algorithmische Operationen auszuführen, was von einer – notwendig algorithmisch arbeitenden – Maschine prinzipiell nicht zu erwarten sei.

Die These von der grundsätzlichen Überlegenheit menschlichen Denkens über die Maschine mag insofern nicht einmal unplausibel klingen, gleichwohl: Daß Gödels Theoreme etwas Derartiges beweisen, kann schon deshalb nicht zutreffen, weil diese gar nichts über Maschinen aussagen. Im übrigen läßt sich zeigen,<sup>3</sup> daß die Unbeweisbarkeit des von Gödel konstruierten Ausdrucks G nicht auf einem grundsätzlichen Mangel formal-algorithmischer Systeme beruht, sondern allein auf der Selbstreferentialität des Ausdrucks G, der ja seine eigene Unbeweisbarkeit aussagt. Daß der Logiker hingegen die Unbeweisbarkeit von G beweisen kann, ist nur so möglich, daß er zu einem Metasystem übergeht. Er formuliert dort denselben Sachverhalt, den auch G ausdrückt, aber diese metasprachliche Aussage über G ist eben nicht mehr selbstreferentiell. G ist unbeweisbar aufgrund seiner Selbstreferentialität; für die metasprachliche Aussage, die denselben Sachverhalt formuliert, entfällt dieser Grund evidentermaßen.

Genau dieser Punkt, scheint mir, ist in der Auseinandersetzung um die mögliche Inanspruchnahme der Gödeltheoreme für das Problem Künstlicher Intelligenz immer wieder verfehlt worden. Für den Mensch-Maschine-Vergleich ist nicht wesentlich, daß es Grenzen formaler Systeme gibt, sondern daß man sich davon durch Übergang auf die Metaebene befreien kann. Zu klären wäre somit, ob die Möglichkeit eines solchen Übergangs nur für den Menschen oder grundsätzlich auch für die Maschine besteht. Gödels Theoreme liefern dafür, landläufigen Meinungen zum Trotz, nicht das geringste Argument. Aus Gödels Überlegungen folgt nichts für oder gegen die Möglichkeit der Maschine, ihrerseits einen Beweis wie der Logiker zu führen. Das ist –

auch nach Gödel – eine offene Frage.

Das Modell der sogenannten Turing-Maschine, das prinzipiell alles einschließt, was Maschinen überhaupt können, ist nicht etwa schon die Antwort auf diese Frage. Denn es sagt uns im Grunde nur, daß die Maschine ein operationsfähiges, algorithmisch beschreibbares System ist. Man kann aber, zumindest heute, nicht behaupten, das menschliche Gehirn sei etwas ganz anderes. Insofern macht der Hinweis auf die generelle Turing-Modellierbarkeit beliebiger Maschinen gerade nicht die spezifische Differenz von Gehirn und Maschine sichtbar. Die bloße Versicherung andererseits, Gehirnprozesse seien, eben als Gehirnprozesse, von prinzipiell anderer Art als Maschinenoperationen, wofür ich nicht die geringste Legitimation sehe, wäre nur eine *Petitio principii*, das heißt das Mensch-Maschine-Problem wäre damit schon vorweg entschieden. Sollte indes auch das Gehirn Turingmodellierbar sein, so würde das bedeuten, daß die Möglichkeiten von Turingmaschinen heute noch völlig unabsehbar sind. Probleme der sogenannten Komplexitätstheorie, wie sie in der Informatik diskutiert werden, bestätigen das in aller Deutlichkeit. Forscher Unmöglichkeitsaussagen sollte man sich daher tunlichst enthalten. Die verfehlt Berufung auf die Gödeltheoreme kann hierfür ein warnendes Beispiel sein.

Anmerkungen

- 1) Hierzu D. Wandschneider, Anfänge des Seelischen in der Natur in der Deutung der Hegelschen Naturphilosophie und in systemtheoretischer Rekonstruktion, in: M.J. Petry (ed.), Hegel und die Naturwissenschaften, Stuttgart 1987.
- 2) ‚Minds, Machines, and Gödel‘, in: A.R. Anderson (ed.), Minds and Machines, Englewood Cliffs, New Jersey 1964.
- 3) D. Wandschneider, Die Gödeltheoreme und das Problem Künstlicher Intelligenz, in: Ethik und Sozialwissenschaften, Bd. 1 (1990).

Autor:

Univ.-Prof. Dr. phil. Dieter Wandschneider ist Inhaber des Lehrstuhls für Philosophie und Wissenschaftstheorie.