

**Seminar SS 2000 – Sozionik:  
Erforschung und Modellierung künstlicher Sozialität**

Günter Hellbardt: "Die Ethik von Agenten"  
Rödiger et al.: " Informatik und Verantwortung"

Richard Kutzky, Juni 2000

Inhaltsverzeichnis:

<b>1. EINLEITUNG .....</b>	<b>2</b>
<b>2. ANSPRUCH UND WIRKLICHKEIT DER INFORMATIONSTECHNIK.....</b>	<b>2</b>
2.1. VERANTWORTUNGSDILEMMATA BEI DER PRODUKTGESTALTUNG.....	4
2.2. VERANTWORTUNGSDILEMMATA UNTER KONKRETEN ARBEITSBEDINGUNGEN ..	5
<b>3. FOLGEN FÜR BENUTZER SOWIE FÜR PRODUKTE .....</b>	<b>6</b>
3.1. ANSATZPUNKTE FÜR VERANTWORTLICHES HANDELN.....	6
<b>4. MENSCH UND MASCHINE.....</b>	<b>7</b>
4.1. MENSCH UND COMPUTER.....	8
4.2. MENSCH UND AGENT .....	8
<b>5. ETHIK VON AGENTEN .....</b>	<b>9</b>
<b>6. DISKUSSION .....</b>	<b>9</b>
<b>7. LITERATUR .....</b>	<b>11</b>

## **1. Einleitung**

Die dieser Arbeit zugrundeliegende Literatur befaßt sich mit der Frage, inwieweit die Informatik und die auf ihr aufbauende Informationstechnik Verantwortung gegenüber der Gesellschaft trägt bezüglich der Verhältnisse, in der ihre Produkte mit den Mitgliedern der Gesellschaft, den Menschen interagieren. Ausserdem wird untersucht, welche Voraussetzungen die Produkte haben müssen um problemlos von der Gesellschaft akzeptiert zu werden und wie bereits in der Entwicklungsphase auf diese Voraussetzungen hingearbeitet werden kann. Da nur einer Minderheit der Gesellschaft die Möglichkeit an die Hand gegeben ist, auf ihre maschinisierten und häufig computerbasierten Arbeitsverhältnisse manipulierend einzuwirken, ist es die Aufgabe der Informatik, dieses Verhältnis in einem Maße zu definieren, die den Bedürfnissen der mit ihnen verbundenen Menschen besser gerecht wird und nicht nur die Bedürfnisse der Maschine berücksichtigt. Im Besonderen wird auf das Verhältnis zwischen Mensch und intelligentem Agenten eingegangen und versucht, die Verantwortungsfrage für selbständiges Handeln des Agenten im Kontext der Verantwortung des Benutzers sowie des Designers dieses Agenten zu beleuchten.

## **2. Anspruch und Wirklichkeit der Informationstechnik**

"Technik ist Anwendung der Wissenschaft im gesellschaftlichen Zusammenhang" [Rödiger et al., 1989]. Die Informationstechnik hat folglich zur Aufgabe, Erkenntnisse der Wissenschaft Informatik in der Gesellschaft zu realisieren. Dabei werden Problemstellungen, die in der Beziehung der Informationstechnik mit der Gesellschaft liegen, von der Informatik oft unterschätzt oder sogar ganz ignoriert. Die Informatik ist aufgrund der fortschreitenden Mathematisierung verschiedenster Anwendungsgebiete in der Lage, sich auf ebenso unterschiedliche wie vielfältige Gebiete zu verbreiten und Lösungen anzubieten. Diese Verbreitung führt aber gerade aufgrund ihrer Vielseitigkeit zu erheblicher Abstraktheit der von ihr angebotenen Lösungen. Diese Abstraktion zieht sich durch alle Konzepte, die für ein Problem gelten wie die algorithmische Herangehensweise oder die Methoden und Datenstrukturen zur Beschreibung der Sachverhalte. Trotz dieser Abstraktheit ist die Zielsetzung des Einsatzes von Computern, vor allem im Arbeitsleben, die Rationalisierung des menschlichen Anteils der Arbeit.

An die Informationstechnik und ihre Produkte werden von der Gesellschaft einige Anforderungen gestellt, die sie erfüllen muß, sollen ihre Produkte von der Gesellschaft akzeptiert werden. Es sind dies die Angemessenheit der benutzten Ressourcen, sowie die Zuverlässigkeit ihrer Produkte hinsichtlich ihrer Einsatzfähigkeit. Aufgrund der vielfältigen Einsatzmöglichkeiten der Informationstechnik und der damit verbunden potentiell hohen Komplexität der Problemstellungen, sind diese Forderungen nicht leicht zu erfüllen. Die Angemessenheit der benutzten Ressourcen bezieht sich in dieser Hinsicht nicht nur auf Mittel rein technischer Art, sondern auch auf die Ressource Mensch und ihre Rolle im maschinellen Arbeitsprozeß. Bei der Modellierung eines Arbeitsprozesses ist es von großer Wichtigkeit, bei der Bestimmung von Schnittstellen zwischen Mensch und Maschine nicht nur auf eine software-ergonomisch ausgerichtete Definition Wert zu legen, sondern im Gegensatz dazu auch den Mensch und seine Bedürfnisse stärker zu berücksichtigen.

Im Sinne der Zuverlässigkeit muß die Informationstechnik bei dem Design ihrer Produkte darauf achten, nach Möglichkeit Fehler zu vermeiden. Es ist mittlerweile

anerkannt, das komplexe Softwaresysteme nicht vollkommen fehlerfrei zu realisieren sind. Im Licht dieses Wissens muß es ein Hauptanliegen der Informationstechnik sein, zuverlässige Testumgebungen für Softwaresysteme zu entwickeln, um die Möglichkeiten für unvorhergesehen Abstürze und Softwareversagen weiter zu verringern. Obwohl dieses Problem jedem Softwareentwickler bekannt ist, wird nach außen hin das Bild des unfehlbaren Computers propagiert.

Darüber hinaus wird der Anschein erweckt, die Informatik und mit ihr die Informationstechnik sei in der Lage, jedes Problem universell lösen zu können. Diese Vorstellung des "allgemeinem Problemlösers" basiert zum großen Teil auf der Geschwindigkeit, in der es die Informationstechnik geschafft hat, sich auf die verschiedensten Anwendungsgebiete auszubreiten. Jedoch ist die Informatik grundsätzlich auf Gebiete beschränkt, in denen mathematische Algorithmen anwendbar sind und Sachverhalte in strukturierten Datensätzen modelliert werden können. Es besteht aber kein Grund zu der Annahme, das alle Probleme, die die Informationstechnik zu lösen den Anspruch erhebt, diese Voraussetzungen erfüllen. Zum Beispiel ist im Bereich der Realisierung von Computerinterpretationen von realen Umgebungen und deren Wechselwirkungen die Aufgabe der vollständigen Modellierung dieser Welten ein offenes Problem, dem meist mit gesunder Abstraktion entgegen getreten wird. Dies führt dann natürlich zu einem abstrakten Weltbild des entsprechenden Agenten und kann folglich zu fehlgeleiteten Schlußfolgerungen führen, wenn Details die Wechselwirkungen der Welt beeinflussen, die außerhalb des Abstraktionsgrad der Agentenwelt liegen (siehe auch Qualification- / Ramnification-Problem [Russel/Norvig, 1995]).

Weiterhin gibt es Anwendungsgebiete, in denen der Einsatz der Informationstechnik grundsätzlich fragwürdig ist, auch wenn er technisch möglich erscheint. Es sind dies Anwendungsgebiete, in denen eine Entscheidung eines Problemfalls hohe Verantwortung bzw. semantische Interpretation erfordert. Dies sind Eigenschaften, die (wie später noch eingehender erläutert wird) dem Computer in maßgeblicher Hinsicht fehlen. Beispielsweise sind Expertensysteme in der Lage, unscharf definierte Aufgabenbereiche zu bearbeiten. Die Lösungen, die sie anbieten, sind jedoch, ohne, daß das von außen zu erkennen wäre, unsicher und bedürfen deshalb immer der abschließenden Beurteilung durch einen Menschen. Ein weiteres Beispiel wäre der Einsatz von vollautomatischen, heuristischen Systemen. Solche Systeme schließen, da sie auf Heuristiken basieren, falsche Ergebnisse mit ein. Diese Systeme in einen vollautomatisch ablaufenden Arbeitsprozeß einzubinden ist daher nicht zu verantworten.

In der Wirtschaft werden Informationssysteme meist zur Organisation oder Veränderung von Arbeitsprozessen eingesetzt. Die Konstruktion solcher Systeme bestimmt zum größten Teil die Beziehung, die Mensch und Maschine im Verhältnis der Arbeit zueinander einnehmen. Wenn Maschine und Mensch in einer Arbeitsbeziehung zueinander stehen, bestimmen Anteil und Initiative, die sie in dieser Beziehung an den Tag legen ihre Rolle bzgl. der Bestimmung von Arbeitssubjekt und Arbeitsobjekt, respektive aktivem und passivem Part dieser Arbeitsbeziehung. Keine Maschine kann in diesem Zusammenhang als vollständig autonom gelten, denn ihr Arbeitszweck hat immer eine reale Grundlage und ist somit um einen Menschen als Subjekt der Arbeit gruppiert. Es ist eine soziale Aufgabe der Informationstechnik, ihre

Produkte relativ zum natürlichen Subjekt des Arbeitsprozesses, dem Menschen, zu bestimmen.

Diese Grundregel wird oft zugunsten des Konzepts der "Vollautomatisierung" übergangen. Die "Vollautomatisierung" strebt eine Eliminierung des menschlichen Arbeitsfaktors an, da er als potentiell fehleranfällig betrachtet wird (im Vergleich zum Arbeitsanteil der Maschine) und somit nach Möglichkeit aus dem Gesamtprodukt eines Arbeitsprozesses herausgehalten werden muß. Resultat dieses Konzeptes ist der ausgeprägte Maschinencharakter eines Arbeitsprozesses, der die Maschine als Zentrum oder Arbeitssubjekt sieht und den Menschen zum Objekt degradiert, der gleichsam von der Maschine bearbeitet wird. Der Mensch wird in solchen Systemen der technischen und organisatorischen Logik des Systems, in dem er arbeitet, völlig unterworfen. Der Rechnereinsatz in einem solchen System hat also seine ursprüngliche unterstützende Funktion zugunsten einer dominierenden Position aufgegeben. Nebeneffekt dieses Konzepts ist die abnehmende Bereitschaft eines Menschen, bei der Arbeit an einem solchen System Bereitschaft zu eigenverantwortlichem Handeln zu übernehmen, wenn diese Möglichkeit überhaupt gegeben sein sollte. Der "Vollautomatisierung" entgegengesetzt befindet sich ein Ansatz, der den Menschen verstärkt mit seinen von der Maschine einzigartigen Fähigkeiten, insbesondere Kreativität oder Improvisation berücksichtigt und der Maschine die Rolle eines Werkzeugs zuordnet.

### **2.1. Verantwortungsdilemmata bei der Produktgestaltung.**

Damit die Informationstechnik ihre Produkte im Sinne eines sozial verträglichen Gesamtbildes definieren kann, muß sie sich schon in der Designphase eines Produktes zwei Anforderungen stellen. Zum Einen muß das zu entwickelnde Produkt einen ausgeprägten "Werkzeugcharakter" besitzen, zum anderen muß sein Einsatz "soziale Zweckbestimmtheit" [Rödiger et al., 1989] aufweisen.

Der Werkzeugcharakter bestimmt das Maß, indem ein Softwareprodukt die Fähigkeiten des Benutzers einschränkt oder im Gegensatz dazu fördert. Der Werkzeugcharakter bestimmt somit die Handhabbarkeit eines Systems. Der Begriff der sozialen Zweckbestimmtheit fragt im Gegensatz dazu, ob ein System zu einem sozial akzeptablem Zweck eingesetzt wird. Dies begründet sich auf der Vielseitigkeit von Informationssystemen, die dadurch auch mißbräuchlich benutzt werden können. Es ist beim Design entsprechender Systeme darauf zu achten, der Anwendung Grenzen zu setzen, wenn sie schützenswerte persönliche oder kulturelle Bereiche berührt.

Beide Leitbilder können letztlich in Konflikt zueinander geraten. Die Verwirklichung beider Prämissen hängt nicht nur von der zugrundeliegenden Technik, sondern auch von der persönlichen Haltung aller am Designprozeß beteiligten Personen ab.

Die Vielseitigkeit von Informationssystemen steht der Konstruktion sozial zweckbestimmter Systeme entgegen, da ein System, das viele zweckoffene Zugriffe erlaubt zum Mißbrauch verleitet. Dieses Problem ist um so stärker, je näher der Zugriff dem Kern der Maschine ist. Maschinennahe Kommandos werden von jeder Anwendung benötigt, sind also in der Regel jedem zugänglich und in ihrer Wirkung uneingeschränkt. Eine Verwirklichung von sozial zweckbestimmten Systemen ist deshalb in aller Regel nur auf einer höheren, abstrakteren Programmebene möglich. Zur Unterstützung dieses Aspektes sind politische und soziale Regelungen notwendig. Die Einhaltung dieser Regeln bedarf aber ebenso einer gemeinschaftlichen

Übereinkunft und verantwortlichen Haltung jedes Einzelnen, da informationstechnische Systeme prinzipiell die Umgehung solcher Regelungen zulassen.

Der Werkzeugcharakter eines Informationssystems ist nur dann gegeben, wenn bei der Entwicklung der Benutzer mit seinen Fähigkeiten im Mittelpunkt steht. Häufig wird der Endbenutzer jedoch nur als eine dem Computer ähnliche informationsverarbeitende Maschine betrachtet. Ein Entwicklungsprozeß, der diesem Menschenbild entgegensteht, muß verschiedene, dem Menschen eigene Aspekte beachten. Die Körperlichkeit des Menschen verlangt, wenn immer möglich, eine Bindung von informationstechnischen Mitteln an sinnlich erfahrbare Praxis. Das ganzheitlich-intuitive Vermögen des Menschen trachtet nach der Möglichkeit, im Arbeitsprozeß eigenständig Ziele bilden und verfolgen zu können. Außerdem will der Benutzer sich eine Übersicht über seinen Arbeitsprozeß und der Wirkung seiner Arbeit im Gesamtgefüge eines Arbeitsvorgangs, verschaffen. Um das zu erreichen, muß das System dem Benutzer Handlungsfreiräume geben, und zwar gleichermaßen in seiner Konfigurierung (adaptive Systeme) und letztendlich seiner Benutzung.

## **2.2. Verantwortungsdilemmata unter konkreten Arbeitsbedingungen**

Damit in der Entwicklung von Informationssystemen die Leitlinien der sozialen Zweckbestimmtheit und des Werkzeugcharakters erfüllt werden können, müssen bestimmte Bedingungen gegeben sein. Moderne Entwicklungstätigkeit ist aufgrund der hohen Komplexität der Entwicklungsprojekte in der Regel Teamarbeit, an der eine große Menge von Entwicklern, meist in der Form eigenständiger und größtenteils unabhängiger Teams arbeiten. Damit die genannten Leitlinien verwirklicht werden können, muß eine transparente Projektorganisation gegeben sein. Auch muß die persönliche Haltung der beteiligten Entwickler auf das Erfüllen der Leitlinien vorhanden sein, denn oftmals bestimmt das Design der Benutzerführung z.B. zum großen Teil den Werkzeugcharakter des Gesamtprodukts, unabhängig davon, wie handhabbar die übrigen Module des Produkts auch sein mögen.

Um die gesteckten Ziele zu erreichen empfiehlt sich schon in der Entwicklungsphase ein ausgiebiges Prototyping, vorzugsweise mit dem Endbenutzer, um die Gebrauchsfähigkeit einer Rohversion in ihren verschiedenen Details zu evaluieren. Anstelle des Prototyping sollten die Entwickler auch angehalten werden, ihre Perspektive in die des Benutzers zu verändern, um ein Gefühl für die Bedürfnisse der Menschen zu bekommen, die das Endprodukt schließlich einzusetzen haben.

Damit diese Bedürfnisse von den Entwicklern besser berücksichtigt werden, müssen sie zuerst verstanden werden, das heißt sie müssen im Gespräch mit dem zukünftigen Endbenutzern ergründet werden. Dies fordert von der Softwareentwicklern neue, möglicherweise ungewohnte kommunikative Fähigkeiten, um das herrschende Mißtrauen, das auf beiden Seiten zuerst vorherrscht, zu überwinden. Die Entwickler fühlen sich in Rollen gedrängt, die wenig mit Informatik zu tun haben, die Endbenutzer füllen sich in der Regel durch den bevorstehenden Einsatz des neuen Informationssystems überrollt oder fürchten, durch die Einführung ihre Kompetenzen ganz oder teilweise zu verlieren.

Desweiteren sind die Entwickler gefordert, im Umgang mit dem Projektmanagement den Rahmen des Machbaren festzustecken und nicht zu verlassen, damit Diskrepanzen

in den Erwartungen des Management und den Fähigkeiten des Produkts nicht entstehen können.

### **3. Folgen für Benutzer sowie für Produkte**

Die Folgen der Einführung und Benutzung von Informationssystemen sind vielfältig und betreffen die Benutzer sowie auch die Produkte, an die von ihrem Umfeld gewisse Erwartungen gestellt werden.

Die bereits erwähnte Hinwendung zu Systemen, die den Menschen als vermeintlich fehleranfälligeres System auszuschließen trachten, steht insofern auf schwachem Fundament, da die Systeme, die ihre Stellen übernehmen, selbst nicht als absolut fehlerfrei betrachtet werden können. Dies hat ihre Ursache zum großen Teil in der Arbeitsteilung beim Entwicklungsvorgang und wird durch das Fehlen zuverlässiger Testumgebungen und Korrektheitsbeweisen verstärkt.

Informationssysteme werden, manchmal ungewollt oder indirekt, zur Kontrolle und Überwachung von Menschen bzw. Mitarbeitern eingesetzt. Obwohl diese Systeme im Zuge wirtschaftlicher Erfordernisse eingeführt werden, kann ihr Einsatz in gegenteiligen Effekten resultieren, wenn der Grad der Überwachung und Disziplinierung zu schwer auf der betreffenden Personengruppe lastet.

Das kulturelle Umfeld des Menschen wird durch die Auseinandersetzung mit seiner äußeren Realität bestimmt. Die Realität, wie sie durch die Maschine vermittelt wird, ist einer Vielzahl ihrer sensorischen Ausdrucksarten beraubt. Dadurch wird die vermittelte Realität formal, gleichgeschaltet. Ein Rechner stellt keine sozialen Ansprüche an seinen Arbeitspartner, hat keine eigenen sozialen Bedürfnisse. Der Umgang mit dem Rechner vermittelt deshalb ein falsches Bild von der Realität, das Denken und Handeln wird auf negative Art beeinflusst und kann auf Dauer zu sozialen Defiziten führen, da die Beziehungen zwischen Handeln und den damit verbundenen Folgen verloren geht.

Die Einführung computergestützter Informationssysteme in ein Unternehmen hat darüber hinaus weitreichende Folgen auch in Bereichen, die zunächst nicht von der Umstellung betroffen scheinen. So kann sich die Unternehmensdarstellung nach außen (Image) ändern, Kunden und Geschäftspartner betreffend. Diese Änderungen sind nicht zwangsläufig positiver Natur, sondern können durchaus auch negativen Charakter haben, beispielsweise durch die Einführung von elektronischen Selbstbedienungsgeräten anstelle von menschlichem Verkäufer-Kunden Kontakten. Die Einführung eines Informationssystems wird oft unter dem Vorwand betrieben, kleine, klar definierte Teilbereiche eines Unternehmens zu verbessern. Als Kriterium für diese Verbesserung werden aber in der Regel nur quantitativ erfassbare Werte angeführt, da sie klar zu definieren sind. Andere, nicht erfassbare, aber nichtsdestotrotz ebenso wichtige Kriterien wie Mitarbeitermotivation, Kundenbeziehungen, etc. bleiben dabei außen vor, obwohl auch sie von einer solchen Umstellung betroffen sind.

#### **3.1. Ansatzpunkte für verantwortliches Handeln**

Die Entwicklung von Softwaresystemen unter den Aspekten der sozialen Zweckbestimmtheit und des Werkzeugcharakters erfordert von den verantwortlichen Personen eine Erweiterung ihrer Perspektive, weg von der Konzentration auf nur technische Belange hin zu den sozialen Folgen und Anforderungen ihrer Arbeit.

Fachfremde Kritik muß ernst genommen werden und darf nicht als technikfremd und nicht realisierbar abgetan werden, soll das Unbehagen des Benutzers, der gezwungenermaßen an einem Informationssystem arbeitet, nicht überhand nehmen.

Die Folgen der Verbreitung der Informationstechnik in der allgemeinen Gesellschaft fordern einen Konsens aller Beteiligten, um Ziele, Folgen und Wirkungen zu klären. Wenn dieser Diskurs nicht ausreicht, müssen politische und gesetzliche Regelungen getroffen werden, um den ungezügelten technischen Fortschritt einzudämmen.

Abgesehen von der Wirkung von Informationssystemen auf die Gesellschaft, muß auch die Wechselbeziehung von einzelnen Repräsentanten dieser Gruppen, von Computern oder Agenten und den mit ihnen oder an ihnen arbeitenden Menschen, den Benutzern oder Entwicklern, beleuchtet werden.

#### **4. Mensch und Maschine**

Wie schon oben erläutert, wird das Umfeld des Benutzers zum großen Teil von der Maschine bestimmt bzw. begrenzt, mit der er arbeitet. Diese Grenze muß nicht zwangsläufig einengend sein. Ein System mit ausgeprägtem Werkzeugcharakter wird im Gegenteil die Möglichkeiten eines Benutzers in großem Maße erweitern. Letztendlich setzt aber die Maschine die Grenze dieses menschlichen Handlungsspielraums. Da sie darüber hinaus auch einen Teil der Realität des Benutzers widerspiegelt, hat sie Einfluß auf die Sichtweise und Intuition des Benutzers und prägt sie in nicht geringfügigem Maße.

Von einer Maschine (in diesem Sinne nicht zu verwechseln mit einem Computersystem oder intelligentem Agentensystem), wird in aller Regel erwartet, das sie zuverlässig arbeitet und ihre Arbeitsergebnisse voraussehbar sind. Individuelle, subjektive oder gar irrationale Handlungen werden ihr nicht zugebilligt, um unproduktive Arbeitsergebnisse zu vermeiden. Im Gegensatz dazu bleibt dem weniger verlässlichen Menschen in dieser Arbeitsbeziehung der nicht automatisierbare Rest der Arbeit, die Kreativität und Improvisation, aus der unter anderem der Fortschritt erwächst.

In diesem Arbeitsverhältnis, in dem Benutzer und Maschine in Konkurrenz zueinander stehen, wird die Maschine vorschnell als Vorbild herangezogen, wenn ein Fehler geschehen ist, und der Benutzer als Fehlerquelle identifiziert ist, wobei die Verantwortung der Maschine meist nicht betrachtet wird.

Mit der Teilung der Arbeit wird auch die Handlungsverantwortung zwischen Benutzer und Maschine geteilt. Man muß aber bei der Bewertung dieser Verantwortung berücksichtigen, daß die Maschine an sich keine Verantwortung tragen kann, da sie kein eigenes Bewußtsein kennt, was aber eine wichtige Voraussetzung für den allgemeinen Verantwortungsbegriff ist. Dieser allgemeine Verantwortungsbegriff gründet darauf, das ein Individuum, um volle Verantwortlichkeit seiner Handlung zu erreichen, auf sich selbst reflektierende Bewußtseinsprozesse kennen muß, sich also sich selbst und seiner Umwelt voll bewußt ist und das es seine Handlungen willentlich wählen und durchführen kann.

Da der Maschine (Computer, Agent,..) die erste Voraussetzung fehlt, spricht man in ihrem Falle von "Deliberationsverantwortung" [Lenk, 1989]: Der Maschine wurde von außerhalb die Vollmacht gegeben, in bestimmten Fällen Entscheidungen "verantwortlich" zu treffen. Diese Überantwortung von Entscheidungsvollmachten werden in der Regel von ihrem Konstrukteur vorgenommen und beruhen auf seiner

allgemeinen Verantwortung, zu entscheiden, ob seine Maschine den Benutzer in einer Entscheidungsfrage, aufgrund ihrer Fähigkeiten, ständig übertreffen kann.

Wenn also ein Benutzer einer ständig wachsenden Zahl von Maschinen gegenüber steht, stellt sich die Frage, in welchem Maße er noch für seine Handlungen zur Verantwortung gezogen werden kann, oder inwieweit die entsprechenden Maschinen bzw. die dahinter stehenden Konstrukteure Verantwortung tragen.

#### **4.1. Mensch und Computer**

Was in diesem Sinne für Maschinen gesagt wurde, gilt auch für Computer. Außerdem muß in Betracht gezogen werden, daß der Computer in seiner Auslegung und Arbeitsweise dem menschlichen Denken sehr ähnelt. Der Computer steht mit dem Benutzer über allgemeine Sprachakte (= vereinbarte Formen der Kommunikation) in Verbindung oder kann auch als Übermittler von Sprachakten zwischen Benutzern dienen. Der Computer ist in diesem Sinne selbst ein Sprachakt, und zwar von seinem Konstrukteur an seinen jeweiligen Benutzer. Diese Kommunikation ist aber in der Regel einseitig und muß deshalb in besonderem Maße der vereinbarten Kommunikationsform entsprechen.

Beispielsweise beinhaltet das Icon einer Textdatei, die ein Softwareentwickler durch sein Programm einem Benutzer anzeigen läßt, die unterschwellige Vereinbarung, das diese Datei in der Lage ist, Text zu speichern, editiert, gelöscht zu werden, usw. Diese allgemeine Aussage ist bedeutungslos, wenn der Programmierer sich nicht an diese Vereinbarung hält.

#### **4.2. Mensch und Agent**

Spitze der derzeitigen Maschinenentwicklung ist der sogenannte intelligente Agent. Ihm liegt ein Leitbild zugrunde, daß von dem System verlangt, "situationsbezogen und unabhängig seine Ziele zu bestimmen und zu verfolgen" [Blandford, 1993].

Derart unabhängige Systeme zu konstruieren ist strenggenommen nicht möglich, da ein Computersystem nicht in der Lage ist, die Semantik seiner Umgebung zu erkennen. Deshalb ist es ihm auch nicht möglich, eigene Ziele und Werte zu bilden, sondern ist immer auf Werte, Ziele und Motivationen angewiesen, die letztendlich als von fremder Seite induziert erkannt werden müssen. Dies wird in der Regel von dem Designer übernommen, der den Agent konstruiert hat. Er legt dem Agenten, oftmals unbewußt, seine eigenen Ziele und Wertvorstellungen zugrunde und läßt das Agentensystem danach handeln. Für den Fall, daß dieses grundlegende Wertsystem vom Benutzer manipulierbar und anpassbar ist, kann nicht mehr von einem unabhängigen Agenten gesprochen werden.

Der Mensch ist einer in ihrer Komplexität immer stärker wachsenden Umwelt ausgesetzt, in der er sich ohne die Hilfe von Computersystemen nicht mehr zurecht finden kann. Es stellt sich daher die Frage, ob er in der Lage ist, sich gegen die Entscheidungen seiner intelligenten Assistentensysteme zu behaupten und seine Freiheit zu behalten, oder ob er vor ihnen aufgrund ihrer umfassenden Verfügung über Fakten und Sachverhalte und ihrer Verarbeitungskapazität kapitulieren muß, so daß diese ihre Ziele, Werte und Überzeugungen durchsetzen können. Wertvorstellungen, die wiederum nicht ihre eigenen sind, sondern auf ihren Konstrukteuren basieren.

## **5. Ethik von Agenten**

Die Entscheidungen und Handlungen des Menschen orientieren sich an seinen moralischen und ethischen Ansichten und Wertvorstellungen. Diese Ansichten haben vielfältige Ursachen, die religiös, sozial oder rational begründet sein können. Wenn eine Gemeinschaft eine Entscheidung treffen will, ist es nötig, daß sich die einzelnen Individuen in einem Diskurs über ihre Wertmaßstäbe und Ansichten auseinandersetzen und dann auf Grund dieses Diskurses eine gemeinsame Entscheidung oder zu vertretende Meinung bestimmen. Grundvoraussetzung für diesen Angleichungsprozeß ist, daß die beteiligten Personen ihre Meinungen ausnahmslos offen darlegen und begründen.

Wenn man sich in der Argumentation des Diskurses eines Agenten bedient, besteht die Gefahr, daß Wertvorstellungen und Vorurteile auf unterschwellige Weise in den Diskurs eintreten, die von keinem der daran Beteiligten vertreten werden, weil sie keinem bewußt sind und trotzdem Einfluß auf das Gesamtergebnis nehmen können. Es sind dies die Wertvorstellungen, die dem Agenten als dessen eigene Ziele bekannt sind und von seinem Konstrukteur herrühren, der in aller Regel am betreffenden Diskurs nicht beteiligt ist.

Der ethische Diskurs erhält dadurch einen Teilnehmer, der die Regeln des Diskurs umgeht. Es ist daher unumgänglich, daß die Teilnahme von Agentensystemen am ethischen Diskurs immer von den menschlichen Teilnehmern kritisch hinterfragt und keinesfalls leichtfertig hingenommen werden darf. Insbesondere befreit der Verweis auf das Ergebnis eines Agentensystems niemanden von der Begründungspflicht seiner Ansichten.

Computer nehmen in immer größerem Maße nicht nur Handlungsdurchführungen vor sondern werden auch in der Planung und Entscheidungsfindung eingesetzt. Sie wählen Fakten aus und handeln eigenständig nach Kriterien, die kritisch hinterfragt werden müssen, wenn das Ergebnis vom Menschen akzeptiert werden soll.

## **6. Diskussion**

Rödiger vertritt mit seiner Argumentation bzgl. der Eigenschaften von Werkzeugcharakter und sozialer Zweckbestimmtheit eines Computersystems bekannte, aber auch höchst unscharfe Begriffe.

Es ist unzweifelhaft, daß ein gut ausgebildeter Werkzeugcharakter eines Systems grundlegend ist für eine gute Beziehung und Akzeptanz der Informationstechnik in der Gesellschaft. Viele seiner Forderungen bzgl. der Entwicklung von Softwaresystemen sind aufgegriffen worden und werden im Software-Engineering praktisch umgesetzt. Die Kundennähe, die enge Beschreibung des zu programmierenden Systems, die Umschulung der Benutzer auf ein Softwareprodukt sowie der andauernde Support des Kunden sind fest verankerte Begriffe des Software-Engineerings und des Prozesses der Softwareentwicklung. Es ist bei der Entwicklung ein großer Schwerpunkt, die Ergonomie eines Softwaresystems auf seine zukünftigen Benutzer zu optimieren. Handlungsfreiräume von Mitarbeitern, Kommunikationsmöglichkeiten, Übersicht und Motivationsschaffung sind von Unternehmensmanagern längst als die Produktivität steigernde Faktoren anerkannt und werden in neu einzuführenden Softwaresystemen entsprechend unterstützt und gefordert, weil bekannt ist, daß diese Systeme einen großen Einfluß auf die mit ihnen arbeitenden Angestellten haben.

Sein Begriff der sozialen Zweckbestimmtheit ist jedoch, so wünschenswert er in seiner Einfachheit ist, so abstrakt und allgemein, das er nie vollständig erlangt werden kann. Was ein sozial akzeptabler Gebrauch einer Software (eines Systems,..) ist, ist ein höchst relativer Begriff und hängt sehr stark vom Standpunkt des jeweiligen Betrachters ab. Ein Verfassungsschützer wird in hohem Maße an der Möglichkeit einer neuen Software zum Ausspähen von vertraulichen oder persönlichen Daten interessiert sein, für ihn ist diese Tätigkeit vielleicht durch für ihn vertretbare Gründe gerechtfertigt. Für den Bürger, der durch diese Tätigkeit in Mitleidenschaft gezogen wird, ist eine solche Software im Gegenzug moralisch nicht zu vertreten.

Als Begründung für viele seiner Ausführungen bemüht er weiterhin die Ansicht, die Informationstechnik und die ihr zugrundeliegende Informatik protegiere, ob bewußt oder unbewußt, das Image des "Allheilmittels" bzgl. der Lösung von Alltagsproblemen, bzw. tue nichts, um dieser landläufigen Meinung der Gesellschaft entgegen zu wirken. Auch das mag bei Veröffentlichung seines Artikels 1989 vielleicht so gewesen sein, mittlerweile jedoch hat die Informatik meiner Ansicht nach längst diesen Nimbus verloren und ist auch in der Gesellschaft als fehleranfällig und keineswegs allmächtig erkannt worden. Zu viele Softwareprodukte und Hardwareentwicklungen haben sich in der Zwischenzeit peinliche Fehler geleistet, die publikumswirksam breitgetreten wurden. Man könnte auch sagen, daß die Informationstechnik ein zu schnelles Tempo vorgelegt und über ihre eigenen Beine gestolpert ist.

Die von Hellbardt vertretene These, daß die Ziele und Werte eines Agentensystems (einer Maschine, etc. ) nur von ihrem Designer abhängen [Hellbardt, 1996], ist insofern richtig, daß diese Ziele in der Mehrzahl der angewendeten Systeme vom Benutzer nicht hinterfragt werden. Es ist jedoch auch hier ein Anliegen des modernen Software-Engineerings, ein Softwareprodukt für den Kunden bis auf die Grundlagen durchschaubar zu machen. Es ist oftmals auch eine Forderung des Kunden, ein fragliches Softwaresystem bis auf die Grundlagen erklärt und dokumentiert zu bekommen. In dem Fall, in dem das Wissen des Kunden für fachspezifische Software in den Entwicklungsprozeß mit einbezogen und der Entwicklung zugrunde gelegt wird, wie das im Entwicklungsprozeß von Großprojekten idealerweise der Fall ist, kann nicht mehr der Designer allein als der Verursacher der "..beliefs and goals" [Blandford, 1993] des Agentensystems gelten.

Ausgehend davon ist auch die Frage der Verantwortungsverteilung der Handlung von Benutzer und Maschine (bzw. ihres Designers) fragwürdig. Diese Verteilungsfrage kennt viele mögliche Lösungsmöglichkeiten und alle sind letzten Endes diskutierbar. Es wäre z.B. möglich, die Verantwortung für die Nutzung eines Systems ganz beim Anwender zu belassen. Dies wird bei einer Vielzahl von Produkten bereits getan, indem der Käufer eines Programms sich mit dem Kauf und dem Einsatz verpflichtet, jedwede Verantwortung zu tragen, die im Zusammenhang mit der Nutzung entstehen kann. Diese Sichtweise entbindet den Entwickler von jeder Verantwortung und ist insofern problematisch, da der Anwender für Vorgänge haftbar gemacht werden kann, von denen er keine Kenntnis hat, für deren Verantwortung ihm eigentlich die "willentliche und freie Entscheidung" [Lenk, 1989] im Sinne der allgemeinen und bewußten Verantwortungsdefinition fehlt.

Ein weiterer Ansatz wäre es, dem Anwender die Verantwortung für von ihm willentlich initiierte Handlungsweisen seines Systems zu geben und dem Designer die

Verantwortung für das Handeln zuzuweisen, das sein System unabhängig vom Benutzer vornimmt. Diese Frage hat naturgemäß auch ethische, philosophische und rechtliche Aspekte, aber abgesehen davon ist es grundsätzlich eine unscharfe Grauzone, wo das Wissen des Benutzers um die Handlungsweise des von ihm benutzten Agentensystems endet, und wo folglich die Verantwortung des Agentendesigns und des Designers beginnen soll, denn verschiedene Benutzer haben verschiedene Voraussetzungen.

Diese Frage bedarf zur Lösung deshalb einer öffentlichen Diskussion unter Berücksichtigung der rechtlichen Fakten, technischen Gegebenheiten und sozialen Zustände der Gesellschaft, für die sie entschieden werden soll.

## **7. Literatur**

Blandford, A. E. , (1993). An agent-theoretic approach to computer participation in dialogue, *Int. J. Man-Machine Studies*, 39:965-998.

Hellbardt, G., (1996). Die Ethik von Agenten, *Informatik-Spektrum*, 19:87-90.

Lenk, H., (1989). Können Informationssysteme moralisch verantwortlich sein? *Informatik-Spektrum*, 12:248-255.

Rödiger, K.H. et al., (1989). Informatik und Verantwortung, *Informatik-Spektrum*, 12:281-289.

Russell, S. and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice Hall International.