

Künstliche Intelligenz und Expertensysteme

Das Thema künstliche Intelligenz (KI) wurde in den letzten Jahren zumindest in der Bundesrepublik Deutschland überwiegend kontrovers diskutiert¹. Gegenwärtig hingegen ist zunehmend ein Trend zur Akzeptanz der neuen Technik zu konstatieren, wohl nicht zuletzt aufgrund des rasanten Aufstiegs des PCs am Arbeitsplatz wie in den privaten Haushalten². Dennoch gibt es Fragen der Computerisierung der Arbeits- und Lebenswelt, die nicht so einfach ad acta gelegt werden sollten. Dafür lassen sich gesellschaftliche und technische Gründe anführen. Denn beispielsweise können Softwarefehler zum Zusammenbruch von Telefonnetzen oder etwa der Flugüberwachung führen. Nicht ganz unberechtigte Ängste bleiben zumindest in einigen Anwendungsbereichen der künstlichen Intelligenz bestehen. Sie erwachsen einer Abhängigkeit von Systemen, die man nicht versteht, von denen aber zunehmend das Funktionieren gesellschaftlicher Abläufe geprägt ist. Darüber hinaus schafft die Vernetzung der Computer Gefährdungen für den einzelnen und die Gesellschaft, nicht nur im militärischen Bereich und an der Börse. Auch könnte sich der grundsätzliche Widerspruch zwischen einer offenen Informationsgesellschaft und dem Datenschutz durch eine forcierte Entwicklung der künstlichen Intelligenz noch ausweiten³.

Trotz einiger Bedenken werden immer neue Expertensysteme zur Lösung komplexer Probleme in Industrie und Handel, im militärischen und medizinischen Bereich angeboten. Viele Projekte sind oder waren spekulativ und ihre Realisierungschancen damit beschränkt. Nachdem heute die Programmziele durch die Entwickler zunehmend bescheidener formuliert werden, sind nun die Verwirklichungsmöglichkeiten größer, auch wenn es noch kaum Verwender von Expertensystemen gibt, die nicht zugleich ihre Programmierer sind. Insgesamt sind die Realisierungsmöglichkeiten der künstlichen Intelligenz noch eher begrenzt. Dies reizt zu Spekulationen und bietet einen willkommenen Anlaß für neue Mythologien. Aber nicht die überzogenen und falschen Erwartungen oder Befürchtungen sind das eigentliche Problem in der Frage nach Spekulation und Realisation hinsichtlich der künstlichen Intelligenz. Dieses liegt vielmehr in der ethischen Frage beschlossen, ob jede technische Spekulation im Bereich der Expertensysteme realisiert werden sollte oder nicht.

Der künstlichen Intelligenz als Teildisziplin der Informatik oder als Teilprojekt der Cognitive Science⁴ geht es um die Modellierung kognitiver Leistungen⁵. Ziel ist die Nachahmung verschiedener menschlicher Fertigkeiten wie Wahrneh-

mung, Fortbewegung, Bildverstehen und Problemlösung. Insbesondere bezeichnet man mit künstlicher Intelligenz die Fähigkeiten einer vom Menschen gemachten Maschine, menschliche Methoden der deduktiven (schlußfolgernden) und induktiven (auf Erfahrung beruhenden) Beschaffung und Anwendung von Wissen und Gedankenführung nachzuahmen.

Expertensysteme sind ein wichtiges Teilgebiet der künstlichen Intelligenz. Hier dient ein regelbasiertes Programmierprinzip der zielgerichteten Strukturierung von Wissen. Ziel ist die automatische und routinierte Problemlösung. Expertensysteme können nicht selbständig agieren oder reagieren. Ihr Vorteil besteht vielmehr darin, daß sie mit Unsicherheiten oder Widersprüchen umgehen können. Als Einsatzgebiet kommt daher die Diagnose, die Expertise und die Beratung in Problemsituationen in Frage. Ferner eröffnen sie Zugänge zu Daten, selektieren sie, entwickeln Konfigurationen von Wissen, planen und leisten Hilfen bei der Strukturierung von Information. Auch für den Unterricht und zur Entscheidungsunterstützung sind derartige Systeme einzusetzen. Branchen mit Interesse für Expertensysteme sind das Militär, die Medizin, die Offshore-Ölbohrung, die chemische Industrie, der Maschinenbau und die Banken.

Expertensysteme sind überall dort sinnvoll, wo routinisierte Arbeitsabläufe gesteuert werden müssen. Auch bei der Fehleranalyse sind sie vielfach anwendbar. Gerade in Produktion und Entwicklung, aber auch in der computergesteuerten Auftragsabwicklung können Expertensysteme herangezogen werden. Sie lassen sich als Speerspitze einer Informatisierung der Arbeitswelt verstehen. Hier versprechen sich die meisten ihrer Befürworter eine Flexibilisierung der Produktion wie der Arbeitszeit. Vor übertriebenen Hoffnungen wie Befürchtungen sei allerdings gewarnt. Im Augenblick ist die Zahl der wirklich anwendungsreifen Systeme wie ihre kommerzielle Bedeutung gering.

Der Vorteil der Expertensysteme liegt in ihrem im Vergleich zum Menschen schnelleren Umgang mit Wissen. Probleme könnten durch die Standardisierung von Kommunikation entstehen. In Banken sollen sie automatisierte Expertisen im Abwickeln von Bankgeschäften erstellen, sich auf Fehlersuche begeben und situationsunabhängige Bewertungen von Fakten und eine problemorientierte Beschaffung von Informationen erzielen. Hier sollen vor allem Routinearbeiten ersetzt werden. Auch die automatisierte Produktion – die Fabrik ohne Arbeiter – kann ohne Expertensysteme nicht auskommen. Die wichtigsten Anwendungsgebiete liegen bislang aber im militärischen Bereich und in der Medizin.

Medizinische Expertensysteme

In der Medizin sollen Expertensysteme das Beratungsgespräch zumindest in einigen Teilbereichen ersetzen. Die hier anfallenden Daten könnten mit den moder-

nen Methoden zuverlässiger verarbeitet und zu komplexerer Information vernetzt werden. Zudem lassen sich medizinische Expertensysteme zur Fehlerkontrolle und zur Grundlagenforschung in unstrukturierten Problembereichen heranziehen. Die Entwicklung medizinischer Expertensysteme befindet sich zwar noch im Anfangsstadium, so daß eine umfassende Folgenabschätzung derzeit nicht möglich ist. Dennoch lassen sich gewisse Entwicklungstendenzen bereits aus den vorliegenden Programmen gewinnen.

Einige von ihnen seien hier genannt: Mycin diene in den frühen 70er Jahren zur Diagnose- und Therapieunterstützung bei Infektionskrankheiten. Weizenbaums natürlichsprachliches Dialogsystem Eliza sollte Hilfestellungen bei der psychologischen Beratung leisten. Medicil kann bei der Diagnose der Ursachen von Bauchschmerz eingesetzt werden. Dieses Programm vergleicht die eingegebenen Daten mit den Informationen von über 6000 Patienten hinsichtlich der möglichen Ursachen von Bauchschmerzen. Help in Gießen schließlich ist ein Konsultationssystem für die Intensivmedizin. Hier wird den nachfragenden Ärzten aktuelles medizinisches Wissen schneller zur Verfügung gestellt. Hier wird also der Arzt und nicht der Patient durch das System beraten.

Expertensysteme sollen so die medizinische Wissensexplosion zu bewältigen helfen und damit zur Kostendämpfung im Gesundheitswesen, das heißt zur Rationalität und Wirtschaftlichkeit des medizinischen Versorgungssystems beitragen. Ihr Einsatz setzt eine EDV-Ausstattung und einen Computer zur Datenerfassung voraus. Für die Arztpraxis sind derartige Systeme noch zu teuer und daher eher für das Krankenhausinformationsnetz geeignet. Zu ihren Vorteilen zählt die Meßwertverarbeitung. Daher erscheint ihre Verwendung besonders empfehlenswert in Diagnosezentren und im Informationsverarbeitungssystem der medizinischen Forschung. Auch zur Standardisierung und Objektivierung ärztlicher Leistungen können Expertensysteme beitragen.

Medizinische Expertensysteme können komplexes Wissen als Basis für ärztliche Entscheidungen bereitstellen. Eine schnellere Therapieentscheidung und Arzneimittelverordnung scheint damit erreichbar zu werden. Durch besseren Datenaustausch wäre es zudem möglich, Doppeluntersuchungen zu vermeiden. Vorläufig beschränkt sich ihr Einsatz aber auf begrenzte und spezifische Diagnosebereiche. Doch bereits bei den ersten Anwendungsfällen zeigt sich, daß der maschinenerarbeitete Vorschlag eine erhebliche Vorentscheidung bedeutet, die gerade bei unerfahrenen Ärzten zur Delegation von Verantwortung an das System führen kann. Die Expertensysteme müßten daher so programmiert werden, daß eine bloße Übernahme von Therapieplänen unmöglich wird. Hier wird eine besondere Ausbildung von Ärzten im Umgang mit Expertensystemen erforderlich sein.

Durch die Entlastung von routinemäßigen Arbeiten wird der Arzt voraussichtlich mehr Zeit für Patienten oder für andere Aufgaben haben. Probleme

könnten dort entstehen, wo durch Expertensysteme der Arzt verdrängt wird. Außerdem besteht die Gefahr der Entpersönlichung und Enthumanisierung des Verhältnisses von Arzt und Patient, die allerdings bei einem reflektierten Umgang beider Betroffenengruppen mit den Systemen als nicht besonders groß anzusehen sein dürfte. Der Patient hat aus ethischer Perspektive zudem Anspruch auf einen umfassenden Datenschutz. Allerdings ist es für die Erstellung der Diagnose durch ein Expertensystem nicht erforderlich, personenbezogene Daten zu verarbeiten. Ebenfalls nicht wünschenswert wäre eine Entwicklung, die Patienten mit Hilfe von Expertensystemen zur Selbstdiagnose verleitet.

Aufgrund der Fehlermöglichkeit der Expertensysteme hat der Arzt die Verpflichtung zur Überprüfung der Ergebnisse und Vorschläge der Maschine. Dies kostet nun wiederum Zeit und kompensiert einen Teil der Zeitersparnis durch das System. Dies muß in Kauf genommen oder auf das System verzichtet werden, denn eine kritiklose Übernahme des von Maschinen erarbeiteten Wissens ist mit dem ärztlichen Ethos so lange nicht vereinbar, wie derartige Vorschläge fehlerhaft sein können. Eine weitere potentielle Gefahr ist die drohende Enthumanisierung infolge der Rationalisierung der Versorgungsabläufe und der Standardisierung und Objektivierung von Kommunikation. Bereits bestehende Tendenzen der Apparatedizin könnten verstärkt werden.

Expertensysteme im militärischen Bereich

Im Golfkrieg wurde es jedem offensichtlich und unweigerlich klar. Ohne Expertensysteme ist auch ein konventioneller Luftkrieg nicht zu führen. Und diese Systeme – oder das sie bedienende Personal – machen bisweilen Fehler. So wurde ein iranischer Jumbo schon vor dem Golfkrieg von einem amerikanischen Zerstörer irrtümlich für einen angreifenden Jagdbomber gehalten und abgeschossen.

Von Anfang an war die KI-Grundlagenforschung in den USA massiv vom Militär gefördert worden. KI wurde eingeführt, weil Menschen der Geschwindigkeit und den Problemen der modernen Luftkriegsführung nicht gewachsen waren. Joseph Weizenbaum meint sogar, daß in einer anderen historischen Situation dies möglicherweise ein Anreiz gewesen wäre, die Aufgabe zu modifizieren, vielleicht auf diese Waffengattung sogar ganz zu verzichten⁶. Aber der Computer sei nun einmal eingeführt worden. Damit habe man neue, kaum mehr verantwortbare Abhängigkeiten geschaffen⁷.

Das Verteidigungsministerium hat in der Entwicklung der Mikroelektronik von Anfang an in den USA die zentrale Rolle gespielt, zunächst insbesondere bei der Entwicklung von Raketen⁸. Allerdings sind wohl mittlerweile fast alle modernen Waffensysteme und nicht nur SDI nicht möglich ohne die intelligenten Rechner, die ein einprogrammiertes Ziel haargenau für diverse Geschosse anzu-

steuern erlauben. Militärische Roboter, Expertensysteme für die Entscheidungsfindung und Autopiloten für Flugzeuge und Kampffahrzeuge sind in der Entwicklung. Die Informatik ist zur kriegsentscheidenden Grundlagenwissenschaft geworden, ähnlich der Kernphysik zur Zeit der ersten Atombomben.

Auch das System der atomaren Abschreckung funktioniert ohne Expertensysteme nicht. Gegen Interkontinentalraketen wurden schon bald elektronische Frühwarnsysteme wie Norad und das Luftverteidigungssystem Sage entwickelt⁹. Das Problem hierbei besteht in den Fehlalarmen, von denen es bisher mindestens vier mit weitreichender Alarmierung gegeben hat. Mindestens viermal stand die Menschheit in der jüngst vergangenen Geschichte vor einem Atomkrieg, ausgelöst durch einen Irrtum eines Expertensystems. Technische Verbesserungen bringen hier nur marginal mehr Sicherheit. Gefordert sind politische Lösungen, wie dies die Entwicklung im Ost-West-Konflikt in den letzten Jahren gezeigt hat. Es muß deutlich werden, daß es nicht die Zuverlässigkeit der Technik ist, die hier Sicherheit schafft, sondern der verantwortungsbewußte Umgang des Menschen mit der Technik.

Mit Darpa (Defense Advanced Research Projects Agency) hat das amerikanische Verteidigungsministerium im Oktober 1983 ein strategisches Forschungsprogramm angeworfen, das für zehn Jahre geplant ist. Für die ersten fünf Jahre sollen diesem bereits 600 Millionen Dollar zur Verfügung gestellt werden. Die Forschung erstreckt sich bevorzugt auf drei Bereiche: Entwicklung von Kriegssrobotern und Kriegspanzern, Computerunterstützung von Kampfpiloten und Managementsysteme für die Kampfführung von Flugzeugträgern.

SDI ist ein computerisiertes zentrales Schlachtenführungsprogramm, das mit ungeheuer hohem Entwicklungsaufwand eingerichtet werden soll. Wie bei vielen militärischen Expertensystemen wird der Mensch aus allen oder doch den meisten Entscheidungsprozessen ganz oder weitgehend verdrängt werden. Dies ist mit nicht unerheblichen Folgeproblemen auch ethischer Art verbunden, die später diskutiert werden sollen. In der Sachlogik des Einsatzes von Expertensystemen im militärischen Bereich überlassen wir zunehmend die Entscheidung über Krieg und Frieden im wesentlichen Maschinen, die wir nicht mehr kontrollieren können. So könnten wir uns schließlich auch der Verantwortung für die Auslöschung der Menschheit entziehen wollen; denn die Aufrechterhaltung der atomaren Abschreckung garantieren Expertensysteme, programmiert von vielen und verschiedenen Experten zu unterschiedlichen Zeiten und mit je eigenen Methoden, so daß eine Konsistenz und Kohärenz des Programms nicht mehr gegeben sein kann. Letztlich kann niemand mehr die Strukturen dieser Verteidigungs- und Kriegsführungssysteme voll durchschauen.

Zugegebenermaßen war militärische Forschung schon häufig einer der wesentlichsten Motoren technologischer Forschung und Entwicklung. Außerdem ist es nicht die Militärforschung, die die Kriege letztlich führt. Trotzdem haben auch

die KI-Forscher Verantwortung für ihre Arbeitsergebnisse, insbesondere da Computer keine unfehlbaren Systeme darstellen. Politische Lösungen sind gefordert, keine Forcierung der KI-Forschung im militärischen Bereich. Militärforschung müßte andere Wege gehen, eventuell KI-Programme dazu einsetzen, um alternative, defensive Kriegsstrategien zu modellieren, die ohne Androhung atomarer Vergeltung wirksame Abschreckung oder besser noch Erhaltung des Friedens ermöglichen, um in anderer Weise Sicherheit vor feindlichen Angriffen zu gewährleisten. Denn die meisten der zur Zeit gültigen Konzepte vermitteln letztlich eher das Gefühl der Angst als der Sicherheit.

Folgenabschätzung für Expertensysteme unter ethischen Gesichtspunkten

Ein gewisses Gefährdungspotential beim Einsatz von Expertensystemen ist in der gesellschaftlichen Anwendung daher nicht auszuschließen. Zur Bewertung der Expertensysteme aus ethischer Perspektive ist somit eine Folgenabschätzung erforderlich. Diese muß den programmatischen Kern der vorhersehbaren Folgen herausarbeiten. Zur Bewertung und Steuerung der Technikentwicklung dient die Szenario-Methode. Sie ist charakterisiert durch eine explorative Vorgehensweise bei der Hypothesenbildung unter Einschluß von Literaturanalysen und Experteninterviews. Szenarien entwickeln mögliche Zukünfte. Sie sind keine Vorhersagen oder Prognosen im strengen Sinn¹⁰.

Ein Einsatz von Expertensystemen in großem Stil vor dem Jahr 2000 ist eher unwahrscheinlich. Dann wird sich eine Nachfrage vor allem im Bereich der Produktion ergeben. Zu diesem Zeitpunkt wird voraussichtlich auch die Computerangst abgebaut sein. Die Kompetenz von Maschinen, Daten und Wissen zu verarbeiten, wird dann akzeptiert sein. Die etwas längere Dauer der Einführung der neuen Technik ist sogar wünschenswert. Denn dies erhöht die Sozialverträglichkeit der Informations- und Kommunikationstechniken.

Das zentrale Problem, das Expertensysteme zu bewältigen helfen können, ist die Wissensexplosion. Die Unsicherheit bei medizinischen, militärischen und wirtschaftlichen Entscheidungen nimmt zu. Zusätzlich wächst der Zeitdruck, unter dem Entscheidungen getroffen werden müssen. Das Aufspüren von Störeignissen, die Überprüfung von Prozeßvorgängen und die Identifikation von Handlungsmöglichkeiten können hier Entscheidungshilfen auf kognitiver Ebene bereitstellen und Routinearbeiten bei der Informationssuche ersetzen.

Einen wachsenden Bedarf an Expertensystemen hat neben Militär und Medizin die Wirtschaft. Ihr Einsatz zur Steuerung der Produktion ist vielfältig; die menschenfreie Autoproduktion oder die Steuerung von Kernkraftwerken sind hier die spektakulärsten Beispiele. Doch ihr Einsatz bleibt bislang teuer. Allerdings weisen die neuen Systeme zunehmend Marktorientierung und damit ein

günstigeres Verhältnis von Kosten und Nutzen auf. Relevante Problembereiche ergeben sich vor allem in Produktion, Verwaltung und Handwerk.

Allerdings werfen die modernen Informationstechnologien auch eine Reihe von Fragen grundsätzlicher Natur auf. Sie kreisen um den Zusammenhang von Humanität und Rationalität und ihrer potentiellen Transformation durch künstliche Intelligenz. Der Prozeß technologischen Könnens hat uns einen Zuwachs an Wissen und Verantwortung beschert, den wir immer weniger zu bewältigen scheinen. In dieser Situation bietet die künstliche Intelligenz an, ein problemlösendes und entscheidungsunterstützendes Modell von Rationalität zu entwerfen. Doch ist die Frage nicht von der Hand zu weisen, ob damit nicht der Teufel mit dem Beelzebub ausgetrieben werden soll. Gegen eine unkritische Fortschreibung des Programms der technologischen Moderne plädiere ich für eine Rückbesinnung auf das Projekt der Aufklärung¹¹ und auf das in ihr beschlossene Humanitätsideal wie des Menschenrechtsethos.

Entscheidend für die Bewertung von Expertensystemen ist daher, ob sich ihr Rationalitätsmodell ethisch-aufklärerisch oder entfremdend auswirkt und ob die gesellschaftliche Einführung von Expertensystemen zu einer größeren Verwirklichung der Menschenwürde wie sozialer Gerechtigkeit führt oder diese gerade verhindert. Sollen, wie häufig bei medizinischen Expertensystemen, Kommunikationssituationen durch den Mensch-Maschine-Dialog ersetzt werden, so muß die Integrität der beteiligten Personen, hier des Arztes und des Patienten, gewahrt bleiben. Kommunikation und ihre personale Dimension darf nicht vollständig durch instrumentelle Formen des Informationsaustauschs ersetzt werden. Dies impliziert keine Ablehnung einer Nutzung der neuen Kapazitäten bei der Wissensverarbeitung angesichts der Wissensexplosion. Die Nachvollziehbarkeit der Schritte aber, die zu einem maschinenerarbeiteten Vorschlag geführt haben, sollte im Interesse der sittlichen Zurechenbarkeit einer Entscheidung gewahrt sein, da sonst die Idee der Kritik, der Autonomie der sittlichen Entscheidung oder des Gewissensurteils durch die Heteronomie des vom Expertensystem erstellten Vorschlags ersetzt wird.

Entscheidungsunterstützende Expertensysteme und verantwortbares Handeln

Im Hinblick auf menschliche Entscheidungsfindung gibt es zwei Klassen von Expertensystemen. Weitere Verbreitung fanden zunächst Expertensysteme, die als „Decision Making Systems“ bezeichnet wurden. Sie wurden vor allem beim Militär und bei der Steuerung von Kernkraftanlagen verwendet und könnten auch bei Verkehrsleitsystemen der Zukunft eingesetzt werden. Hier droht eine vollständige Entsubjektivierung des Entscheidungsprozesses, wobei zu fragen ist, ob dies das höhere Ziel, die Rettung von Menschenleben, eventuell rechtferti-

gen könnte. Eine andere Klasse von Expertensystemen nennt man „Decision Support Systems“, also entscheidungsunterstützende Systeme. Letztere, auf individuelle Benutzer oder Kleingruppen zugeschnittene Programme, basieren auf der Annahme, daß besseres Wissen bessere Entscheidungen erzeugt¹².

Planungshilfen bei Entscheidungen, Problemlösungsstrategien, der Mensch-Maschinen-Dialog und seine Modellierung, die Psychologie der Entscheidung und vernetzte Kommunikation sind daher die Forschungsfelder der Decision-Support-Systems-Programmierer. Rationaler aufbereitetes Wissen etwa in Entscheidungsbäumen könnte Handlungen vernünftiger und nachvollziehbarer machen. So ließe sich die Sehnsucht nach einer rationalen Gesellschaft erfüllen¹³.

Auch bei entscheidungsunterstützenden Systemen gibt es genug Probleme, nicht zuletzt methodischer Art. Zunächst sind mathematische Modelle für Ungewißheit zu entwickeln und Fragen der Quantifizierung von Unsicherheit zu lösen. Die Näherung soll über die Feststellung von Graden des Vertrauens gelingen, die dann in einem probabilistischen Netzwerk eingefangen werden sollen. Dafür wurden zwei Modelle vorgeschlagen, die Wettverhaltens-Interpretation und die qualitative Annäherung. Bei beiden Modellen wird jedoch höchstens eine individuelle Einschätzung festgestellt, die nicht als objektiv gelten darf. Daher scheint sich der Ausweg über eine Festsetzung von oberen und unteren Wahrscheinlichkeiten anzubieten¹⁴.

Wie sicher aber sind maschinenerarbeitete Vorschläge, wenn in ihrem Programm unsicheres Wissen, unsichere Maßeinheiten und subjektive Zustimmung vereint sind? Aus methodisch-metaethischer Perspektive gilt es zudem zu unterscheiden zwischen präskriptivem und deskriptivem Anteil an einer Entscheidung¹⁵. Daß das kognitive und das deskriptive Moment an Entscheidungen simulierbar sein könnte, mag noch einleuchten. Doch wie soll das normative Element angemessen einbezogen werden? Sollte dieses programmierbar sein, bedürfte es einer zumindest hinreichend konkreten deontischen Logik und entsprechend abgesicherter und konsensfähiger Basiswertaussagen, aus denen wertende Urteile als Handlungsempfehlungen abgeleitet werden können. Zumindest im letzteren Bereich fehlt noch eine hinreichend konsistente Theorie.

So läßt sich vermuten, daß „Decision Support Systems“ nicht auf sittliche Entscheidungen zurückgreifen. Der Entscheidungsbegriff hat unterschiedliche Bedeutungen. Dabei besagt der logische Begriff der Entscheidung die Determination durch einen Algorithmus. Modelliert wird der Deduktionsprozeß, der zu einem logischen Urteil führt, dem eine Entscheidung zur Handlung folgen kann. Decision Support Systems möchten mehr Rationalität in die Entscheidungen bringen¹⁶. Doch die Frage bleibt, ob dadurch nicht instrumentell orientierte Rationalität sittliche oder praktische Vernunft ersetzt.

Auch die methodischen Probleme sind nicht zu unterschätzen. Grundsätzlich müssen bei Expertensystemen die beiden Bereiche der Ingenieursanalyse des

technischen Systems und die psychologische Beschreibung menschlicher Fähigkeiten in einer zusammenhängenden Theorie erfaßt werden¹⁷. Aber läßt sich Klugheit oder Tugend programmieren? Sind sittliche Einstellungen menschliche Fähigkeiten, die nach einem Algorithmus ablaufen? Nach traditionellem philosophischem Verständnis läßt sich „praktische Vernunft“ und „sittliche Rationalität“ nicht als Algorithmus darstellen oder gar begreifen. Zudem führt dieser Ansatz in methodische Schwierigkeiten. Denn die übergreifende Theorie, die den Spalt zwischen Ingenieursanalyse und der menschlichen Psyche überbrücken soll, müßte im Sinn der „Decision Support Systems“ wieder maschinenmäßig modellierbar sein. Dies zieht die Dilemmata der Geist-Gehirn-Problematik nach sich¹⁸. „Decision Support Systems“ operieren also weniger mit wirklich sittlichen Entscheidungen, sondern basieren eher auf spieltheoretischen Konzepten, in denen Präferenzfunktionen im Sinn ökonomischer Theorien¹⁹ den Kern bilden. Sie könnten dann programmiert sein nach der Regel der Nutzenmaximierung, also das Programm des rationalen Egoisten abspulen. Dies impliziert aber noch keineswegs moralisches Handeln.

Sicher wäre es übertrieben, die sittliche Intuition des Experten so in den Vordergrund stellen zu wollen, wie die Gebrüder Dreyfus dies tun. Auch Intuitionen sind in nicht geringem Ausmaß fehlbar. Und hier ist das Plädoyer der KI für mehr Rationalität in Entscheidungen durch Strukturierung ihrer wissensmäßigen Voraussetzungen akzeptabel. Das Problem beginnt dort, wo Entscheidungen nicht mehr unterstützt, sondern ersetzt werden sollen. Aber welcher Art kann die Entscheidung sein, wenn der Benutzer den Vorschlag nicht mehr kontrollieren kann, den das System anbietet, oder gar nicht mehr zur Entscheidung zugelassen wird? Wenn Decision-Support-Systems-Programme nicht so aufgebaut werden, daß das Wissen in allen seinen Schritten zumindest dimensionsmäßig vom Entscheidungsträger kontrolliert werden kann, dann wird den Benutzern eine entfremdete Rationalität aufgenötigt, die Aufklärung zunichte macht.

Jedoch erscheint es als durchaus möglich, die Situationsanalyse und die vorhersehbaren Folgen einer Entscheidung, also die kognitiven Aspekte der Entscheidungssituation, zu simulieren und so zur Entscheidungsfindung beizutragen. Die ethische Bewertung von Zielen und Mitteln kann die Maschine nicht übernehmen, allerdings zur Erhellung der Frage beitragen, welche eventuellen Folgen die Verwirklichung eines Ziels hat. Hier liegen die sittlich vertretbaren Konstruktionsziele entscheidungsunterstützender Expertensysteme. Sobald aber der Eindruck erweckt wird, die Maschine unterstütze die Entscheidung nicht nur, sondern ersetze sie, wird die Idee neuzeitlicher Rationalität und der sittlichen Autonomie des Menschen untergraben. Entscheidungen sind dann auch nicht zurechenbar.

Die Idee sittlicher Rationalität fordert, daß die sittlich zurechenbare Entscheidung die Entscheidung einer Person bleiben muß. Aber für bloße Regelanwen-

derung übernimmt man keine Verantwortung²⁰. Genausowenig wird man sich rechtfertigen wollen für die Übernahme eines Maschinenvorschlags. Doch auch dessen Annahme ist eine persönliche Entscheidung, für die Rechenschaft gefordert werden kann. Die subjektlose Maschine kann Verantwortung nicht übernehmen, wohl aber der, der Vorschläge akzeptiert. Trauen sich nicht Anwender von „Decision Support Systems“ zu viel zu, wenn sie glauben, unüberschaubare Entscheidungsvorschläge annehmen zu dürfen? Die mit Emphase vorgetragene These vom Prinzip Verantwortung kann so langfristig von entscheidungsersetzenden Systemen ausgehöhlt werden. Sie passen in gegenwärtige gesellschaftliche Entwicklungstendenzen. Ein Bedarf hinsichtlich entscheidungsunterstützender Expertensysteme besteht schon allein deshalb, weil Weisheit und Klugheit schwer zu rechtfertigen sind²¹. So erscheint es verständlich, wenn man sich hinter maschinenproduzierten Entscheidungen verstecken möchte.

Eine rationale Strukturierung der Wissensbasis für Entscheidungen durch Expertensysteme ist in den meisten Fällen aus ethischen Gründen nicht abzulehnen. Doch die Ersetzung von menschlich-sittlichen Entscheidungen durch Computersimulationen gefährdet das Konzept aufgeklärten Selbstdenkens und der Emanzipation²². Durch Expertensysteme fallen Entscheidungen ohne Berücksichtigung sittlicher Grundsätze nach Kosten-Nutzen-Analysen. Ihre Berechnung erfolgt aber nach spieltheoretisch vorgegebenen Rahmenbedingungen, die zwar Komplexität reduzieren, aber auch Entscheidungsmöglichkeiten festlegen. Der geniale dritte Weg, die aristotelische Mitte zwischen zwei Übertreibungen oder die Suche nach dem rechten Maß sind dem Computer wohl verschlossen, insbesondere wenn sie außerhalb des eingegebenen Horizonts liegen. Nicht selten sind die produktivsten Lösungen für ein Dilemma diejenigen, die die Voraussetzungen verwerfen, die in dieses geführt haben. Und wo entscheidungsersetzende Systeme eingesetzt werden müssen, wie zum Beispiel im militärischen Bereich oder bei der Steuerung von Kernkraftwerken, ist zu vermuten, daß diesen Techniken die humane Angemessenheit fehlt.

Expertensysteme zwischen technokratischem Handeln und verantwortbarer Steuerung der Technologieentwicklung

Expertensysteme erscheinen so als ambivalent. Sie verstärken als entscheidungsersetzende Systeme im allgemeinen Tendenzen einer neuen Weise technokratischen Handelns in überindividueller Form. Andererseits helfen sie durch Wissensmanagement Komplexität zu reduzieren und ermöglichen so in hochkomplexen Entscheidungssituationen verantwortliches Handeln.

Formen subjektlosen und instrumentellen – technokratischen – Interagierens und Entscheidens wurden vorbereitet und eingeübt bei Entscheidungen in

Teams, Firmen und Parteien, die durch wachsende Komplexität und Undurchschaubarkeit Verantwortung verschleiern. Diese Tendenz könnte durch die technischen Möglichkeiten der KI-Forschung zusätzliche Impulse erhalten und zu Konzeptionen des Entscheidens und Verantwortens führen, die zunehmend unkontrollierbar und theoretisch inkonsistent sind und daher gesellschaftlich destabilisierend wirken.

Sowohl die Expertensysteme wie der gesellschaftliche Umgang mit ihnen müssen daher unter ethischen Ansprüchen gesehen werden. Die technologische Gesellschaft muß – um dem Technokratievorwurf entgegen zu können – Formen der Verantwortung von kollektivem Handeln entwickeln, bei dem auch entscheidungsunterstützende Systeme zur Reduktion von Komplexität konsultiert werden können. Dabei bleibt der Ansatzpunkt die individuelle, persönlich zuschreibbare und zu rechtfertigende Zuordnung von sittlicher Verantwortung von Handlungen trotz aller gemeinsam getragenen Entscheidungen. Dies fordert die Idee sittlicher Rationalität und Autonomie: Wenn ich den Vorschlag eines Computers übernehme, trage ich die Verantwortung, auch wenn den Programmierer oder die Gruppe der Spezialisten, die ein Programm erstellt haben, eine Mitverantwortung für programmbedingte Fehlentscheidungen trifft.

Eine rationale Strukturierung der Wissensbasis für Entscheidungen wird man nicht ablehnen. Doch die Ersetzung von menschlich-sittlichen Entscheidungen durch Computersimulationen gefährdet das Konzept aufgeklärten Selbstdenkens und der Emanzipation. Dort, wo im Rahmen der KI-Anwendung vom Benutzer maschinenerarbeitete Vorschläge nicht mehr kontrolliert werden können und bloß vorgeschlagene Entscheidungen übernommen werden sollen, muß man eine Verletzung des Humanitätsideals unterstellen, wie es sich in der Aufklärung herausgebildet hat. Dies ist zu vermeiden, damit Expertensysteme nicht als Teilstruktur technokratischen Handelns etabliert werden. Dazu sind bei der Realisierung von Spekulationen im KI-Bereich nicht nur technische Möglichkeiten zu berücksichtigen, sondern auch die Human- und die Sozialverträglichkeit.

Die Gefahren eines Abdriftens von Expertensystemen ins Technokratische bei der Unterstützung von Entscheidungen sind nicht zu unterschätzen. Sie sind je nach Anwendungsbereich durchaus spezifisch, weisen jedoch einige gemeinsame Strukturmerkmale auf. Die Gefahr einer Uniformierung der Denkmuster und Entscheidungsweisen wie einer Automatisierung und Standardisierung von Kommunikation sind nicht von der Hand zu weisen. Entscheidung ersetzende Expertensysteme sind jedoch noch problematischer. Bei der militärischen Anwendung zum Beispiel liegt möglicherweise die Entscheidung über Krieg und Frieden in der Hand von nicht fehlerfreien Systemen. Die Zuschreibung von Verantwortung bei Entscheidungen unter Einsatz von Expertensystemen ist vielfach ungelöst. Rechtslücken bestehen. Hinzu kommen fehlende oder nur schwach ausgeprägte Kontrollmöglichkeiten und ein oft unzureichender Datenschutz.

Hinsichtlich der Expertensysteme erwachsen die Gefahren für die aufgeklärte Humanität des Menschen aus zwei Quellen, nämlich aus der Art und Weise ihrer Konstruktion und aus dem mehr oder weniger verantwortlichen Umgang der Gesellschaft mit diesen neuen technischen Möglichkeiten. Ethik muß in beiden Bereichen darauf hinarbeiten, daß der Einsatz der Expertensysteme die Gesellschaft nicht enthumanisiert, entindividualisiert und depersonalisiert. Aus ethischer Perspektive sind diese Systeme dem Menschen anzupassen. Sie sollen seine Kreativität und Entscheidungsfähigkeit fördern, nicht diese aufzuheben versuchen. Nur dann wird eine humane Verwendung von Expertensystemen in den Grenzen des Projekts der Aufklärung zu rechtfertigen sein. Nur so kann die Eingliederung der Expertensysteme in eine technokratische Gesellschaftsstruktur, in der die Maschine den Menschen weitgehend beherrscht, verhindert werden.

ANMERKUNGEN

- ¹ Künstliche Intelligenz, hrsg. v. B. Irrgang, J. Klawitter (Stuttgart 1990) 33–46.
- ² Weizenbaum contra Haefner: Sind Computer die besseren Menschen?, hrsg. v. M. Haller (Zürich 1990) 10, 26.
- ³ Ebd. 56.
- ⁴ F. Rose, *Ins Herz des Verstandes. Auf dem Weg zur künstlichen Intelligenz* (München 1985) 53–57.
- ⁵ G. Götz, Möglichkeiten der Automatisierung kognitiver Leistungen. Zur Problematik der „Künstlichen Intelligenz“, in: *Technik und Gesellschaft* Jb. 4 (Frankfurt 1986) 179.
- ⁶ J. Weizenbaum, *Die Macht der Computer und die Ohnmacht der Vernunft* (Frankfurt 1978) 50–52.
- ⁷ R. Trappl, *Impacts of Artificial Intelligence* (Amsterdam 1986) 37.
- ⁸ Rose, a. a. O. 47.
- ⁹ R. Keil-Slawik, *Das kognitive Schlachtfeld*, in: *Künstliche Intelligenz*, a. a. O. 84.
- ¹⁰ Enquete-Kommission „Einschätzung und Bewertung von Technikfolgen. Gestaltung von Rahmenbedingungen der technischen Entwicklung“. Materialien zur Drucksache 10/6801 (Bonn 1987).
- ¹¹ H. Hastedt, *Aufklärung und Technik. Grundprobleme einer Ethik der Technik* (Frankfurt 1991) 224–258.
- ¹² G. S. Henk, *Paradoxes around DSS*, in: *Decision Support Systems. Theory and Application*, hrsg. v. C. W. Holsapple, A. B. Whinston (Berlin 1987) 7.
- ¹³ S. v. Zelewski, *Das Leistungspotential der Künstlichen Intelligenz. Eine informationstechnisch-betriebswirtschaftliche Analyse* (Witterschlick-Bonn 1986) 240, 261.
- ¹⁴ D. Dubois, H. Pride, *Recent Models of Uncertainty and Imprecision as a Basis for Decision Theory. Towards less normative Frameworks*, in: E. Hollnagel, G. Mancini, D. Woods, *Intelligent Decision Support in Process Environments* (Berlin 1986) 3, 5f., 10.
- ¹⁵ S. F. Garriba, *The Use of Information Structures in Risky Decisions*, ebd. 39.
- ¹⁶ G. Volta, *Time and Decision*, ebd. 50, 55.
- ¹⁷ J. Rasmussen, *A Framework for Cognitive Task Analysis in System Design*, ebd. 195.
- ¹⁸ *Denken und Denkenlassen. Problemfeld Künstliche Intelligenz*, hrsg. v. B. Irrgang, J. Klawitter, J. Schmidt (Berlin 1992) 132–134.
- ¹⁹ H. Putnam, *The mental Life of some machines*, in: *Mind, Language and Reality, Philosophical Papers 2* (Cambridge/USA 1975, 1986) 409.
- ²⁰ H. L. Dreyfus, St. E. Dreyfus, *Künstliche Intelligenz. Von den Grenzen der Denkmachine und dem Wert der Intuition* (Reinbek 1987) 49.
- ²¹ Ebd. 259.
- ²² J. Weizenbaum, a. a. O. 338.