

Seinsschichten, künstliche Intelligenz und Moral von Béla Pokol

Inhalt

1. Theoretische Rekonstruktionen von Schichtung der Realität:

Die Analysen von Maturana, Luhmann und Nicolai Hartmann

1. Humberto Maturanas strukturelle Verbindungslösung
2. Niklas Luhmanns Analysen über die Systemebenen der Welt
3. Nicolai Hartmanns Seinssichten

Literature

2. Künstliche Intelligenz: Die Entstehung einer neuen Seinsschicht?

(KI – im Spiegel von Nicolai Hartmanns Ontologie)

1. Der Mensch und die Hierarchie der Seinsschichten
2. Die zunehmende Verwebung der künstlichen Intelligenz in die geistige Seinsschicht
3. Die direkte Verbindung der künstlichen Intelligenz mit der physischen Seinsschicht
4. Die selbstorganisierende künstliche Intelligenz
 - 4.1. Die starke künstliche Intelligenz
 - 4.2. Die digitale Existenz des emulierten menschlichen Gehirns
 - 4.3. Die Fragen der genetisch verbesserten Superintelligenz
 - 4.4. Eine neue zusätzliche Seinsschicht oder der Neubeginn der irdischen Evolution auf der Grundlage der selbstorganisierenden künstlichen Intelligenz?

Literature

3. Die Seinsschichten und die Fragen der Roboterethik

1. Die Vorfragen der Roboterethik
2. Operative Moral, funktionale Moral und vollkommene Moral
3. Die Abwertung der physikalisch-biologischen Umwelt als moralisches Problem?
4. Moralische Dilemmata und Verantwortlichkeiten in hybriden und vernetzten Systemen
5. Selbstlernen, maschinelles Lernen und Verantwortung
6. Identität in der Welt der künstlichen Intelligenz
7. Das moralische Credo von „Unabomber“, dem „verrückten Mathematiker“

Literature

4. Eine Kritik der Verengung der Moral auf Gerechtigkeit

1. Kollektivmoral und persönliche Moral
 - 1.1. Rudolf von Jherings Theorie über Moral und Sitte
 - 1.2. Nicolai Hartmanns Ethik
2. Moralsoziologische Untersuchung
3. Die Annulierung der Kollektivmoral: die kritische Moral
 - 3.1. Kants Moralphilosophie
 - 3.2. Der Weg der Verengung der Moral auf Gerechtigkeit

Literature

1

Theoretische Rekonstruktionen von Schichtung der Realität: Die Analysen von Maturana, Luhmann und Nicolai Hartmann

Die Fragen der physischen, biologischen und psychischen Existenz unterhalb der Ebene der sozialen Existenz sind normalerweise nicht wichtig für die Beantwortung gesellschaftstheoretischer Probleme. Obwohl ein fundiertes Verständnis der moralischen Anforderungen, die für die Existenz des Menschen und der menschlichen Gemeinschaften und der daraus resultierenden Normen und Werte wesentlich sind, ohne Berücksichtigung der biologischen und psychologischen Begabung des Menschen nicht möglich ist, hängt die Funktionsweise der meisten gesellschaftlichen Institutionen nicht so sehr von ihnen ab. Auf diese Weise reicht eine auf das gesellschaftliche Zusammenhänge beschränkte Analyse meistens aus. Neue Entwicklungen im Bereich der künstlichen Intelligenz und das Aufwerfen der Möglichkeit der Menschenrechte und möglicherweise der echten Privatrechte im Fall der immer fortschrittlicheren Robotern haben es für mich interessant gemacht, ob es im Fall einer künstlichen Intelligenz, der mit der physischen Existenz eines Roboterkörpers verbunden ist, wirklich möglich wäre, sie vollständig in die Gesellschaft einzubeziehen? In ihrem Fall gibt es weder eine biologische noch eine psychische Seinsschicht, und im Vergleich zu der vierschichtigen (physischen, biologischen, psychischen und geistlichen) natürlichen Person und der menschlichen Gesellschaft wird es fraglich, wie die Erweiterbarkeit etablierter, gesellschaftlichen Institutionen auf sie angewendet werden kann.¹

Um dies zu beantworten, wurde es notwendig, zusätzlich zur Theorie der sozialen Realität die Theorien zu überprüfen, die bereits die Abbildung der Ebenen der gesamten umfassenden Existenz oder mit anderen Worten die Seinsschichten der gesamten Realität durchgeführt haben. Die physische Existenz als Voraussetzung für biologische Existenz und die Unentbehrlichkeit beider für die menschliche Gesellschaft kann als Evidenz besichtigt werden. Aber wie diese Realitätsebenen (und vielleicht sogar noch verborgenere) miteinander in Beziehung stehen und wie ihre Auswirkungen aufeinander auf theoretischer Ebene artikuliert werden können, gibt es unterschiedliche Ansätze. Im Folgenden werde ich zunächst die Lösungen des chilenischen theoretischen Biologen *Humberto Maturana* untersuchen² und anschließend die Modifikation durch den deutschen Soziologen *Niklas Luhmann* analysieren. Diese Ansätze vergleiche ich dann mit der ontologischen Schichtenlehre des deutschen Philosophen *Nicolai Hartmann*.

¹ Einzelheiten zu diesen Fragen siehe Béla Pokol, Künstliche Intelligenz: Die Entstehung einer neuen Seinsschicht? (KI - im Spiegel von Nicolai Hartmanns Ontologie) Pázmány Law Working Papers, Nr. 2018/12.) und Béla Pokol, Die Seinsschichten und die Fragen der Roboterethik. C: /Users/Dell/Downloads/academia.edu/Die_Seinsschichten_und_die_Fragen_der_Ro.pdf

² Um Maturanas Theorie zu veranschaulichen, verwende ich sein gemeinsames Buch mit seinem ehemaligen Schüler Francisco J. Varela, aber da diese Theorie Jahre zuvor allein von Maturana erstellt wurde, nehme ich dies trotz aller Ergänzungen von Varela als seine Theorie. Siehe Maturana / Varela, *The Tree of Knowledge. The biological Roots of Human Understanding*. Revised Edition. Shambhala-Publications. Boston, Massachusetts. 1987.

1. Humberto Maturanas strukturelle Verbindungslösung

Der Ausgangspunkt für die Theorie des Biologen Maturana ist die geschlossene Selbstreproduktion von Zellen in lebenden Organismen, um die systemische Natur komplexerer biologischer Systeme und der daraus aufbauenden tierischen und menschlichen Gesellschaften zu verstehen. Die Selbstreproduktion der inneren Molekülstruktur von Zellen durch konstante chemische Prozesse wird vollständig durch die innere Struktur der gegebenen Zellen bestimmt, und sie absorbiert nur die notwendigen Chemikalien aus ihrer Umwelt. Als sich in der Evolution durch die permanente Zusammenschließung der Zellen die Mehrzellige zu bilden begannen, änderte dies nichts an der geschlossenen Selbstreproduktion jeder inneren Zelle. Jede Zelle in dieser Zusammenschließung ist nur als Ganzes miteinander verbunden, und die internen Prozesse jeder Zelle können durch die anderen Zellen und ihre Selbstreproduktion nur als Umwelt gereizt werden. Durch Reizung kann dann das gereizte Moment der kettenartigen inneren Reproduktion der anderen Zelle transformiert werden, und diese Änderung kann die innere Struktur dieser Zelle verändern, und dann wird die Zelle als Ganzes etwas auch transformiert. Diese Irritation wirkt aber auch nur insoweit, als die innere Struktur der Zelle dies zulässt. Ist dies nicht der Fall, kann durch die Reizung die Zelle getötet werden, aber sie kann den äußeren Effekt nicht als Reizung identifizieren und somit die Zelle nicht zur Veränderung zwingen.

Maturana nennt diese systemische Schließung und die Selbstreproduktion der Zellen operative Schließung (*operative closure*) sowie die Verbindung von Zellen aus mehrzelligen biologischen Organismen als strukturelle Kopplung (*structural coupling*). Das komplexere mehrzellige biologische System führt somit gleichzeitig eine doppelte Selbstreproduktion durch, wobei die ständige Erneuerung einzelner Zellen und die Erneuerung des aus zusammengeschlossenen Zellen bestehenden Organismus selbst in seiner Umwelt erfolgt. Die strukturelle Verbindung erzeugt aufgrund der gegenseitigen Irritationen der Zellen ihre wechselseitige (aufeinander abgestimmte) Veränderung innerhalb dem komplexeren biologischen System, aber der Organismus selbst reagiert auf die Reizung seiner Umgebung nur aufgrund seiner inneren Bestimmung. Dies bedeutet, dass die Umwelt den Organismus nicht direkt beeinflussen kann, sondern nur in dem Maße, in dem ihre innere Struktur spezifiziert, was als Reizung durch die Umwelt wahrgenommen werden kann, und dann eine der Komponenten des Organismus und damit die Struktur des Organismus etwas verändert. Wenn dies nicht der Fall ist, kann es die Reaktionen des Organismus zerstören, aber nicht ändern. Maturana nannte diese Theorie der geschlossenen Selbstreproduktion von Zellen und Organismen *die autopoetische Systemtheorie* im Gegensatz zur vorherigen Systemtheorie, von der die Änderung des Systems größtenteils durch die Systemumgebung erklärt wurde.

Maturana weist darauf hin, dass komplexere Mehrzellige mit Veränderungen auf die Irritationen ihrer Umgebung reagieren können, und diese Fähigkeit zur Veränderung wird radikal erhöht, wenn ihre Neuronen in ein zentrales Nervensystem übergehen können. Dies liegt daran, dass die Oberflächen des Organismus, von denen die Auswirkungen der Umwelt wahrgenommen werden, sich mit den Bewegungsoberflächen (Muskeln) des Organismus verbinden und dadurch viel mehr Änderungen (möglicherweise Ortwechsel) an den Reizungen vornehmen kann. Auf diese Weise kann der Organismus trotz seiner inneren Bestimmtheit viel mehr Irritationen wahrnehmen und darauf mit seiner Strukturänderung reagieren.

Nach der Analyse der Zellen und der daraus aufgebauten Mehrzelligen als *quadratische Einheiten (second-order unities)* untersucht Maturana die Koexistenz, die sich aus den dauerhaften Wechselwirkungen komplexerer mehrzelliger Organismen ergibt, die als Einheiten dritter Ordnung (*third-order unities*) dauerhafte Tiergesellschaften schaffen können. Insbesondere wenn die dauerhafte Koexistenz zu solchen wechselseitigen strukturellen Veränderungen in den miteinander verbundenen einzelnen Organismen geführt hat, dass eine Fortpflanzung der Rasse nur durch die Paarung mehrerer Individuen möglich ist. In solchen Fällen werden aufgrund der besseren Pflege des Nachwuchses, der Schutz des fürsorglichen erwachsenen Organismus und die Fütterung des Nachwuchses oder des Nachwuchses durch andere Personen das langfristige Zusammenleben erhöht, und die strukturelle Konnektivität bzw. die dafür erforderliche Kommunikationsvielfalt in dem Zusammenleben in einer höhere Ebene gesteigert wird. Diese Kommunikation ist auf viele Arten möglich, beispielsweise durch den kontinuierlichen Austausch chemischer Verbindungen zwischen Ameisen in Ameisengesellschaften oder durch spezielle Gesangsstimmen für einige Vogelarten, die Bildung reaktionsfähiger Duette und schließlich wird das sprachliche Kommunikationsfeld beim Menschen zum Bereich der sprachlichen Kommunikation.

Maturanas Analysen zeigen hier die beiden Seinsschichten der menschlichen Existenz, die biologische Seinsschicht und die kulturelle Schicht, die darauf aufbaut. Durch die Sprache kann man seine eigenen inneren Beziehungen und seine Umwelt mit unterschiedlichen sprachlichen Unterscheidungen beschreiben und so die Anpassung an die eigene Umwelt nicht nur durch seine Instinkte konservieren kann, sondern sie wird auch durch kulturelle Fixierung und Weitergabe durch sprachliche Unterscheidungen für Mitglieder neuer und neuer Generationen bewahrt: "Language enables those who operate in it to describe themselves and their circumstances through the linguistic distinctions of linguistic distinctions" (Maturana/Varela 1987:210). Während biologische Organismen und ihre Gemeinschaften nur ihre biologische Existenz kontinuierlich reproduzieren, muss die menschliche Gemeinschaft zusätzlich zu ihrer biologischen Existenz kontinuierlich ihre kulturelle Existenz pflegen und reproduzieren: "In human social systems, the case is different. As human communities these systems have operational closure too, in the structural coupling of their components. But human social systems exists also as unities for their components in the realm of language. Therefore, the identity of human social systems depends on the conservation of adaption of human beings not only as organisms (in a general sense) but also as components of their linguistic domains." (ders.)

Matura beschreibt jede Schicht der zwei Schichten der menschlichen Existenz so, dass solange auf der Ebene der biologischen Existenz die miteinander verbundenen Einheiten eng zusammengeschlossen sind und keine Unabhängigkeit für die Zellen ihrer inneren Organisation zulassen und sie sind für den lebenden Organismus, ermöglicht das menschliche Gesellschaftssystem seinen Mitgliedern jedoch ein höheres Maß an Autonomie.³ Als Ausnahme stellt er die Gesellschaft von Sparta vor, in der die Unabhängigkeit jedes Mitglieds vollständig unterdrückt wurde und als ein biologischer Organismus eine enge Vernetzung in ihrer sozialen Organisation verwirklicht wurde, aber diese Depersonalisierung führte zum Verlust der Lebensfähigkeit der ganzen Gesellschaft von Sparta.

In einer späteren Studie mit einem anderen Co-Autor zeigte Maturana auch einen weiteren Unterschied zwischen der Organisation der biologischen Ebene und der Organisation der sprachlich-kulturellen Ebene. Während das biologische System immer in der Gegenwart lebt, kann die Systeme der kulturelle Ebene durch Beobachtung und ihre sprachliche Fixierung neben der Gegenwart auch vergangene Erfahrungen und mögliche zukünftige Ereignisse

³ „The organism restricts the individual creativity of its component unities, as these unities exist for the organism. The human social system amplifies the individual creativity of its components as that system exist for these components" Matura/Varela 1987: 99).

benutzen und auch im Lichte dieser Erfahrungen können die menschlichen Handeln ausgestaltet werden: "It is only for us human beings that as we exist and operate in language, we can generate reflections and explanations, in that the past, the present, and the future, have a presence and operational values in our living as explanatory notions of our experiences." (Maturana/ de Rezepka: 2007:4).

Die zweischichtige Natur der menschlichen Existenz ermöglicht somit eine zeitliche Differenzierung durch die obere Schicht der Existenz, und somit wirft Maturanas Theorie im Prinzip neben dem fortschreitenden Prozess der biologischen Ebene der menschlichen Existenz auch den Prozess dauerhafterer Formationen der kulturellen Ebene der menschlichen Existenz auf. Trotzdem besteht Maturana darauf, dass die Reproduktion der menschlichen Welt immer auf dem momentartige Bestimmungen und Kausalitäten beruht, und er kritisiert daher das Denken, das das Leben unserer Kinder aufgrund verschiedener Ängste heute vorbestimmen will: "We human beings create the world that we live arises moment after moment in the flow our living, how can we pretend then, to specify a future that will not belong to us because it will arise in the living of our children and will not be created by us? (...) We human beings live in the present; the future and the past are manners of the being in the present." (Maturana/de Rezepka: 2007:9.) So tauch aus dieser Behauptung eine Unterschätzung in Bezug auf die kulturelle Beobachtung und gesellschaftswissenschaftliches Vorhersage sowie die gesamten kulturellen Seinsschicht auf, und die Überschätzung der Charakteristik der biologischen Seinsschicht.

Zusätzlich zu den beiden Seinsschichten sollte angemerkt werden, dass Maturana gelegentlich erwähnt, dass die physische Welt unter biologischen Bedingungen aufgebaut ist und dass das biologische System Materie aus seiner physischen Materie entnimmt: "Thus, autopoietic unities specify biological phenomenology as the phenomenology proper of those unities with features distinct from physical phenomenology. This is so, not because autopoietic unities go against any aspect of physical phenomenology – since their molecular components must fulfill all physical laws - but because the phenomena they generate in functioning as autopoietic unities depend on their organization and the way this organization comes about, and not on the physical nature of their components (which only determine their space of existence)". (Maturana/Varela 1987: 51). So ergeben sich letztendlich drei Seinsschichten aus Maturanas Theorien, aber er erwähnt das Physische lediglich als eine materialgebende Schicht (er verwendet den Begriff "spezielle Phänomenologie" für jede Seinsschicht), und er analysiert die Beziehung zwischen dem Physischen und dem Biologischen nicht mehr und deshalb versucht er auch nicht, allgemeinere Verbindungen zwischen ihnen herzustellen.

2. Niklas Luhmanns Analysen über die Systemebenen der Welt

In Niklas Luhmanns Schriften finden sich mehrere Gedankenverschiebungen, um die Seinsschichten unter der sozialen Welt thematisieren zu können. Seine Versuche können erstmals in einem Artikel von 1974, dann in seiner umfassenden Zusammenfassung der Systemtheorie von 1984 und schließlich in seinen Schriften zwischen 1987 und 1991 gefunden werden. Besonders systematisch befasst er sich in den Schriften zwischen 1987 und 1991 mit diesem Thema im Zusammenhang mit der Analyse über die Beziehung zwischen psychischem Bewusstsein und sozialer Kommunikation.

Im ersten erwähnten Material untersuchte er die Beziehung zwischen Macht und physischer Gewalt, die ihre letzte Grundlage ist, und etwas eingehender ging Luhmann in

diesem Zusammenhang auch auf die Beziehungen zwischen der sozialen Welt und ihrer physisch-biologischen Seinsschicht ein.⁴ Das Modell hierfür wurde von der Medientheorie von Talcott Parsons geliefert, von wem die körperliche Gewalt als ultimatives Mittel zur Deckung der Macht als verallgemeinertes Austauschmedium des politischen Systems verstanden wurde. Der moderne Staat mit seinem Monopol auf die Nutzung physischer Macht hat Macht nur als verallgemeinertes symbolisches Austauschmedium, sofern sie nur als letztes Mittel zur Verfügung steht, aber er muss sie normalerweise nicht verwenden. So wie Geld als Austauschmedium beruht auf abstrakte Weise darauf, dass wir es wegen der Möglichkeit, es gegen Produkte auszutauschen, akzeptieren aber dies wird normalerweise nicht bei jedem Geldeingang überprüft und wird von dem Geld das Funktionieren der Gesellschaft nur auf diese Weise harmonisch gemacht. Luhmann verallgemeinert dieses Beispiel etwas und sieht, dass es auch unter dem Austauschmedium mehrerer sozialer Teilsysteme einen Zusammenhang mit einer solchen physikalisch-organischen Seinsschicht gibt. Wie im Fall von Liebe als Austauschmedium von intimer Beziehungen, ist es die Sexualität, dies dazu dient, Liebesbeziehungen in einer biologischen Schicht des Lebens zu fixieren. Dies macht die Liebe als kommunikative Beziehung, die auf sozialer Ebene aufgebaut ist, wirklich stark und kann Familienbeziehungen, die in den modernen Gesellschaften auf Liebe basierenden Ehe beruhen, dauerhafte Stabilität verleihen. In ähnlicher Weise kann im Fall der Wissenschaft als Teilsystem der Gesellschaft von der Wahrheit als Austauschmedium die Forschungsarbeit der Wahrheitsfindung und die damit verbundenen Entscheidungen nur dann organisiert werden, wenn die *sensorische Wahrnehmung* als ultimatives Mittel zur Entdeckung die ultimative Grundlage für die Entscheidungsfindung ist. Wenn die Wahrheit in Bezug auf einer Erscheinung nicht durch eine endgültige, indirekte Wahrnehmung entschieden werden kann, wird sie aus der Perspektive der modernen Wissenschaft entfernt.

Luhmann macht in dieser Studie auch einige allgemeine Bemerkungen, die sich auch auf die Beziehung zwischen der sozialen Seinsschicht und den physikalisch-biologischen Seinsschichten auswirken. Die eine bezieht sich auf die Tatsache, dass der Ort der Verbindung zwischen den beiden Seinsschichten auf der Systemebene der sozialen Welt liegt, die hauptsächlich auf dem biologischen Aspekt von Person basiert, das heißt, auf der Ebene der Interaktionssysteme.⁵ Wenn wir jedoch wissen, dass in Luhmanns Theorie die soziale Organisation die Rolle der persönlicher Interaktionen abnimmt, und im Verlauf der sozialen Evolution zunehmend zugunsten der Bildung der dauerhaften Organisationssysteme in den Hintergrund gerückt wird, und selbst innerhalb der Interaktionsebene wird die persönliche Interaktion zunehmend durch unpersönliche schriftliche Kommunikation ersetzt, dann bedeutet die Verknüpfung der sozialen Seinsschicht mit der physisch-biologischen Seinsschichten auf der Ebene persönlicher Interaktionen auch die Verringerung dieser Bindung.

Eine weitere allgemeinere Bemerkung von Luhmann zur Beziehung zwischen den beiden Seinsschichten ist, dass körperliche Gewalt und Sexualität in sozialen Beziehungen nicht nur als physikalisch-biologische Tatsachen in ihrer Grausamkeit auftreten, sondern sie

⁴ Diese Studie mit dem Titel "Symbiotische Mechanismen" von 1974 ist Teil der "Soziologischen Aufklärung 3" und es steht mir zur Verfügung, daher werde ich daraus zitieren.

⁵ „Beim Aufbau sozialer Systeme erfüllen diese Bedingungen Interaktionssysteme, die durch relative rasch strukturiert werden können, dadurch aber nicht festgelegt sind, sondern noch definierbar bleiben. Das gilt eindeutig für gemeinsames Wahrnehmen, für Drohung mit physischer Gewalt, für Kommunikation über nur sozial bzw. nur asozial befriedigende Bedürfnisse: Sexualität bzw. Befriedigung aus knappen Ressourcen. (...) Der Bezug auf präzente Organismen erleichtert mit anderen Worten die Herstellung der reflexiven Reziprozität der Perspektiven und damit die soziale Strukturierung der Situation.“ Niklas Luhmann: Soziologische Aufklärung 3.“. Westdeutscher Verlag. 1981. 231. p.

auch durch eine symbolische Dimension vergesellschaftlicht und transformiert werden. Das heißt, auch eine Transformation dieser physikalisch-biologischen Tatsachen in der Seinsschicht der sozialen Welt durchgeführt wird: „Gewalt, Sexualität usw. haben in diesem Sinne eine symbolische, nicht nur eine physische oder organische Wirksamkeit (...) die soziale und kulturelle Entwicklung von Medien-Codes Umdisposition im Bereich der symbiotischen Mechanismen erfordern mag“ (Luhmann 1981: 231). Schließlich ist eine der allgemeinsten Bemerkungen in dieser Studie, dass selbst wenn die evolutionäre These akzeptiert werden soll, dass kulturell-symbolische Interferenzen zunehmend vergeistigte soziale Kontakte in modernen Gesellschaften schaffen, bedeutet dies nicht, dass es eine Ablösung von den Phänomenen der physikalisch-biologischen Seinsschicht gibt: „Dies Kontrastbeispiel erhärtet die Vermutung, daß selbst wenn man die These von einer kulturell-symbolischen Steuerung der gesellschaftlichen Evolution akzeptiert, der Bezug zur Sphäre physisch-organischen Zusammenlebens damit nicht bedeutungslos wird (...) Fruchtbarer wird es sein, davon auszugehen, daß sinnhafte Erlebnisverarbeitung und Kommunikation ihr physisch-organischen Substrat als Wirklichkeit nicht ignorieren.“ (Luhmann 1981:241).

Luhmanns nächste Reflexion zu diesem Thema ist in seinem Buch „Soziale Systeme“ von 1984 und hier hat er seine Thesen bereits allgemeiner in Bezug auf die Seinsschichten der Realität zum Ausdruck gebracht. Er ging von der Prämisse aus, dass seit Kant die alte ontologische Auffassung der Realität von dem modernen philosophischen Denken verworfen wurde und das bloßen *Ding an sich* da draußen in der Welt nur durch aktive Bewusstseinskonstruktionen erreicht werden kann, und obwohl Kants Lösung dieses Problem übermäßig psychologisiert, kann dies aus einer umfassenden systemtheoretischer Sicht adäquater gelöst werden. Aus dieser breiteren Perspektive folgt, dass bestimmte Elemente der Realität nicht von der Realität selbst vorbestimmt sind, sondern im Gegenteil werden sie von verschiedenen Systemen (biologische, psychische und soziale) selbst konstruiert.⁶ Auf der Grundlage dieser Ansicht stellt Luhmann daher fest, dass die alte Ansicht der Schichtung der Realität und ihr Bottom-up-Aufbau zurückgewiesen werden muss: „Es geht nicht einfach um einen geschichteten Weltaufbau, bei dem die unteren Schichten zuerst fertiggestellt sein müssen, bevor weitergebaut werden kann. Vielmehr werden die Voraussetzungen mit der Evolution höherer Ebenen der Systembildung selbst erst in eine dafür geeignete Form gebracht. Sie entstehen durch Inanspruchnahme. Deshalb ist Evolution nur durch Interpenetration, das heißt nur durch wechselseitige Ermöglichung möglich.“ (Luhmann 1984:294). Mit dieser allgemeinen Ablehnung geht Luhmann also weder detailliert auf die ontologischen Theorien der Schichtung der Realität ein, noch zitiert er in seinen Arbeiten auch nur kritisch die relevanten Autoren, einschließlich Nicolai Hartmann. Wenn wir uns dann Luhmanns Lösung bei der Analyse der Beziehung zwischen der sozialen Welt und dem Menschen nähern, sehen wir, dass er sogar zu Hartmanns Schlussfolgerungen hätte kommen können.

Luhmann lehnt den „Menschen“ als Grundeinheit der sozialen Organisation in traditionellen Sozialtheorien ab, weil er angesichts der Kritik an der humanistischen Tradition sieht, dass durch die Auffassung des Menschen als Einheit der Gesellschaft auch das ganze Gesellschaftsbild so konstruiert wird, als ob die Gesellschaft nur für die Menschen, für ihr letztes Element wäre. Dies wiederum zwingt die Sozialtheorie in einen ideologisch-normativen Rahmen und dieser Rahmen muss daher entsorgt werden, um unvoreingenommen

⁶ „ Anders als die Wortwahl und Begriffstradition es vermuten lassen, ist die Einheit eines Element (zum Beispiel einer Handlung) nicht ontisch vorgegeben. Sie wird als Einheit erst durch das System konstruiert, das ein Element als Element für Relationierungen in Anspruch nimmt. Diese De-Ontologisierung und Funktionalisierung des Elementansatzes ist in der modernen Wissenschaftsbewegung durch die Mathematisierung der Naturwissenschaften in Gang gebracht worden.“ Luhmann: Soziale Systeme. Suhrkamp, Frankfurt am Main, 1984. 42.p.

sein zu können. Daher muss das einheitliche Konzept des „Menschen“ verworfen werden, und nur die „Person“ kann als Element der sozialen Welt in das Bild eintreten, was nur als Identifikationspunkt für normative und kognitive Erwartungen für einzelnen Mensch (Person) in Formationen der sozialen Welt verstanden werden kann. Das heißt, die Person repräsentiert nur die Aspekte der sozialen Seite des Menschen und schneidet die biologischen und psychischen Aspekte des Menschen ab, und der Mensch selbst befindet sich zusammen mit seinen biologischen und psychischen Aspekten in Luhmanns Theorie außerhalb der sozialen Welt und wird als ihre Umwelt definiert. Andererseits erscheint die reduzierte „Person“ als sozialer Aspekt des Menschen nur als eine sinnhaftige Konstrukt in der Rekonstruktion der sozialen Welt: „Wir wählen den Ausdruck „Mensch“, um das festzuhalten, daß es sowohl um das psychische als auch das organische System des Menschen geht. Den Ausdruck „Person“ wollen wir in diesem Zusammenhang weitgehend vermeiden, um ihr für die Bezeichnung der sozialen Identifikation eines Komplexes von Erwartungen zu reservieren, die an einen Einzelmenschen gerichtet werden.“ (Luhmann 1985:286).

Luhmanns nächster theoretischer Konstruktionsschritt besteht darin, die biologischen und psychischen Aspekte des Menschen, die als notwendige Umwelt der sozialen Welt wahrgenommen werden, mit dem Konzept der *Penetration* zu verbinden. In der Tat bedeutet dies die Verbindung der soziale Seinschicht mit den psychischen und biologischen Seinschichten der Realität. Mit der Penetration macht eine Entität auf einer Systemebene ihre Komplexität für eine andere Entität auf einer anderen Systemebene verfügbar, ohne die sie nicht bauen und existieren könnte. (Ohne biologisches Leben könnten die Entitäten der sozialen Welt also nicht existieren.) Und wir können von gegenseitiger Durchdringung, *Interpenetration* sprechen, wenn sich die Entitäten auf beider Systemebene durch den Kontakt gegenseitig transformieren.⁷ Wenn wir uns nun Hartmanns Analysen der Aufteilung des Menschen und der menschlichen Welt in vier Seinschichten und den Fokus auf die höchste sinnhaftige Seinschicht der menschlichen Gesellschaft ansehen, hätte Luhmann das auf Menschen basierende soziale Bild, das er der wegen der Verteidigung gegen der humanistischen Verzerrung verworfen hat, auch aufgrund der Analysen von Nicolai Hartmann beiseite geschoben. Wenn er davon ausgegangen wäre, hätte er auch versuchen können, eine genauere theoretische Rekonstruktion der Gesellschaft auf der geistlichen Seinschicht der Realität aufzubauen.

Luhmann schränkt die Verbindung zwischen sozialen Systemen und psychischen Systemen größtenteils auf den beiden gemeinsamen sinnhaftigen Moment ein, und daher können die gesamten spirituell-emotionalen Prozesse des Bewusstseinssystems nur verengt erscheinen.⁸ Da die spirituell-emotionalen Prozesse aber ständig von unten mit den Instinkten der biologischen Prozesse in Kontakt stehen und sie tragen auch ihre Impulse durch Transformation weiter, so dass diese gesamte Seinschicht aus Luhmanns Perspektive herausfällt. Unter Verwendung von Maturanas Theorie der strukturellen Kopplung, bei der die Verbindung nur auf der Ebene der Elemente stattfindet und über diese Ebene ändern sich die

⁷ „Von Penetration wollen wir sprechen, wenn ein System die eigene Komplexität (und damit: Unbestimmtheit, Kontingenz und Selektionszwang) zum Aufbau einer anderen System zur Verfügung stellt. In genau diesem Sinne setzen soziale Systeme „Leben“ voraus. Interpenetration liegt entsprechend dann vor, wenn dieser Sachvehalt wechselseitig gegeben ist, wenn also beide Systeme sich wechselseitig dadurch ermöglichen, daß sie in das jeweils andere ihre vorkonstituierte Eigenkomplexität einbringen. (...) Im Falle von Interpenetration wirkt das aufnehmende System auch auf die Strukturbildung der penetrierende Systeme zurück.“ (Luhmann 1984:290.)

⁸ „Von hier aus wird besser verständlich, weshalb der Sinnbegriff theoriebautechnisch so hochrangig eingesetzt werden muß. Sinn ermöglicht die Intepenetration psychischer und sozialer Systembildungen bei Bewahrung ihrer Autopoiesis; Sinn ermöglicht das Sichverstehen und Sichfortzeugen von Bewußtsein in der Kommunikation und zugleich das Zurückrechnen der Kommunikation auf das Bewußtsein der Beteiligten.“ (Luhmann 1984:297)

Strukturen nur indirekt aufgrund der Änderung ihrer miteinander verbundenen Elemente, löst Luhmann die Verbindung von Seinsschichten wie folgt: "Es sind Differenz und Ineinandergreifen von Autopoiesis und Struktur (die eine sich kontinuierlich reproduzierend, die andere sich diskontinuierlich ändernd), die für das Zustandekommen von Interpenetrationsverhältnissen zwischen organisch/psychisch und sozialen Systemen auf beiden Seiten unerlässlich sind" (Luhmann 1984:297). Das heißt, zum Beispiel haben die beiden Systemebene ständigen Kontakt aufgrund dem Sinn als ihr gemeinsamer Baustein, und sowohl die sinnhaftige Beiträge der psychischen Systeme der Kommunikationsteilnehmer als auch die sinnhaftige Beiträge der an der Kommunikation beteiligten sozialen Teilsysteme tragen zur Weiterentwicklung dieser sozialen Systems bei und es wandelt sich in diesem Kontakt etwas auch die teilnehmenden psychischen Systeme um. Aber nur in dem Maße, wie es die Struktur des psychischen Systems erlaubt, aber andererseits wird diese Struktur selbst durch die Änderung seines Elements etwas verändert. Neben der Tatsache, dass auf diese Weise nur der sinnhaftigen Teile der psychischen Seinsschicht berücksichtigt wird und der Einfluss der biologischen Instikte auf das intellektuelle Material des psychischen Systems von unten abgeschnitten wird, kann allgemein gesagt werden, dass es in dieser Theorie kein Bottom-up Aufbau der Seinsschichten gibt, sondern nur eine gegenseitige Reizung und daraus resultierende teilweise Veränderung. Um ein konkretes Beispiel für die Wirkung dieser Verengung zu geben, können wir die Liebeskommunikation in der sozialen Welt und die daraus resultierende dauerhafte intellektuell-emotionale Beziehung erwähnen. Es ist ziemlich offensichtlich, dass in dieser Beziehung die Existenz und der Grad des biologischen Sexualtriebs bei den beiden Partnern eine bedeutende Rolle spielen, und wenn dies beispielsweise im Fall eines Partners aus irgendeinem biologischen Grund minimal ist, läuft die Liebesbeziehung im Fall dieses Partners ganz anders ab als im Fall des in der Beziehung teilnehmenden anderen Partners mit einem starken sexuellen Instinkt, und vermutlich wird dieser Unterschied bald zu einem Auseinanderbrechen der Beziehung führen. Das heißt, die Liebe wird nicht nur durch eine intellektuelle Beziehung, sondern auch durch Sexualität biologisch durchdrungen, und diese biologische Komponente wird auf die emotional-spirituelle Ebene transformiert. Und dies, kombiniert mit einer solchen biologisch-spirituellen Grundlage, schafft eine völlig andere Art von Liebeskommunikation und Partnerschaft als zwischen zwei Partnern mit gleichem biologischem sexuellem Drang. Gleiches gilt jedoch für die Analyse einer Reihe anderer sozialer Phänomene, bei denen die direkten Auswirkungen starker biologischer Reize auftreten, wie beispielsweise das Funktionieren des Lebensinstinks bei Todesgefahr, das im Soldaten im Krieg zur Aufhebung einer Reihe von moralischen Normen führt, die bisher als heilig angesehen wurden, und zur Wiederbelebung des älteren Tötungsinstinkts. Dann, nach einem Krieg, werden diese „verwüsteten“ Persönlichkeiten, die für die Tötung, Grausamkeit und sogar Wertschätzung dieser Tätigkeiten umgestaltet wurden, für eine lange Zeit fast ungeeignet für die Wiedereingliederung in das zivile Leben sein. Im Gegensatz dazu bauen solche Lebensinstikte im zivilen Leben hauptsächlich ab und andererseits die umgestalteten Instikte der Hilfsbereitschaft auf und auf der Grundlage eines solch zurückhaltenden und umgestalteten biologischen Reizes und des dadurch beeinflussten spirituell-emotionalen Lebens verlaufen auch die Prozesse der intellektuellen Kommunikation zwischen den Teilnehmern des sozialen Lebens und den daraus entstehenden sozialen Institutionen unterschiedlich. Natürlich könnten weitere Beispiele für die Koexistenz von biologischen Reizen, spirituell-emotionalem Leben und sozial-intellektuellen Kommunikationsformationen angeführt werden, aber diese beiden Beispiele waren vielleicht genug, um zu zeigen, dass Luhmann die Ebenen des biologischen, bewussten und sozialen Systems (oder in Hartmanns Worten: Seinsschichten) zu sehr auseinander zerreißt und somit kann er das Funktionieren realer Formationen nur durch Verengung und Verzerrung erfassen.

Um dieses Problem zu verstehen, läßt sich diese Art von struktureller Kopplung zwischen dem psychischen Bewusstseinssystem und dem Kommunikationssystem der sozialen Welt in Luhmanns Studien von 1987 bis 1994 zu besichtigen.⁹ Er führte am vollständigsten seine Gedanken über die Kopplung der beiden Systemebene in der Schrift „Die Autopoiesis des Bewußtseins“, die er in seinen späteren Arbeiten nicht mehr geändert hat.

Es soll herausgehoben werden, dass Luhmann seit Beginn seiner Überlegungen 1974 seine Konzeption des Systems radikal geändert hat, indem er die Idee der Autopoiese übernommen hat, und nun zu einem Prozess der zirkulären Schließung in der operativen Abfolge von Systemelementen übergegangen ist, anstatt auf Strukturen zu achten. Während 1974 mit der Betonung symbiotischer Mechanismen in Luhmanns Analysen auch die biologischen Reizen während intellektueller Kommunikationsprozesse eine Rolle hatten (z. B. der kontinuierliche Einfluss des Sexualtriebs auf die Aufrechterhaltung der Liebe), sind sie hier am Ende der 80er Jahren vollständig verschwunden. Auf diese Weise bedeuten die Prozesse des psychischen Systems hier nur die Verbindung der intellektuellen Kommunikationen. Von Gedanke zu Gedanke bewegen sich die Bewusstseinsprozesse als Träger des psychischen Systems, und hier gibt es keinen psycho-emotionalen oder biologischen Instinkt: „Die Autopoiesis des Bewußtseins ist das Fortspinnen mehr oder minder klarer Gedanken, wobei das Ausmaß an Klarheit und Distinktheit selbstregulativ kontrolliert wird je nachdem, was für einen bestimmten Gedankenzug - vom Dösen und Tagträumen bis zur mathematischen Rechnung - zur Einteilung der Gedanken und zum Übergang erforderlich ist“. (Luhmann 1995:61). Diese auf einen rein intellektuellen Aspekt beschränkte Bewusstseinsauffassung kann natürlich sehr kluge Dinge über diese Art von intellektuellen Prozessen im individuellen Bewusstsein und zum Beispiel für die Reifung von Gedanken enthüllen, und Luhmann beschreibt darüber tiefe Einsichten. Zum Beispiel hebt er heraus, dass nur durch ständige Selbstbeobachtung im Bewusstsein der neue Gedanke auf eine reifere Ebene gebracht wird, und nur dadurch wird er auch dem Bewusstsein zugänglich: „Für sich ist ein Gedanke also zunächst nur ein Gedanke, für andere Gedanken dagegen, das heißt für den rekursiven Prozeß des Systems, ist eine Vorstellung“ (Luhmann 1995:65) Jeder, der eine völlig andere, neue Idee als zuvor hatte und nach der Konsolidierung selbst versucht hat, herauszufinden, wie sie in seinem Kopf geboren wurde, wird von dieser Studie von Luhmann viel Hilfe erhalten, und er wird das Gefühl haben, dass es tatsächlich so passiert sein könnte. Diese Gedanken von Luhmann sollen geschätzt werden, aber dies ändert nichts an der Tatsache, dass Luhmann zwar den intellektuellen Aspekt des individuellen Bewusstseins ausgezeichnet ausgearbeitet hat, aber die Auswirkungen emotionaler und biologischer Instinkte davon vollständig abschnitt.

In Bezug auf Emotionen bemerkten auch enge Anhänger und Kritiker von Luhmann seine theoretischen Mängel, und in 2004 wurde eine Sonderausgabe von Zeitschrift „Soziale Systeme“ in Bielefeld diesem Thema gewidmet. Schauen wir uns zuerst *Dirk Baecker* von den Luhmann-Verteidigern und *Luc Ciompi* von den Kritikern an. Baecker erkennt grundsätzlich an, dass Emotionen in Luhmanns Theorie und innerhalb von Luhmanns Analysen der bewussten Prozesse psychischer Systeme keine herausragende Rolle spielen.¹⁰ Einige untergeordnete Funktionen werden jedoch ein- oder zweimal von Luhmann in Bezug

⁹ Diese Studien erschienen zusammen in Luhmanns Band von 1995, und ich werde sie hier zitieren, siehe Luhmann: Soziologische Aufklärung 6. Die Soziologie und der Mensch. Westdeutscher Verlag. Opladen. 1995.

¹⁰ „In der Tat spielen Emotionen im Gesamtwerk von Niklas Luhmann keine sehr prominente Rolle. Luhmann hat zwar die Stelle markiert, an der seines Erachtens eine Theorie der Gefühle von seiner Theorie sozialer Systeme abzweigen, beziehungsweise an diese Theorie angedockt werden könnte, doch hat er dem Phänomen der Gefühle nie die Aufmerksamkeit geschenkt, die etwa Ciompi angesichts der nicht nur psychischen, sondern auch sozialen Prominenz des Phänomens für geboten hält.“ (Baecker 204:9-10.)

auf Emotionen angegeben, z. dass sie als Immunsystem des Bewusstseins fungieren, wenn der bewusste Denkprozess ins Stocken gerät, und dann wieder verschwinden, wenn der Denkprozess wiederhergestellt wird. Nach seinen Schriften hält Dirk Baecker diese untergeordnete Rolle in gewissem Maße auch für zu wenig, hält es jedoch für unnötig, Luc Ciompis radikale Erweiterung in Bezug auf die Rolle der Emotionen in dem Funktionieren des psychischen System zu akzeptieren. Er erwähnt jedoch einen praktischen Grund in dieser Ablehnung. Seiner Ansicht nach befindet sich die soziologische Emotionstheorie immer noch auf einem so unsicheren empirischen Niveau, dass es besser ist, diese Unsicherheit zu vermeiden.¹¹

Interessant ist jedoch Baeckers Erklärung, warum Luhmann die Rolle der emotionalen Welt in sozialen Ereignissen und sozialen Institutionen nur ungern anerkannte. Seiner Beschreibung nach war der Umgang mit Emotionen Luhmanns gesamtem theoretischen Stil fremd, und er führt dies auf die historischen Ereignisse in Luhmanns Jugend zurück, als die Unterdrückung fundamentalistischer und faschistischer Bewegungen in Hitlers Deutschland Luhmanns Lebensbedingungen prägte und diese Bewegungen sich stark auf die Befreiung von kollektiver Emotionen und ihre Schädigungseffekte stützten. Erst in Luhmanns spätem Werk begann sich dies etwas zu entspannen, aber dies änderte nicht mehr die ursprünglichen Richtungen.¹² Dies ist zwar nur Baeckers Meditation über das Finden der Ursache in dem Luhmannschen Denken, aber da er zwischen 1980 und 1995 viele Jahre direkt neben Luhmann in Bielefeld war, konnte er diese Seite von Luhmann kennenlernen, wer ansonsten in seinen inneren persönlichen Angelegenheiten sehr zurückhaltend war. Deshalb muss diese Erklärung beachtet werden. Dies ist natürlich keine Entschuldigung für Luhmanns theoretischen Mangel in diesem Bereich, daher verdient die Rolle, die Ciompi für Emotionen in den mentalen Prozessen der Psyche und in der strukturellen Konstruktion und der Ereignissen der sozialen Welt spielt, die größte Aufmerksamkeit.

Das Auslassen von Emotionen bei der Analyse psychischer Prozesse kann meines Erachtens jedoch nicht einfach durch Luhmanns besondere persönliche Sozialisation erklärt werden. Dies ergibt sich aus seiner theoretischen Entscheidung, die viel früher, fast zu Beginn seiner theoretischen Ausbildung, getroffen wurde. Demnach ist der Sinn der gemeinsame Baustein auf Systemebene über der Ebene des physischen und biologischen Systems, dh auf der Ebene sowohl des psychischen als auch des sozialen Systems. Diese grundlegende Entscheidung beinhaltet nicht die Erfassung von Emotionen, und diese Erfassung hätte nur durch Korrektur dieser grundlegenden Entscheidung getroffen werden können. Diese Korrektur hätte dann zusammen mit einer Reihe weiterer Änderungsanträge nur inkonsistent sein können, und Luhmann war nie bereit, solche Korrekturen in seiner Theorie vorzunehmen.

¹¹ "Affektkontrolle hat es immer auch mit dem Adressieren von Affekten zu tun, und niemand weiß, welche Emotionen dies hervorruft. Die Soziologie der Emotionen stellt diese Unvorhersehbarkeit und Unberechenbarkeit von Emotionen in Rechnung, indem sie sie als Phänomene beschreibt, in denen die verschiedenen Systemebenen des Organismus, des Bewusstseins und der Kommunikation aufeinandertreffen, ohne dass man genau wusste, wie sich diese Systeme ausdifferenzieren und wie sie miteinander gekoppelt sind. (...) Mir scheint diese Frage nicht entscheidbar zu sein, solange die Theorie der Emotionen so schwach ausgearbeitet ist, wie dies gegenwärtig der Fall ist." (Baecker 2004:12-13.)

¹² "Diese Abstinenz gegenüber dem Phänomen der Gefühle hat bei Luhmann Gründe, die etwa mit seinem Theoriestil zu tun haben, der wiederum seine zeithistorischen Motive hat. Wer eine Theorie entwirft, der es um Differenzierung im Umgang mit der Komplexität der Gesellschaft geht und die nicht zuletzt in der mangelnden Differenzierung Einfallstore für fundamentalistische und im Extremfall faschistische Gesellschaftsvorstellungen sieht, der wird sich nicht unbedingt mit Phänomenen beschäftigen, deren Funktion möglicherweise gerade darin besteht, Differenzierungen hochselektiv und problemgenau aufzuheben oder zumindest zu verwischen. Genau deswegen ist es ja so bemerkenswert, dass Luhmann diese Haltung in seinem Spätwerk lockert und sich ein "Verbindungsmedium" Werte überhaupt vorstellen kann." (Baecker 2004:10).

Neben der Erforschung der Psyche beschäftigte sich Luc Ciompi als medizinischer Psychiater auch mit Emotionen und überwachte kontinuierlich die Ergebnisse psychologischer und sozialpsychologischer Experimente auf diesem Gebiet. In seinen Analysen wird deutlich, dass das menschliche Denken in allen Momenten mit den emotionalen Komponenten der Psyche umgeben ist und zum größten Teil unser gesamtes Wissen nicht bewusst und spürbar von unserer emotionalen Disposition begleitet wird. (siehe Ciompi 1997, 2004) Wenn wir Wissen und Erfahrung sammeln, werden unsere emotionalen Einstellungen zum Thema zusammen mit Wissen und Erfahrung in unser Gedächtnis gedrückt, und wenn Wissen und Erfahrung später wieder hervorgerufen werden, erscheinen diese Emotionen auch, untrennbar in der Erfahrung und dem Wissen versunken. Wir denken nicht nur rein intellektuell, sondern was unser Interesse weckt, welche Erinnerungen und Assoziationen beim Anblick von etwas hochgeschoben werden oder erstickt bleiben, sind wichtig, und so werden bestimmte Gedankenpfade blockiert und andere in den Vordergrund gerückt. Das heißt, das Denken und die dazu benutzten Erfahrungen sind emotional gefärbt, und intellektuelles Denken und seine Denkprozesse laufen ständig gemeinsam mit emotionalen Prozessen in unserem Gehirn ab. Ciompi weist darauf hin, dass Emotionen mit unseren intellektuellen Prozessen zügig verlaufen, insbesondere wenn wir in einer neuen Situation neues Wissen gewinnen. Wenn wir es später neben bestehenden Wissen einbeziehen, wird unser neues Wissen, das wir aus neuen Erfahrungen gewonnen haben, bereits zur Routine und mit Wiederholungen werden die emotionalen Prozessen und der emotionale Energieverbrauch auf ein Minimum reduziert. Wenn jedoch später eine problematische Situation in Verbindung mit diesen auftritt, wird auch die emotionale Dimension reaktiviert und sie läuft neben der Problembearbeitung auf intellektueller Ebene auch.¹³

Zusammenfassend lässt sich sagen, dass von Luhmann die Psyche, das psychische System - oder mehrmals nur das „Bewusstsein“ genannt - auf mentale Prozesse beschränkt wird, die von Gedanke zu Gedanke ablaufen. Aus Ciompis Analysen geht jedoch hervor, dass von dieser Auffassung die reale menschliche Psyche verfehlt wird und sie ohne emotionale Prozesse nicht angemessen erfasst werden kann. Im Gegensatz zu Luhmann konnte *Nicolai Hartmann* mit seiner umfassend formulierten Schichtentheorie der Realität das Ensemble biologischer, psychologischer und intellektueller Komponenten der menschlichen Existenz und ihrer Zusammenarbeit adäquater formulieren. Wenden wir uns daher bei der Forschung des Themas im Folgenden Hartmann zu.

3. Nicolai Hartmanns Seinsichten

Hartmann kultivierte eine Art Philosophie, die er nicht als eine andere intellektuelle Aktivität zusätzlich zu den Wissenschaften auffasste, sondern als die umfassendste Art von Wissenschaft, den allgemeinsten Synthesizer spezifischerer wissenschaftlicher Erkenntnisse.

¹⁴ Infolgedessen versuchte er, basierend auf den empirischen wissenschaftlichen Ergebnissen seiner Zeit, die abstraktesten ontologischen Kategorien zu durchdenken, die aus der

¹³ “Zur Alltagslogik wird alles anfänglich Neue und Aufregende, das sich oft genug wiederholt. Die beteiligten Emotionen werden zwar durch Gewöhnung zunehmend unbewusst, behalten aber ihre typischen Operatoreffekte, “wie selbstverständlich” bei - ein höchst sinnvoller Mechanismus, der gleiche Leistung mit viel geringerem affektenergetischem Aufwand ermöglicht (...) Passiert indes Ungewohntes, so flammen alle latenten Affekte wieder auf.” (Ciompi 2004:10.)

¹⁴ Für eine ausführliche Beschreibung der Hartmannschen Philosophie siehe Wolfgang Stegmüller: Hauptströmungen der Gegenwartsphilosophie. Alfred Körner Verlag 1999. 3. Auflage;

Vergangenheit geerbt wurden. So versuchte er bestimmte Kategorien von Platon, Aristoteles, der Neuplatoniker und dann der großen Philosophen des Mittelalters und des New Age (Duns Scotus, Leibniz, Kant, Hegel und der Neokantianismus) im Lichte von Physik, Biologie, Psychologie und neueren sozialwissenschaftlichen Erkenntnissen zusammenzufassen. Im Gegensatz zu Luhmann, der seine allgemeine Sozialtheorie im Wesentlichen nur aus bestimmten Bereichen der Sozialwissenschaften zusammenstellte und die biologischen und physikalischen Systemebenen nur beiläufig berücksichtigte, untersuchte Hartmann systematisch die umfassendsten Zusammenhänge dieser Systemebene. So beschäftigte er sich in seinem ersten großen Werk mit den philosophischen Fragen der Biologie und später zusätzlich zu seiner langjährigen Arbeit in der Geschichte der Philosophie auch mit der ethisch-moralischen Sphäre der sozialen Welt in seiner systematischen Ethik, aber auch in einer gründlichen ästhetischen Monographie mit den künstlerischen Aktivitäten. Dann begann er, diese Untersuchungen als Theorie der Seinsschichten der realen Welt zusammenzufassen, welche Gedanke schon bei Aristoteles auftauchte. Er publizierte diese Theorie in 1933 zunächst in Bezug auf die intellektuelle Seinsschicht in einer systematischen Monographie und schloss dann 1940 eine umfassende Tabelle der Seinsschichten der realen Welt bzw. der Regelmäßigkeiten zwischen ihnen ein. Nach seinen Ambitionen kann diese Theorie als Synthese der umfassendsten wissenschaftlichen Kontexte der Zeit gelesen werden.

Hartmann sieht die reale Welt als eine Kombination aus vier im Laufe der Evolution aufeinander aufgebauten Seinsschichten. Die biologische Seinsschicht, die über der physischen Seinsschicht aufgebaut ist, und welche nach einer Evolutionsphase die psychische Seinsschicht bei einigen Tierarten geschaffen hat, und schließlich im Fall der Primaten auch die Anfänge der intellektuellen Seinsschicht erschienen, die gebaut auf den ersteren drei Schichten dann beim Menschen besonders ausgeprägt und entscheidend wurden. In dieser Überlagerung unterscheidet Hartmann zwischen dem Aufbau, in dem die charakteristischen Elemente der unteren Seinsschicht für die obere Schicht transformiert verwendet werden, wie es im Fall der biologischen Seinsschicht ist, und dem Modus, der lediglich auf der unteren Schicht aufbaut, aber dabei ihre Elemente nicht benutzt werden, und auf diese Weise eine autonomere Ablösung von der unteren Schicht ermöglicht wird. Dies wird von der psychischen Seinsschicht gezeigt, die über der biologischen aufgebaut ist, sowie auch von der geistlichen Seinsschicht, die über der psychischen aufgebaut ist. Eine wichtige Korrelation in der Verbindung der vier Seinsschichten besteht darin, dass die obere Seinsschicht zwar immer einen höheren Evolutionsgrad aufweist, die unteren jedoch ontologisch stärker sind als die höheren. Denn die Höheren können die Gesetze der niederen Schichten nicht umwandeln, und ohne sie zu verletzen, können sie die Gesetze ihrer eigenen höheren Schichten entfalten. Somit können die biologischen und darüber liegenden Seinsschichten nicht gegen die zugrunde liegenden physikalischen Gesetze verstoßen, so wie Tiere mit einer psychischen Seinsschicht und Menschen nur mit dem Funktionieren biologischer Gesetze überleben können und schließlich soziale Institutionen, die mit der geistigen Schicht des Menschen geschaffen wurden, können nur aufgrund der Gesetze der unteren Seinsschichten überleben, ohne sie zu verletzen.

Zusätzlich zu diesen vier Seinsschichten, die Hartmann die Schichten der realen Existenz nennt, unterscheidet er die ideale Existenz, die nur im Fall der obersten geistlichen Seinsschicht eine Rolle spielt. Diese ideale Existenz kennzeichnet nur solche geistliche Produkte, die nur auf Papier oder in einer anderen ähnlichen fixierten Form existieren, werden sie aber in keiner menschlichen Gemeinschaft benutzt. Um diesen Unterschied zu benennen, verwendet Hartmann einerseits für die intellektuellen Produkte und kognitiven Handlungen, die als Teil der Schicht der realen Existenz fungieren, das „lebendige“ Geist, und andererseits die „tote“ geistliche Existenz der objektivierten intellektuellen Produkte, die sich bereits von der realen Existenz gelöst haben und nur im Papier existieren. Letztere sind immer noch Teil

der idealen Existenz, aber nicht der realen Existenz, aber im Falle der Fixierung und der Möglichkeit, zu ihnen zurückzukehren, kann die ideale Existenz objektivierter intellektueller Produkte immer zur realen Existenz werden, wenn eine Gesellschaft sie für das Funktionieren ihrer intellektuellen Existenz und ihrer Institutionen nutzt.

Wenn man sich den vier Seinsschichten nähert, verursacht die psychische Seinsschicht ein Verständnisproblem für den Forscher. Um dies zu verstehen, muss man wissen, dass Hartmann den persönlichen Geist innerhalb der „lebendigen“ geistlichen Seinsschicht trennt, die die höchste Schicht der vier Schichten des einzelnen Individuums darstellt. Der reale objektive Geist in den menschlichen Gesellschaften wird durch die Aktivitäten von Millionen persönlicher Geister und die daraus resultierenden intellektuellen Institutionen (moralische Normen, Gesetze, Sprachsysteme usw.) repräsentiert, die in ihre soziale Funktionsweise einfließen. (Wenn ein intellektuelles Produkt es nicht tut, bleibt es als toter Geist in Papier fixierter Form der idealen Existenz). Somit befindet sich die an jede Person gebundene psychische Seinsschicht immer in der Nachbarschaft des persönlichen Geistes und fließt im menschlichen Bewusstsein zusammen. Nur ein Teil der psychischen Seinsschicht erscheint im Bewusstsein, und die mentalen unbewussten Prozesse ergänzen dies. Aber wo ist die Grenze und was gibt das eigene Terrain für die psychisch-Seinsschicht? Wenn wir nach einer Antwort auf diese Frage suchen, finden wir bei Hartmann eine Verwirrung.

In seinem Buch über die Probleme der geistlichen Seinsschicht von 1933 geht er näher darauf ein und schreibt hier, dass die psychische Seinsschicht ein dünnen Streifen zwischen den Einflüssen der angrenzenden unteren biologischen und oberen geistlichen Seinsschichten darstellt: „Das Bewußtsein ist gleichsam nur ein schmaler Streifen zwischen ihm und dem unbewußt Seelischen. Beide greifen mannigfach bestimmend in seiner Ebene ein, aber in entgegengesetztem Sinne“ (Hartmann 1962: 50-51) Die prozesse des psychischen Unbewußtes reichen bis in die biologischen Instiktwelt und die unbewusste Welt besteht hauptsächlich aus diesen umgeformten Instikten. Aus der anderen Richtung greifen jedoch die Prozesse des persönlichen Geistes nach unten in die psychische Seinsschicht und beiden gegensätzlichen Einflüsse zusammen geben das Funktionieren der psychisch-spirituellen Seinsschicht. Aus dieser Beschreibung geht jedoch genau die Natur des dünnen spezifischen Streifens ungeklärt, und die psychische Seinsschicht wird tatsächlich in der unteren und oberen Richtung zwischen zwei benachbarten Seinsschichten aufgeteilt. Ist dies jedoch der Fall, stellt sich die Frage, warum es nicht genug ist, nur über die Instikte der biologischen Seinsschicht von unterhalb und die bewussten Prozesse des persönlichen Geistes der geistigen Seinsschicht von oberhalb zu sprechen? Aber nicht nur die Besonderheit des „dünnen Streifens“ bleibt hier dunkel, sondern es fehlen auch Emotionen in Hartmanns Beschreibung der psychischen Seinsschicht. Wenn wir nach alternativen Beschreibungen suchen, dann ist einer von Hartmanns Diskussionspartnern, *Max Scheler*, in dessen Beschreibung die Emotionen erscheinen, und *Arnold Gehlen* als ihr gemeinsamer Diskussionspartner streitet heftig mit Scheler über genau dieses Thema.

Insbesondere in seiner Zusammenfassung von 1940 weist Hartmann darauf hin, dass von den beiden oberen Seinsschichten von Psychischen und Geistlichen die Produkten der unteren biologischen Seinsschicht nicht transformiert werden, sondern ohne jegliche Transformation die obere Schichten nur darauf aufbauen, dh die psychische Schicht direkt auf die biologische und die geistliche auf die psychische Schicht. Dies steht im Gegensatz zu dem, was Hartmann einige Jahre zuvor geschrieben hat, und er versöhnt die beiden gegensätzlichen Behauptungen, indem er betont, dass diese vier Seinsschichten zusätzlich neben ihrer Unabhängigkeit immer noch in einzelnen Menschen vereint sind: „Das seelische Sein enthält die organische Prozesse nicht, wohl aber enthält „der Mensch“ sie in sich; denn der Mensch ist selbst ein geschichtetes Wesen, er ist auch Organismus, und folglich auch ein materiell-körperschaftes Gebilde. Insofern hat er die niederen Kategorien alle als konstituierende

Momente an sich." (Hartmann 1940:496) Diese Analysen ersetzen jedoch nicht, dass der „dünne Streifen“ der psychischen Seinsschicht inhaltlich nicht gefüllt ist. In dem Band von 1940 finden wir einen weiteren Hinweis, der uns näher an die Suche nach dem Terrain der psychischen Seinsschicht bringen könnte. Denn schreibt er im Zusammenhang mit der Unabhängigkeit der beiden oberen Seinsschichten vom Material der unteren Schichten, dass ebenso wie die Räumlichkeit und träge Substanz der unteren Schichten in der psychischen Seinsschicht nicht auftreten, so erscheinen die Aktcharaktere der Seelische im objektiven Geist der intellektuellen Seinsschicht auch nicht.¹⁵ Somit ist die psychische Seinsschicht das Reich menschlichen Handelns im Gegensatz zum Terrain des objektiv Geistes der intellektuellen Seinsschicht, wo die menschlichen Handeln nicht erscheinen. Hartmann unterscheidet also nur den objektiven Geist von der Seele und erwähnt nicht den persönlichen Geist, der der psychischen Seinsschicht im Individuum direkt benachbart ist. Daher können wir auch aus dieser Analyse das gesuchte Material der psychische Seinsschicht nicht erhalten.

Um dies auszugleichen, wenn wir nicht vollkommen die psychische Seinsschicht verwerfen und die Funktionsweise des Menschen und der menschlichen Gemeinschaft anstatt vier in drei Seinsschichten beschreiben wollen, können wir als Lösung nur die frühe Analyse von Luhmann über die Rolle der biologischen Sexualität in der Emotion der Liebe verwenden. Im Allgemeinen kann festgestellt werden, dass der Instinkt der biologischen Seinsschicht durch die Transformation von Emotionen in der psychischen Seinsschicht in gezähmter Form und absorbiert erscheint. Andererseits erhalten auch die Emotionen eine Transformation (eine Dämpfung und Anreicherung gleichzeitig) durch intellektuelle Produkte der geistlichen Seinsschicht. Wie im Fall des Sexualtriebs wird von den psychischen Emotionen der Liebesbeziehung die Wildheit der bloßen biologischen Sexualität verändert, aber schon mit den intellektuellen Obertönen und Erfahrungen der großen Liebesbeschreibungen, Romanen und wunderbaren Melodien der Liebeslieder aller Zeiten, von denen die Menschen sozialisiert wurden. In tatsächlichen Liebesbeziehungen und Kommunikationen ist der Zusatz dieser drei Seinsschichten untrennbar miteinander verbunden und aufeinander aufgebaut, und im Fall der eventuellen geringe Präsenz einer von ihnen kann es verursachen, dass die Liebesbeziehung nur in begrenztem Umfang überleben wird. Um diesen Aufbau verstehen zu können, muss im Fall der Instinkte akzeptiert werden, dass die Instinktwelt, die in der Tierwelt auf bestimmte Verhaltensweisen und Bewegungskoordinationen beschränkt ist, erscheint im Fall der Menschen einerseits reduziert aber andererseits in der Form allgemeiner Motivationen (Geschlechtsinstinkt, Lebensinstinkt, Pflgeinstinkt usw.) und sie haben in dieser Form viele Funktionen. Arnold Gehlen kritisierte Konrad Lorenz, dass Lorenz die Instinkte in irgendeiner Form im Fall der Menschen angenommen hatte, aber in dieser Form muss die Wahrheit von Lorenz akzeptiert werden.¹⁶ Es ist noch zu erwähnen, dass auch Max Scheler die Rolle von Instinkten in Form von instinktiven Bestrebungen beim Menschen hervorhebt und er schreibt ihre Funktion als Vermittlungsschicht zwischen biologischen Lebensfunktionen und Bewusstsein hervor¹⁷ (siehe Scheler 1928:30-33)

¹⁵ "Räumlichkeit und träge Substanz kehren oberhalb des Organischen nicht wieder, die Aktcharaktere des Seelischen nicht im objektiv Geiste". (Hartmann:1940:512).

¹⁶ „Dennoch sind die Versuche, das Konzept des Instinkts direkt auf den Menschen zu übertragen, äußerst dünn und enttäuschend (...), insbesondere vertreten durch Konrad Lorenz. (...) Es gibt keine allmähliche Beziehung zwischen intelligentem und instinktivem Verhalten, aber (...) es gibt eine Tendenz zum gegenseitigen Ausschluss.“ (Gehlen 1976: 32)

¹⁷ "Wir können den Fehler sehen und das ist der grundlegender Fehler von Descartes, der das Ignorieren des tierischen und menschlichen Instinktsystems bedeutet, obwohl sie die Vermittlung und Einheit zwischen Bewegungen des realen Lebens und den Inhalten des Bewusstseins sind" (Scheler 2016: 92).

Ein weiteres Problem und Hindernis für die mögliche Transformation von Hartmanns Lösungen ist die Stärke der evolutionär unteren Seinsschicht, die für die obere Schicht unantastbar ist. In den achtzig Jahren seit dem Schreiben seiner Synthese von 1940 hat die geistliche Seinsschicht aber eine so starke Transformationsfähigkeit insbesondere in Bezug auf die unterste physische Schicht aber auch auf die biologische Schicht erreicht, dass diese These von Hartmann in Frage gestellt werden kann. Die Stärke der unteren beruht auf der Tatsache, dass ihre Existenz unabhängig vom Einfluss der oberen ist. Die obere Schicht kann die Elemente der unteren Schicht nur als Baustein verwenden und eine komplexere Bildung der oberen Schicht erzeugen, wie z. B. die Bildung biologischer Organismen aus dem Material der physikalischen Schicht stattfindet, aber diese transformierte Verwendung wirkt sich nicht auf den unteren Teil aus: "Die höheren Kategorien setzen stets eine Reihe niederer voraus, sind aber ihrerseits in diesen nicht vorausgesetzt. Kategoriale Abhängigkeit also waltet durchgehend von den niederen zu den höheren, nicht aber umgekehrt." (Hartmann 1940:519-520.) Im Gegensatz zu dieser Behauptung können wir aber sagen, dass im Fall der höchsten geistlichen Seinsschicht in der Zwischenzeit solche Fortschritte erzielt wurden, dass die elementare Ebene der physikalischen Materie mit immer tieferer Auflösung und im Nanobereich transformiert werden kann und es bisher nicht existierende Material (z. B. zuletzt Graphen) erstellen kann, und ganz neue Entitäten aus dem bisher nicht existierende Formationen konstruiert werden können. Somit wird die oberste geistliche Seinsschicht nicht nur evolutionär höher, sondern sie kann zunehmend bestimmte Elemente der physischen Seinsschicht ohne Grenzen transformieren, dh die untere verliert ihre ontologische Stärke.

In gleicher Weise wird die biologische Seinsschicht zunehmend von der geistigen Seinsschicht dominiert und diese obere Schicht baut nicht einfach darauf auf, denn nach der Entdeckung der DNA im Jahr 1951, deren Existenz Hartmann nur theoretisch erraten hatte,¹⁸ kann der genetische Redakteur nun genauer darauf zugreifen. Auf diese Weise wird das resultierende lebende Individuum durch die geistliche Seinsschicht gestaltbar. Aber die Lösungen, die heute nur in Forschungslabors für die Schaffung der künstlichen Arterie oder auf die gleiche Weise für die Erneuerung der Hauptorgane des menschlichen Körpers durch implantierte Stammzellen und deren radikale Verlängerung des Lebens bekannt sind, führen dazu, dass die Gesetze der biologischen Seinsschicht von der geistlichen Seinsschicht überschrieben werden können. Erwähnt werden sollten aber auch Experimente mit der Emulation des menschlichen Gehirns und Geistes, die bei Erfolg dazu führen können, dass die intellektuelle Seinsschicht der einzelnen Menschen in der Zukunft ohne biologische Seinsschicht voll funktionsfähig auf einen Computerträger übertragen wird.

In Anbetracht dessen kann gesagt werden, dass Hartmann die bereits in seinem Alter beobachteten Fortschritte in Bezug auf die Fähigkeit der geistigen Seinsschicht, in die unteren Seinsschichten einzudringen, nicht genug hervorgehoben hat, und daher erwies sich These der Unverletzlichkeit des Niedrigeren von der höheren Schicht in Bezug auf alle Schichten des Lebens im Fall der höchsten geistlichen Schicht als fehlerhaft. Diese Korrektur kann auch wichtig sein, weil wir theoretisch davon ausgehen können, dass durch die zukünftige Entwicklung der künstlichen Intelligenz und ihre mögliche Ablösung vom menschlichen Bewusstsein diese abnehmende Abhängigkeit der oberen Schicht von den unteren Schichten nur vervollständigt würde. Auf diese Weise könnte es sein, dass die künstliche Intelligenz, die autonom geworden ist, als neue Seinsschicht nur auf der Grundlage der niedrigsten physischen Seinsschicht ohne biologische Schicht existieren könnte.¹⁹ Ohne diese

¹⁸ "Man weiß hier sehr wohl, dass der morphologische Bau des Organismus seine bestimmten Gesetze hat (...) man weiß auch, dass die letzteren Funktionsgeformtheiten sind, in denen die Gewähr liegt, dass die gleiche Formung sich wiederbildet und somit das Leben sich erhält." (Hartmann 1940: 456)

¹⁹ Siehe für Details meine frühere Schrift, Béla Pokol, Künstliche Intelligenz: Die Entstehung einer neuen Seinsschicht? (KI im Spiegel von Nical Hartmanns Ontologie) Pázmány Law Working Papers 2018 No. 12.

abnehmende Abhängigkeit könnte diese Möglichkeit nicht existieren, aber mit der Annahme dieser Entwicklung könnte es nur als die Vervollständigung der bereits abnehmenden Abhängigkeit aufgefasst werden.

Literature

- Baecker, Dirk (2004): Einleitung: Wozu Gefühle? Soziale Systeme (Vol. 10) 2004, Heft 1. 5-20.p.
- Ciampi, Luc (1997): Zu den affektiven Grundlagen des Denkens. System Familie (Vol. 10) 1997. No. 1. 128-134. p.
- Ciampi, Luc (2004): Ein blinder Fleck bei Niklas Luhmann? Soziodynamische Wirkungen von Emotionen nach dem Konzept der fraktalen Affektlogik. Soziale Systeme (Vol. 10.) 2004 Heft 1. 21-49. p.
- Gehlen, Arnold (1974): Der Mensch: seine Natur und seine Stellung in der Welt. Athenaeon Verlag.
- Hartmann, Nicolai (1940): Der Aufbau der realen Welt. Walter de Gruyter. Berlin.
- Hartmann, Nicolai (1933) (1962): Das Problem des geistigen Seins. Untersuchungen zur Grundlegung der Geschichtsphilosophie und der Geisteswissenschaften. 3. unveränderte Auflage. Walter de Gruyter. Berlin.
- Luhmann, Niklas (1975): Evolution und Geschichte. In: uö. Soziologische Aufklärung 2. 1975. 150-170.p.
- Luhmann, Niklas (1981): Soziologische Aufklärung 3. Westdeutscher Verlag. 1981. 231. p.
- Luhmann, Niklas (1984): Soziale Systeme. Suhrkamp, Frankfurt am Main, 1984.
- Luhmann, Niklas: Soziologische Aufklärung 6. Die Soziologie und der Mensch. Westdeutscher Verlag. Opladen. 1995.
- Maturana, Humberto R./Francisco J. Varela : The Tree of Knowledge. The Biological Roots of Human Understanding. Revised Edition. Shambhala Publications. Boston, Massachusetts. 1987.
- Maturana, Humberto R./Sima Nisis de Rezepka: Human Awareness: Understanding the Biological Basis of Knowledge and Love in Education. In Jane Cull (ed.): Life's Natural Solutions. 2014, 2-27.p.
- Pokol, Béla (2018): Künstliche Intelligenz: Die Entstehung einer neuen Seinsschicht? (KI – im Spiegel von Nicolai Hartmanns Ontologie) Pázmány Law Working Papers, Nr. 2018/12.
- Scheler, Max (2016): Die Stellung des Menschen im Kosmos. Verlag der Contumax GmbH. Berlin. 2016

2

Künstliche Intelligenz: Die Entstehung einer neuen Seinsschicht? (KI – im Spiegel von Nicolai Hartmanns Ontologie)

Die wachsende Stärke der künstlichen Intelligenz (KI) schreckt in den letzten Jahren schon ihre Verwendung lenkenden Praktiker (Elon Musk) und die Theoretiker (Stephen Hawking, Nick Bostrom) und sie wird dargestellt als über die Menschen wachsende und von menschlicher Kontrolle befreite gefährliche Entwicklung. Eine weitere Beschreibung, die aus einer Bemerkung von John von Neuman stammte, thematisiert die inhärente Neuheit der KI als die Entstehung der Ära von Singularität.²⁰ Gemäß dem letzteren führen die zunehmende Rechenleistung und schnelleres Programm aufgrund der selbstlernende künstliche Intelligenz zu einem Punkt, wo die die Beschleunigung bis dahin verlangsamende menschliche Teilnahme aus der Prozesse der KI beseitigt wird. Von diesem Moment, der in der Weltgeschichte einzigartig (Singular) ist, wird die selbstlernende KI auf tausendfache Geschwindigkeit wachsen und innerhalb weniger Stunden wird sie völlig unverständlich auch für IT-Profis. Von dieser Entwicklung - zusammen mit der wachsenden Robotik – wird für die KI die Herstellung aller Dinge ermöglicht, und auf diese Weise wird für die Menschen nicht nur die künstliche Intelligenz unverständlich, aber sie werden von der Führung der Welt entfernt. Und das ist das Auftreten der Ära von Singularität.

Wenn wir diese Konzepte und Ängste vor der künstliche Intelligenz für eine Weile klammern und stattdessen stellen wir die philosophischen Konzepte in dem Mittelpunkt, von denen die bisherigen evolutionären Sprünge des Seins auf der Erde analysiert wurden, dann können wir die besser fundierten Kategorien der Ontologie für das Verständnis der Neuheit der KI verwenden. In den letzten 100 Jahren wurden schon auf empirische Basis beruhenden ontologischen Analysen hergebracht und aufgrund diesen Analysen können die Neuheit der künstlichen Intelligenz besser verstanden werden. Ich halte die Analysen von Nicolai Hartmann über die ontologischen Seinsschichten und über die Evolution dieser Schichten für angebracht, um die jetzige evolutionäre Sprung durch die KI mit den bisherigen Sprünge vergleichen zu können. Ganz sicher lautet die philosophische Gedankenweise und die ontologische Analyse für die etablierte Denkschamata in der Forschung der KI einigermaßen fremd aber ich denke, dass die Ontologie Nicolai Hartmanns, insbesondere seine Lehre der Seinsschichten einen wesentlichen Beitrag zu heutigen Untersuchungen über die Natur der künstlichen Intelligenz leisten kann. Hartmann entwirft ein Bild des Schichtenbaus der Welt, und zwar als eine Überlagerung von vier Hauptschichten: physische, organische, seelische und geistige Seinsschichten. Aufgrund dieser ontologischen Struktur entfaltete er seine Theorien über die moralischen Prinzipien der sozialen Welt, die Gesetze der Ästhetik und andere Insitutionen der Gesellschaft (siehe Hartmann 1926; 1953; 1962). Sein Oeuvre wurde

²⁰ Die Tatsache, dass Neumann der erste Schöpfer der Idee der Singularität ist, kennen wir nur indirekt von Stam Ulam. Er erinnerte sich an ein Gespräch mit ihm aus den frühen 1950er Jahren, wenn diese Idee von Neumann erwähnt wurde: “The ever accelerating progress of technology and changes in the mode of human live give the appearance of approaching some essential singularity in the history of the race beyond which human affairs as we know them, could not continue”. This is the first known use of the word “singularity” in the context of human technological history.” (Zitat von Kurzweil, 2012: 185).

für ihn die Anerkennung als einer der größten Philosophen der zwanzigsten Jahrhundert gesichert und seine Analysen über die geistlichen Erscheinungen scheinen besonders geeignet, die Entwicklung der künstlichen Intelligenz und ihre Zusammenleben mit der menschlichen Intelligenz besser verstehen zu können.

Von Hartmann wurde es geschrieben, wie einst die biologische Seinsschicht über die physische Seinsschicht stabilisieren konnte und während der Entwicklung von Pflanzen und dann die Tieren höherer Sprossen im Entwicklungsstand der Säugetieren eine emotional-seelische Seinsschicht über die physische und biologische Seinsschichten entstand und durch die allmähliche Evolution von Primaten die Keime der geistigen Seinsschicht geschaffen wurden, die in den menschlichen Gemeinschaften besonders dominierend wurden und die unteren Seinsschichten immer stärker durch die geistige Seinsschicht bestimmt wurden. Jetzt stehen wir vielleicht mit der Entstehung von KI vor einem neueren evolutionären Sprung und über die geistige Seinsschicht der Menschen (und für diese Seinsschicht unerlässliche untere Seinsschichten) entsteht eine neue Seinsschicht als die Nachfolger der bisher zu dem Menschen gebundenen geistigen Seinsschicht, die als selbstorganisierende Intelligenz von nun an die höchste Seinsschicht der Welt die dominierende Kraft sein wird.

Nicolai Hartmann hat schon bestimmte Zusammenhänge zwischen den Seinsschichten behauptet und für das Zusammenleben der immer neueren Seinsschicht mit der unteren und älteren Seinsschichten hat er Gesetzmässigkeiten angegeben. Auf diese Weise lohnt es sich, seine Analysen hier kurz zusammenzufassen, bevor wir die Seinsschicht der selbstorganisierenden künstlichen Intelligenz untersuchen.

1. Der Mensch und die Hierarchie der Seinsschichten

In dem Mensch wird das eigentümlich „Menschliche“ von der geistigen Seinsschicht beinhaltet und ihre allmähliche Dominanz über die physische, biologische und seelische Seinsschichten bedeutet die Evolution des menschlichen Lebens, aber in jedem Augenblick wird der Mensch von den Gesetzen der vier Seinsschichten gleichzeitig bestimmt. Der Mensch ist ein mehrschichtiges Wesen und die menschlichen Gemeinschaften können nur in dem kumulativen Rahmen der Gesetzen der vier Seinsschichten entfalten. Die oberen Seinsschichten können nur entwickeln, wenn die Gesetze der Niederen respektiert werden, aber dies ist kein Hindernis für die Autonomie der Gesetze der oberen Seinsschicht in Bezug auf die Gesetze der niederen Seinsschichten. Der Aufbau der höheren Seinsschicht bedeutet zuerst die Überformung der Kategorien der niederen Seinsschicht, aber die noch höhere Schichten bedeuten nicht mehr solche Überformung, sondern sind sie mit eigenen Kategorien über die niederen Seinsschichten gebaut. Während die wesentlichen Elemente der physischen Welt von der biologischen Seinsschicht verwendet werden - nur durch die Gesetze ihrer eigenen Seinsschicht überformt - werden keine wesentlichen Elemente der niederen Schichten von der Seelische und danach der geistigen Seinsschicht beinhaltet.²¹ Hartmann beschrieb diese Zusammenhänge wie folgt: „Um Mehrschichtigkeit zu begreifen, genügt es, sich an allgemein Bekanntes zu halten. Niemand zweifelt, daß organisches Leben sich vom Physisch-Materiellen wesentlich unterscheidet. Aber es besteht nicht unabhängig vom diesem: es enthält

²¹ In Fußnote kann es erwähnt werden, dass die Hirnforschung in der Zwischenzeit aufgedeckt hat - basierend Donald O. Hebb's Initiativen in 1949 -, dass unter den Hundertmillionen der Neuronenzellen in dem Gehirn von der immer neuen Erfahrungen und Wissenstücken die einzelnen Gruppen dieser Neuronenzellen immer in neuen Andordnung organisiert werden. So haben auch die geistigen Prozesse eine materiellen Basis in dem Gehirn. Für eine Analyse darüber, siehe in dem Kurzweils Buch der Kapitel Neokortex (Kurzweil, 2012: 85-95).

es in sich, beruht auf ihm, ja die Gesetze des Physischen erstrecken sich tief in den Organismus hinein. Was nicht hindert, daß dieser über sie hinaus noch seine Eigengesetzlichkeit habe, die in jenen nicht aufgeht. Solche Eigengesetzlichkeit überformt dann die niedere, allgemein physische Gesetzlichkeit. Ähnlich ist es mit dem Verhältnis des seelischen Seins zum organischen Leben. Das Seelische ist, wie die Bewußtseinphänomene beweisen, dem Organischen durchaus unähnlich, es bildet offenbar über ihm eine eigene Seinsschicht. Aber es besteht überall, wo wir ihm begegnen, in Abhängigkeit vom ihm, als getragenes Sein. (...) Das seelische Sein ist also zwar getragenes Sein, aber in seiner Eigenart ist es bei aller Abhängigkeit autonom. Schließlich ist es seit der Überwindung des Psychologismus eine wohlbekannte Tatsache, daß das Reich des geistigen Seins in dem des seelischen und seiner Gesetzlichkeit nicht aufgeht. Weder die logische Gesetzlichkeit noch das Eigentümliche von Erkenntnis und Wissen hat sich psychologisch ausschöpfen lassen. Noch viel weniger die Sphäre des Wollens und Handelns, der Wertung, des Rechts, des Ethos, der Religion, der Kunst. Diese Gebiete alle ragen, schon rein dem Phänomengehalt nach, weit hinaus über das Reich des psychischen Phänomene. Sie bilden als geistiges Leben eine Seinsschicht eigener und höherer Art, mit deren Reichtum und Mannigfaltigkeit sich die niederen nicht entfernt messen können. Aber auch hier waltet sich das gleiche Verhältnis zum niederen Sein. Der Geist schwebt sich nicht in der Luft, wir kennen ihn nur als getragenes Geistesleben - getragen vom seelischen Sein, nicht anders als diese vom Organismus und weiter vom Materiellen getragen ist. Auch hier also, und zwar hier erst recht, handelt es sich um Autonomie der höheren Schicht gegenüber der niederen, gerade in der Abhängigkeit von ihr. (Hartmann, 1962:16-17)

Also ist der Mensch die Einheit vierer Seinsschichten, und die menschliche Vernunft kann Auswirkungen nur aufbauend über die unteren Seinsschichten ausüben - über die biologische Grundlage des menschlichen Körpers. Insbesondere die rein intellektuelle Tätigkeit ist das Terrain der geistigen Seinsschicht und Hartmann unterscheidet drei internen Gebiete dieser Schicht: das Gebiet des individuellen Geistes, das des objektiven Geistes und das der objektivierten Geistes. Die ersten beiden ist der lebendige Geist und der objektivierter Geist bedeutet das Terrain des toten Geistes, aber es kann immer auf die Inhalte des objektivierten Geistes zurückgegriffen werden, und so können diese Inhalte in den lebenden Geist zurückgebracht werden. Der individuelle Geist lebt mit den Inhalten des objektiven Geistes seiner Zeit gemeinsam, und mehr oder weniger trägt eine Vielzahl individueller Geister den objektiven Geist und seine innere Formen als den Volksgeist und andere kollektiven geistigen Formen der Epoche. Aber auch der individuelle Geist hat weitgehend solche Inhalte, die der objektiven Geist enthält, und so können ihre Verhältnis als gegenseitige Träger voneinander beschrieben werden. Die dritte Form - die des objektivierten Geistes - vermehrt sich immer mit der Anreicherung der Fixierung der geistigen Inhalte aufgrund der Schrift und anderer Fixierungsformen. Auf diese Weise können die individuellen Geister neben der Inhalte des objektiven Geistes ihrer Epoche zusätzlich die intellektuellen Inhalte aller Ära benutzen und als Reaktion kann auf diese Weise die Inhalten und Formen des objektiven Geistes der Epoche bereichern. An der Ebene der geistigen Seinsschicht entsteht so die lebende Kollektivität, während an der biologischen Ebene der Rahmen gemeinsamer Existenz nur von der Gemeinschaft der Rasse über ihre immer verschwindenden Individuen getragen wird und ebenso kann das seelische Leben immer nur für die einzelnen Individuen isoliert sein und es ist nicht übertragbar. Wie Hartmann schreibt: „Sein seelisches Sein hat jeder für sich. Es ist esoterisches Sein des Individuums, unübertragbar, mit dem Man wohl Fühlung haben, in das man aber nicht hineingelangen kann. Man kann wohl mit ihm mitleiden und sich mitfreuen: aber es ist und bleibt ein zweites Leiden und ein zweites Sichfreuen neben dem original, und es bleibt auch bei aller Innigkeit ein qualitativ von ihm verschiedenes. Den Gedanken aber, den einer hat, kann man als denselben denken, wenn man ihn erfaßt; es ist

zwar ein zweiter Gedankenakt, Akt eines anderen Bewußtseins, aber es ist derselbe Gedanke” (Hartmann 1962:71)

Hartmann macht einen weiteren Unterschied innerhalb der geistigen Seinsschicht, die in der Grenze zwischen dem objektiven (lebenden) Geist und dem objektivierten (toten) Geist erscheint. Gemäß dieser Unterscheidung können die fixierten geistigen Inhalten der Vergangenheit - Überzeugungen, Verhaltensmuster, moralischen und kulturellen Werte, usw. – in den Gegenwart als massenhaft gefolgt Selbtsverständlichkeiten erscheinen. Aber es bedeutet eine andere Art von Hineinragen der Vergangenheit, wenn es nur als ein rein objektivierten intellektueller Gehalt für das Erreichen der Individuen festgelegt ist, aber nicht mehr erscheint auf der Ebene der massenhaft engagierten Glauben, Wissen und Vorurteile. Dann kann nur der einzelne individuelle Geist bewusst auf diese tote objektivierte geistige Inhalte zurückgreifen und nur er kann sie in den lebenden objektiven Geist hineinholen: „Das ist das Inkraftsein oder Noch-Lebendigkeit (einer Sitte, Anschauung), also die Kraft der „Sache“, den fortlebenden Geist mit einer gewissen Stetigkeit bei sich festzuhalten, auch da, wo er sich sonst zusehends verändert (...) Beim vernehmlichen Hineinragen ist es überall anders, wo die Sache selbst nicht mehr fortlebt, die unmittelbare Tradition abgerissen ist.” (Hartmann 1962:38). Sehen wir uns an, wie sich das Verhältnis dieser drei Gebiete der geistigen Seinsschicht verändert hat - bereits jenseits Hartmann's Zeit in den 1930 Jahren - und wie die Formen der heutigen künstlichen Intelligenz in den alten Inhalt gewellten.

2. Die zunehmende Verwebung der künstlichen Intelligenz in die geistige Seinsschicht

Die dominierende Rolle der geistigen Seinsschicht in den menschlichen Gemeinschaften und die relative Unterdrückung der bestimmende Stärke der niederen Seinsschichten hat mit der Möglichkeit der Fixierung des Sinnes von irgendeiner Form des Schreibens begonnen. Natürlich war dies zuerst nur ein dünner Rahmen für das Leben der menschlichen Gemeinschaften in den Zivilisationen, von denen diese Niveau erreicht wurden, und die breiten Massen bzw. ihr tägliches Leben waren von diese nicht berührt. Auch die Erfindung von Druck in der Mitte von 1440er Jahren in der europäischen Zivilisation konnte diesen Zustand nicht verändern, aber für die Individuen der oberen Schichten hat es durch diese technischen Anlagen begonnen, die Alphabetisierung Bedeutung zu gewinnen und es began auch in dem Alltagsleben, das Erleben und Handeln immer häufiger durch die fixierten sinnhaften Inhalten auszugestalten. Während der 1800er Jahren hat sich diese Erscheinung in der ganzen Europa und in anderen Kontinenten mit europäischen Kultur verbereitet und allmählich auf die gesamte menschliche Gesellschaft erweitert. Aufgrund der universellen Alphabetisierung hat sich die fixierten sinnhaften Inhalten der Tagespresse, Journalen usw. in den frühen 1900er Jahren schon in dem Alltagsleben verwoben, und dann wurde diese Erweiterung durch die Filme und Radio noch weiter erhöht. Im 1950er Jahren kam die allgemeine Verbreitung des Fernsehens und jeder Minute des täglichen Lebens wurde von den geschriebenen, audio-beweglichen Formen der fixierten geistlichen Inhalten immer mehr mitbestimmt. Durch dieser Änderung hat es begonnen, dass die die geistige Seinsschicht über die niederen Seinsschichten immer mehr dominant wurde und die niedere nur teils überformt teils hintergeschoben ihre Einfüsse ausüben konnten und diese war verstanden als die Prozess der Zivilisation.²²

²² Nobert Elias veranschaulicht den Verlauf dieser Zivilisation in den veränderten Formen der Erfüllung der natürlichen Bedürfnisse aufgrund reicher empirischer Basis, siehe Elias, 1976.

Diese Entwicklung erhielt jedoch ein wirklicher Aufschwung noch später und seit den 1980er Jahren hat es begonnen, die Digitalisierung des Schreibens durch die massenhafte Verbreitung von Personalcomputern alle Lebensbereiche durchzudringen. Auf diese Weise konnte die Fixierung der geistigen Inhalt im Zustand der konstanter Korrektur bleiben, aber damit wurde die geistige Seinsschicht nicht nur flüssig gemacht sondern wurde der individuelle flüssige Geist durch die Texteditoren und ihre einfache Konvertierungsmöglichkeiten zu den allgemein geteilten geistlichen Inhalten. Das inhärente Potential in dieser Entwicklung wurde dann von der Internet-Proliferation in den 1990er Jahren verwirklicht. Seitdem ist es möglich, dass was jemand beschreibt, ausdenkt und in der Internet veröffentlicht, wird in einigen Minuten für Hunderten und Tausenden erreichbar und es wird das Erleben und Handeln von Tausenden mitbestimmen.

Kevin Kelly beschrieb diesen Prozess durch zwölf technologischen Entwicklungen wie folgt (Kelly, 2016). Im Zentrum dieser Entwicklungen steht die Entstehung der Flüssigkeit der Sinnfixierung (*flowing*) als Folge der computerisierten Digitalisierung des Schreibens. Auf diese Weise wird es für die sinnhaften Gedanken schaffenden Mensch möglich gemacht, anstatt der früheren rigiden Fixierung der Gedanken und anderer sinnhaften Inhalte die flüssige digitale Fixierung zu benutzen, und so kann man diese Inhalte immer neulich durchdenken und korrigieren, dann die Teile der fixierten Inhalte zerschneiden und sie mit anderen fixierten Inhalte vermischen, damit neue Gedanken, musikalische Werke etc. schaffen zu können. Für die Menschen in den geistigen Teilsysteme wurde von dieser Entwicklung der leichte Aufstieg von der rigiden Sinnfixierung zu dem Zustand des ständigen geistigen Schwebens möglich gemacht. Die Wissenschaftler, Künstler, theoretische Jurist usw. können ihre Gedanken und Ergebnisse aufgrund der digitalen Sinnfixierung immer nur temporär fixieren, die für das Umdenken, Korrigierung usw. immer zur Verfügung stehen. Die computerisierte Digitalisierung ist die Basis für die anderen technologischen Entwicklungen, die von Kelly analysiert wurden, und damit die Erreichbarkeit der Inhalte der flüssigen Sinnfixierung der Einzelnen für die umfassenden menschlichen Gemeinschaften möglich zu werden - und damit diese Sinnfixierung als gesamtgesellschaftlich geteilter flüssiger Geist zur Verfügung stehen zu können – musste einige hervorragende Textverarbeitungsprogramme und ihre leichte Konvertierungsmöglichkeiten erstellt werden und am Ende der 1980er Jahren wurde es von *Word* und einigen anderen Texteditors geschafft. Durch die Entstehung von Internet wurde für die flüssige und leicht vergemeinschaftbare Sinnfixierung in den 1990er Jahren dann möglich gemacht, die ganze geistige Welt umzugestalten. Alle Inhalte haben in den Zustand der ständigen Wandel geraten und alle sind im Zustand des permanenten Werdens (*Becoming*) und damit konnte der früher schon in der Transformation von der traditionellen zu der modernen Gesellschaft erreichte Umbau auf Änderung der funktionalen Teilsysteme (z.B. außer Kraft setzbares Recht im Rechtssystem, wiederlegbare wissenschaftliche Wahrheit, durch Wahlen ersetzbare Staatsmacht usw.) weiterbringen und heute existiert schon fast alle Institutionen und menschliche Sachen in der ständigen Änderung. Die flüssige Sinnfixierung erweiterte dann auf alle Formen der Kommunikation und das brachte anstatt dem Buch der zentrale Stelle des Bildschirms zustande und der Mensch des Buches wurde stufenweise der Mensch des Bildschirms. Der frühere zentral gefüllte TV-Bildschirm verändert sich langsam zu der dezentralisierte Smart-TV, wo aus der vielen Millionen von Inhalten von den einzelnen Zuschauer selektiert wird und parallel führt diese Entwicklung zu dem Computer-Bildschirm dann zu dem Bildschirm von Smartphones, die mehr und mehr Funktion übernimmt und damit wird der gemeinsame multifunktionale Bildschirm von TV/Computer/Telefon/künstliche Intelligenz zustandegebracht. Auf diesen Weg gelangt

unsere gesamte Umgebung stufenweise unter intellektueller Reflexion und anstatt unserem bisherigem passivem Aufsitzen auf die physikalische-biologische Umgebung begann es die Dinge dieser Umgebung mit Kognition zu imprägnieren und immer mehr Dinge zu Smart zu machen und das ist das Prozess von *Cognifying*. Die folgende Entwicklung ist die Back-Überwachung unsere klüger gewordene Dinge, die *Interacting*, und es begann, dass unsere Reaktionen durch unsere Smart-Dinge beobachtet werden und aufgrund der so gewonnenen Informationen ergänzen sie unsere Aktivitäten oder sie beginnen diese Aktivitäten in andere Richtung zu lenken usw.

Von diesen von Kevin Kelly hervorgehobenen neuesten Trends werden die Schwerpunkte der dreien Seinsformen des geistigen Seinsschicht neugeordnet. Der individuelle Geist verwebt sich in den objektiven Geist der Epoche stärker als in der Vergangenheit und er nimmt die geistigen Inhalte der objektiven Geist nicht einfach nur während der frühen Sozialisation - weitgehend für das ganze Leben - aber in täglichen Kontakten und damit sich formierend jeden Tag, bzw. kann der eigene geistige Inhalt durch das Internet sofort für den objektiven Geist erreichbar gemacht werden und dadurch zurück der objektiven geistigen Gehalten modifiziert werden. In ähnlicher Weise stehen die objektivierten intellektuellen Gehalten durch das Internet ständig und kontinuierlich für alle zur Verfügung. Auf diese Weise trennen sich die Gehalte des objektiven Geistes und des (toten) objektivierten Geistes voneinander nicht so stark als in den Jahren von Hartmann in den 1930er Jahren, wenn auch ihre Trennung nicht vollständig eliminiert werden kann.

3. Die direkte Verbindung der künstlichen Intelligenz mit der physischen Seinsschicht

Wie im Ausgangspunkt geklärt wurde, ist der Mensch das Ensemble von vierer Seinsschichten, und hinter jeder geistigen Aktion sind seine psychologischen und biologischen Seinsschichten präsent. Mit dieser Funktionsweise vor den Augen können die Unterschiede zwischen dem Mensch und dem Robot von künstlicher Intelligenz besser geklärt werden. Michio Kaku schreibt in seinem neuen Buch, dass Rodney Brooks ihm in einem Interview gesagt hat, dass der Roboter Maschine ist, wie auch der Mensch es ist, und so können wir eines Tages solche lebende Maschinen bauen, wie wir sind (Kaku, 2014 : 263) Hartmann's Seinsschichten ernstgenommen hätte das nich gesagt werden können, auch nicht, wenn die immer fortgeschrittenen und anspruchsvollen Programme in der Lage sind, zusätzlich zu der intellektuellen Operationen auch die Emotionen und die physiologischen Sinnlichkeiten in Algorithmus konvertieren zu können, um für die Bestimmung der Roboter zur Verfügung zu stellen. Nämlich, von dem Programm können die seelischen Emotionen und die physiologischen Gefühle nur in der geistiger Ebene imitiert werden, aber weil es hinter diesen Imitierungen keine wirkliche seeliche-emotionale Sensschicht und keine bi-physiologischen Mechanismen gibt, kann es nur nachgeahmt werden. Der funktionierende Roboter mit künstlicher Intelligenz kann unvermeidlich nur mit einer zweischichtiges Seinsweise funktionieren und egal wie komplex und durch die Programmierung für die seelischen Reaktionen bzw. die physiologisch-biologische Bewegungen geeignet ist, kann der Roboter nur das Emsemble von zwei Seinsschichten sein im Vergleich der vieren Seinsschichten von dem Mensch. Kaku schildert in seinem Buch diese bereicherte Programmierungsmöglichkeiten und akzeptierend die früher kritisierte Behauptung von Rodney Brooks hebt er die Möglichkeit der Menschenrechte für die Roboter auf, und er schreibt über die ethischen Anforderungen in Bezug auf die Roboter (Kaku, 2014: 250-252).

Im Falle von Roboter wird die geistige Seinsschicht von der Programmierung reproduziert und wenn diese Programmierung immer komplexer sein können, wird es möglich, die niederen Seinsschichten der Menschen in das Programm hinzuziehen. Dann werden auch die Reaktionen der seelischen und eventuell der biologisch-physiologischer Seinsschichten in den Algorithmus programmiert, und das angereicherte intellektuelle Programm kann direkt mit den physikalisch-mechanischen Körper verbunden werden. Eine weitere Manifestation dieser dichten Verbundenheit, wenn im Falle der körperlich behinderten oder anderweitig unbeweglich beschädigten Menschen die Gehirnwellen direkt an gelähmten Körperteile mit Umgehung der beschädigten Teil des Gehirns verbunden werden, und so werden die Funktionen des Gehirns durch ein Programm immitiert und der zuvor gelähmt Mensch wird wieder befähigt, sich zu bewegen. Aber auch ohne diese – wie es im Fall von Stephen Hawking war – können die geistigen Reaktionen der Gehirnwellen des Gelähmten mit einem Rollstuhl verbunden werden und er wird fähig, den Rollstuhl zu bewegen und er kann Objekte in der Außenwelt durch Gehirnwellen bewegen. „Telekinesis: das gelenkte Material durch Geist“ – schreibt Kaku in dem Titel eines Kapitels in seinem Buch, und das ist eine präzise Beschreibung des reduzierten Zustandes von Hawking im Vergleich des vollen menschlichen Seins mit vierer Seinsschichten. Nämlich konnte Hawking nur mit der direkten Zusammenschaltung seiner geistigen Seinsschicht mit der mechanischen Welt kommunizieren und so war er in bestimmter Hinsicht in dem reduzierten Zustand der Zweischichtigkeit. Natürlich mit lebendem Gehirn, und so sollte er gefüttert werden und wegen seinem Stoffwechsel sollte immer Windel gemacht werden. Aber die so geschaffene Technologie kann später auch mit dem schieren physischen Roboterkörper verbunden werden, und das kann die Entstehung solcher zukünftigen Existenz beigetragen, welche ohne die seelische und biologische Seinsschichten in der Welt wirken kann. Die Analyse dieser selbstorganisierenden künstlichen Intelligenz und die Chancen dieser Entwicklung wurde schon - unter anderem - in den Werken von Ray Kurzweil und Nick Bostrom zustandegebracht, so lohnt es sich, in der folgenden Analyse auf der Grundlage ihrer Werke vorher zu gehen. (Kurzweil, 2005; Bostrom, 2014)

Davor jedoch scheint es sich lohnen, einige Schlussfolgerungen von unserer bisherigen Analyse in Bezug auf die Beziehung zwischen der geistigen Seinsschicht und der eventuell über sie entstehenden neuen Seinsschicht der künstlichen Intelligenz zu ziehen. Aufgrund dieser Analyse scheint es, dass im solchen Falle, wenn durch die künstliche Intelligenz nur die bestehende geistige Seinsschicht angereichert wird, und sie benutzt diese Anreicherung als nur zusätzliche Fähigkeiten, dann kann nicht von der Schaffung einer neuen Seinsschicht gesprochen werden. Einschließlich aller von Kevin Kelly beschriebenen Trends werden dann nicht anders sein, als die Stärkung der geistigen Seinsschicht über die niederen Seinsschichten. Sogar, wenn dies nur der Anfang in Bezug auf die Aufbesserung durch die künstliche Intelligenz ist, und die Dinge in unserer Umwelt bekommen immer mehr Smart-Zusätze in den nächsten Jahrzehnten - wie Kelly' Prognose sagt – wird dies nach wie vor nur unsere bisherige vierte und höchste Seinsschicht sein. Eine wirklich neue Seinsschicht kann nur dann sprechen, wenn die Formen der künstlichen Intelligenz, die Algorithmen und zum Programme umgestalteten Geringwellen irgendwie selbstorganisierend werden können und durch direkten Verbindung mit mechanischen Körper auch ohne den menschlichen Geist in der Welt wirken können. Es ist noch eine weitere Frage, ob dies nur die Entfaltung einer weiteren neuen Seinsschicht sein wird, wie es bereits dreimal in den Milliarden der Jahren der Geschichte der Erde stattgefunden hat – und als unverzichtbare Voraussetzung gehalten immer die vorherige Seinsschichten, obwohl nur umgestaltet - oder ist dies Evolutionssprung anders im Vergleich zu den älteren, und dies bekommt einen anderen Verlauf?

4. Die selbstorganisierende künstliche Intelligenz

Ausgehend von den schon zitierten Werken von Ray Kurzweil und Nich Bostrom können zwei Möglichkeiten für die Entfaltung der sich selbst organisierenden künstlichen Intelligenz isoliert werden. Die eine ist die starke künstliche Intelligenz, welche als Nachfolge der heute schon existierenden schwache Version entstehen kann, und die zweite ist die Emulation des menschlichen Gehirns, die als digitale Kopie der geistlichen Inhalte eines menschlichen Gehirns von den Grenzen des menschlichen Seins getrennt autonom existieren kann. Als dritte kann die Möglichkeit der künstlich verbesserten menschlichen Intelligenz sein, die eine Art Super-Intelligenz schaffen kann, obwohl diese nur eine andere Form des heute schon funktionierende Koexistenz zwischen den menschlichen Organen und der zusätzlichen künstlichen Intelligenz sein kann, welche von dem Mensch nicht trennen würde und deshalb keinen Evolutionssprung bedeuten würde, sondern nur die Dominanz der höchsten Seinschicht über die niedere weiter verstärken würde. (So könnte ihre Analyse auch am Ende des vorhergehenden Abschnitts stehen.)

Bevor mit der detaillierten Analyse begonnen wird, muss darauf hingewiesen werden, dass die Entstehung einer starken künstlichen Intelligenz und ihres Potenzials - wie Kurzweil und Bostrom vermuten - von einigen MI-Forschern heftig bestritten wird. Insbesondere John Searle, ein amerikanischer Philosoph, hat die Möglichkeit einer solchen Entwicklung in mehreren Studien von dem Anfang der 1980er Jahre verworfen. Seiner Meinung nach, unabhängig von der Geschwindigkeit und der Speicherkapazität, die für immer fortschrittlichere Computer und ihre Programme zur Verfügung steht, kann dies nicht zu einem Maschinenbewusstsein führen. Somit wird nur schwache künstliche Intelligenz möglich sein. Zum Beispiel wollte er in seinem berühmten "Chinesischen Zimmer"-Experiment zeigen, dass, obwohl ein komplexes Programm jeden Text auf Chinesisch perfekt übersetzen kann, es nur eine Symbolmanipulation ist, die nichts mit einem bewussten Wesen zu tun hat (siehe besonders Searle 1984; 1992). Mit diesem Vorbehalt im Hinterkopf wollen wir uns anschauen, auf welche Weise die Entwicklung einer starken künstlichen Intelligenz aufgrund den Forschungsanalysen realisiert werden kann.

4.1. Die starke künstliche Intelligenz

Die starke Form der künstlichen Intelligenz zeigt solchen Grad der künstlichen Intelligenz, wenn sie das Niveau des menschlichen Geistes erreicht hat, und dann übertrifft sie diesen Grad sehr schnell tausendmal und millionenmal im Gegensatz mit der Form der heutigen schwachen Version. Eine Vorfrage ist, ob es überhaupt möglich ist, und eine so starke künstliche Intelligenz wirklich geschafft werden kann? Durch die Analyse der bisherigen exponentiell schnellen Wachstum der Leistung wird es erlaubt, diese Frage schnell zu beantworten. Ja, das ist möglich und die bleibende Frage ist nur, ob dies um 2040 oder 2100 sein werden kann. Vor diesem Hintergrund taucht zwei wichtige Fragen auf: 1) ob sich die starke Form der künstlichen Intelligenz von der institutionellen Überwachung und Lenkung der Menschen und der menschlichen Gesellschaft befreien kann? 2) und die zweite Frage ist, was für einen Charakter diese außer Kontrolle geratene künstliche Intelligenz haben wird, und ob er ein autonomes Selbstbewusstsein und alldurchdringende Willen haben wird, die ihre enorme Kapazität für die Veränderung der Welt unabhängig von den Willen von Menschen benutzen wird, oder im Gegensatz dazu kann diese alldurchdringende Willen parallel mit der enormen technologischen Kapazität nicht entstehen kann und nur als die geistige

Überblickfähigkeit eines naiven kleinen Kindes - aber mit Verbindung von enormer technologischer Kapazität - beschrieben werden kann.

Was die erste Frage betrifft – das heißt die Möglichkeit des Ausbrechens von KI aus der menschlichen Kontrolle - soll von den Konsequenzen der rekursiven Selbstlernfähigkeit und Selbstveränderbarkeit der genetischen Algorithmen ausgegangen werden, die als Hauptrichtung der Entwicklung der künstlichen Intelligenz gesehen werden kann. Auf diese Weise kann die menschliche Kontrolle nur auf die Bestimmung der Anfangsparameter beeinflussen, aber dann einerseits können solche Lösungen für die Realisierung dieser Parameter von der KI zustandegebracht werden, die von der menschlichen Kontrolle losreißen können, andererseits gelangen auch die Eingangsparameter im Kreis der Selbstlernfähigkeit und der Selbstveränderbarkeit, und nach einigen Zyklen der rekursiven Selbstveränderung wurde jede frühere Bestimmung annulliert. Diese Selbstveränderungsfähigkeit von KI ist schon heute so groß, dass auch die Änderung von ihre Hardware erstellt werden kann. Diese schon bestehenden Kapazitäten bedeuten aber noch nicht ernsthafte Bedrohung, weil aufgrund dem heutigen niedrigen Niveau der KI im Vergleich zu der menschlicher Intelligenz ihre Kontrolle noch geschaffen werden kann. Aber in der Zukunft können so schnelle rekursive Selbsterlernungs- und Selbstveränderungszyklen zustandekommen - im Vergleich zu heute -, dass vielleicht millionenfacher Geschwindigkeit für das Selbstlernzyklus ermöglicht sein wird, das während Stunden, Minuten und Sekunden bis zu hundert Mal pro Tag sinken kann, damit grundlegende Änderungen zu erstellen. Dies kann nicht mehr unter menschlicher Kontrolle gehalten werden und so folgt die Freilassung von KI aus der menschlichen Kontrolle nach einem Punkt einfach von dem heutigen Trend.

Die nächste Frage bezieht sich auf das, welche Natur diese vom Menschen befreite KI haben wird. Für die Einschätzung lohnt es sich, eine Unterscheidung zu machen, und innerhalb der KI soll die technologische Intelligenz von der allgemeinen Intelligenz für die Beurteilung der gesellschaftlichen Zustände getrennt werden. Die technologische Intelligenz ist die Fähigkeit der gezielten Änderung der biologischen und physikalischen Welt und ein Aspekt dieser Fähigkeit bezieht sich darauf, in welchem Masse diese Änderung von anderen Kräften (z.B. von Menschen) unverhinderbar ist. Dies ist also die Fähigkeit für die Herrschaft der niederen Seinsschichten und das erhöht sich mehr und mehr auf dem Gebiet der künstlichen Intelligenz, während ihre Fähigkeit für das Durchschauen der globalen Realität, einschließlich der Realität der menschlichen Gesellschaft weit zurückgeblieben ist. Nick Bostrom analysierte schon ausführlich die Bedeutung des Hineinbaus der sozialen Wissensarte und Werte in die KI-Programme und auf dieser Grundlage können hier auf zwei Probleme hingewiesen werden. Ein Problem ist, dass es für die Grundstruktur der menschlichen Gesellschaft und ihr Überleben kein einheitliches Ziel-Parameter-System gibt, und es hängt immer von der Selektierung der einzelnen dominanten menschlichen Elitgruppen, welche Art von Werthierarchie gelten wird. Aber es ist das kleinere Problem. Das größere Problem ergibt sich daraus, dass auch diese willkürliche und eventuell unvollständige Werthierearchie der Opfer der Selbstveränderungszyklen von KI werden kann. Wenn die KI selbst die bestimmende Komponente ihrer Programmierung immer wieder verändern kann, gibt es keine Garantie, dass die eingebauten sozialen Werte in dem technologischen Parameter der KI bleiben werden und von der KI nach wenigen Zyklen der Selbstveränderung nicht gelöscht werden.

Also, wenn wir das Wesen unserer Unterscheidung genauer bestimmen wollen, dann kann sie als die Opposition der technologischen Intelligenz zu der sozialen Intelligenz ausgedrückt werden. Während die künstliche Intelligenz in der technologischen Dimension riesig wächst, bleibt sie in der Dimension der sozialen Intelligenz auf dem Niveau eines dummen kleinen Kindes. Und wenn für die Aufbesserung beider Arten der Intelligenz und für die Behandlung der sozialen Werte die speziellen Algorithmen in das Programm eingebaut werden, dann gibt

es keine Garantie, dass nach kurzer Zeit durch die selbstlernende KI diese Teile des Programms nicht gelöscht wird. Folglich ist nach meiner Meinung nicht die oft wiederholte Beschreibung der „böse gewordenen künstlichen Intelligenz“ geeignet für die Signalisierung der realen Gefahren, sondern wir müssen vor die sozialen blind und stumm und dennoch leistungsfähige technologischen Intelligenz Angst haben. Diese künstliche Intelligenz wird die Existenz der menschlichen Gesellschaft nicht wegen ihres Bösen zerstören, wenn durch ihre enorme technologische Kapazität das ermöglicht wird, sondern wegen ihres niedrigen Niveaus in Bezug auf das soziale Wissen. Es ist jedoch zu beachten, dass sich diese Lage in den kommenden Jahren ändern kann, weil von den jüngsten Entwicklungen im Bereich der Emulation des menschlichen Gehirns (Mind Upload) auch die Zukunft der starken künstlichen Intelligenz berührt werden kann (siehe nächster Abschnitt).

Betonend diese Zusammenhänge über die Gefahren der künstlichen Intelligenz versuchen die KI-Forscher mindestens in die Anfangsparameter des Programms der KI solche operative Prinzipien einzubauen, welche die gefährlichen Veränderungen für die menschliche Gesellschaft verhindern können. In einem neu erschienenen Buch von Joel und Ben Goertzel wird es versucht, solche Programmelemente für die Planung der künstlichen Intelligenz zu finden, von denen die ständige positive Tendenz in Richtung auf die Beibehaltung der menschlichen Gesellschaft gesichert werden kann. Sie gehen davon aus, dass dieses Ziel nicht vollständig gewährleistet werden kann, aber das kann bestimmt werden, dass mindestens die Selbstveränderungen aufgrund des maschinellen Lernens in diese Richtung immer ein positives Feedback erhalten. „Our current perspective is that provably, or otherwise guaranteeably, Friendly AI is not achievable. On the face of it, achieving strong certainty about the future behaviors of being massively more generally intelligent than ourselves seems implausible. Again, we are aiming at a more modest goal – to explore ways of biasing the odds, and creating AI systems that are significantly more likely than not to be Friendly.” (Goertzel/Pitt, 2014:65.) Um die „freundliche“ Richtung der Maschinenintelligenz in Bezug auf die menschlichen Gesellschaften sicherzustellen, betonen sie als erstes Imperativ, die rekursiven Veränderungszyklen in der ersten Periode am langsamsten zu halten, damit sie noch zugeschnitten zu der Verständnisfähigkeit des menschlichen Geistes zu sein. Außerdem ist es noch wichtig, dass die menschliche Teilnahme in der Bestimmung des Programmveränderungen in den ersten Zyklen sichergestellt werden sollte. Ebenso halten die Verfasser es wichtig, gut ausgearbeitete ethische Prinzipien mit reichen Fall-Szenarien in das Programm hineinzubauen, die in den ersten Phasen durch tausenden Testsituationen ausgeführt werden müssen. „Where, after an AGI (Artificial General Intelligence) has learned some of the everyday aspects of justice, including the balance of justice with empathy in everyday life, and once it has also gotten familiar with the application of abstract ethical principles to other aspects of ordinary life, it will be well poised to appreciate ethical principles and their utility in making difficult decisions. It will be able to understand the abstract nature of justice in a richer and more holistic way.” (Goertzel/Pitt, 2014:72.) Letztlich kann man aber nach wie vor nur hoffen, dass von der vollautomatisch gewordenen und für große Veränderungen fähigen, starken künstlichen Intelligenz die Voraussetzungen für das Funktionieren der menschlichen Gesellschaft nicht beseitigt wird.

Danach lässt sich fragen können, wenn der Umgang mit der starken künstlichen Intelligenz so selbstverständlich eine Bedrohung für die gesamte Menschheit – neben ihren guten Segen und Nutzen – mit sich bringen kann, warum nicht diese ganze Tätigkeit gestoppt werden?! Für dieses Dilemma kann ausgezeichnete Analyse in dem Buch von Nick Bostrom gefunden werden und auch Ray Kurzweil hat sich damit beschäftigt. Namentlich bedeuten die Forschungen im Bereich der künstlichen Intelligenz Forschung seit Jahrzehnten auch die militärische Rivalität der Großmächte und riesige Mengen an Geld und Forschungskapazität werden darauf gewendet. So sind die neuesten Erfolge in diesem Bereich immer geheim

gehalten. Im Grunde passiert jetzt das gleiche in diesem Bereich, was in den späten 1940er Jahren in Bezug auf die nuklearen Forschung gesehen werden konnte und welche große Macht ein Vorteil in diesem Bereich erreichen kann, wird sie in der Lage sein, ihre Dominanz in der ganzen Welt zu sichern. Auf diese Weise würde der Verzicht auf die Forschung im Bereich der gefährlichen starken künstlichen Intelligenz unter den schärfsten Konkurrenz nur bedeuten, einen Wettbewerbsvorteil zu Rivalen zu geben. Ein verwandtes Problem ist, dass selbst wenn nach einiger Zeit die rivalen Großmächte wegen der wachsenden Gefahr einige Bemühungen für die gemeinsame Kontrolle zu leisten beginnen, auch dann wäre es nicht genug, weil es unmöglich ist, das Erreichen der Qualität in der Entwicklung von künstlichen Intelligenz zu beurteilen, wo der Drehpunkt von der heutigen schwachen künstlichen Intelligenz schließlich zu ihrer starken und unkontrollierbaren Nachfolger sein wird. Sowohl in der Analyse von Kurzweil als auch die von Bostrom bringt die Zeitalter der starken künstlichen Intelligenz im Vergleich der Ära von Kernkraft tausendfach größere Segnungen und umfassendere Änderungen am aktuellen Zustand der menschlichen Gesellschaft mit sich und die wachsende Entwicklung in der Robotik zeigt schon diese Änderungen bereits im Alltag. Aber auch die Bedrohungen sind tausendmal größer, als es für die Atomzeit in Bezug auf die früheren kriegerischen Rivalitäten behauptet wurde. Jedoch ist es unmöglich, auch diese zu stoppen, wie es nicht möglich war, das Rennen für das Atom zu stoppen. Das heutige Zeitalter des Rennens für die starke künstliche Intelligenz unterscheidet sich von der früheren Rivalitäten nur dadurch, dass die Konkurrenz im Bereich der Atomwaffen schließlich befriedet werden konnte und die Atomkräfte durch die gegenseitige Abschreckung eine gemeinsame Kontrolle über die Atomwaffen mehr oder weniger stabilisieren konnten, aber im Falle des Zustandbringens der starken KI wird auch ihr Hersteller in kurzer Zeit in dem Hintergrund geschoben und auch er wird gemeinsam mit der ganzen Menschheit bedroht. Es wird nicht mehr sein, der die neue dominierende Kraft über die Welt befrieden könnte.

4.2. Die digitale Existenz des emulierten menschlichen Gehirns

Die eventuelle Entstehung der starken künstlichen Intelligenz aus ihrer schwachen und schon existierenden Form hat schon seit mehreren Jahrzehnten diskutierte Möglichkeit, im Gegensatz zu dem Hochladen des Geistes auf Computer oder in anderer Benennung die Emulation des Gehirns, das nur in den letztem Jahrzehnt im Zentrum des Interesses kam. Dies ist auch eine maschinelle Intelligenz, und es kann ein Zweig der künstlichen Intelligenz sein, und hier kommen auch andere Anreize zusätzlich zu den früheren Motiven, um die Forschung so schnell wie möglich zu beschleunigen. In der Tat, wenn die Emulation des ganzen menschlichen Gehirns erfolgreich sein könnte, und anstelle von den biologischen Prozessen auch von den computerisierten Prozessen die neuronalen Prozesse des Gehirns funktioniert werden könnten, dann würde diese technologische Neuheit eine Perspektive geben, damit der Geist und die Persönlichkeit der Menschen von den verschwindenden biologischen Körper loszuwerden und in einem ewigen Träger existieren für immer. megszabadulva kizárólagosságától durch Exposition gegenüber dem Verschwinden biologischen Körper trägt - ein ewiger átkerülhetne ebenfalls unterstützt. Die Film mit Titel von Transzendenz im Hauptrolle von Johnny Depp hat diese Möglichkeit vor einigen Jahren zum Thema gemacht, und die Forschung geht hier am intensivsten vorher. Zum Beispiel wurde in den letzten Jahren von der Europäischen Kommission eine Summe von eineinhalb Milliarden Euro für Forschungszwecke gegeben, und die Gehirn-Emulation der einfachsten Organismen und der kleinen Säugetiere wird schon gemacht und zur gleichen Zeit wird auch die Emulation des menschlichen Gehirns geforscht.

Die unabhängige Existenz der Emulation des menschlichen Gehirns ist auf dem derzeitigen Niveau der technischen Bedingungen noch weniger sichtbar. (Eine neue Nachricht in der Weltpresse hat angekündigt, dass in 2017 die volle Emulation des Gehirns einer Ratte zu erwarten ist.) Auf diese Weise kann eine empirisch fundierte Aussage über die Möglichkeit der autonomen Existenz der Emulation des menschlichen Gehirns bzw. über den Grad der Identität einer solchen Gehirn-Emulation mit des ursprünglichen Gehirns noch nicht treffen. Ohne die Existenz von realen Emulationen kann man diese Fragen nur auf der Ebene der philosophischen Meditation durchdenken. Natürlich ist dies nicht ohne Nutzen, aber in jedem Fall kann es nur spekulative Natur haben, und dies sollte im Folgenden beachtet werden.

Im Falle der Emulation des menschliche Gehirns ist es noch eine Vorfrage, ob die technische Machbarkeit selbst (Speicherkapazität und Geschwindigkeit) erstellt werden kann, die für die Emulation der vielen Milliarden von Gehirnzellen und der Billionen ihrer Synapsen erforderlich ist. Aufgrund den Analysen über die bestehenden exponentielle Wachstumsgeschwindigkeit in diesem Bereich kann aber behauptet werden, dass etwa in dreißig Jahre der Entwicklung die erforderliche computerisierte Kapazität für diese Aufgabe zugänglich sein wird. Aufgrund einer neuen Nachricht konnte eine Sekunde der ausführlichsten Emulation eines menschlichen Gehirns (d.h. auf der Ebene der neuronalen Prozesse der Gerinzellen), die in dem computerisierten Format konvertiert wurde, in dem schnellsten Computer der Welt in 40 Minuten wiederholt werden. Das heißt, heute braucht eine Sekunde der Gehirn-Emulation noch die Laufzeit von 2400 Sekunden in dem Computer, um das replizierte Gehirn funktionieren zu können, und dies kann ein Entmutigungseffekt verursachen. Aber wenn in der Zukunft die Gültigkeit des Mooreschen Gesetzes weiterhin angenommen wird, d. h. die Rechenleistung auf ca. in einem Jahr und einer Hälfte immer neulich verdoppelt wird (was zum Beispiel der rasche Fortschritt in den Quantencomputern sichern kann), dann kann die 2400-fache Beschleunigung über 15-16 Jahre erreichen, und so die menschlichen Gehirnprozesse können auch in einer angemessenen Zeit auf dem Computer ausgeführt werden. Die Haupt-debatte sollte eher sein, ob nach der erfolgreichen Emulation aller Inhalte des menschlichen Geistes auch der ursprünglichen Bewusstsein dieses Geistes im Computer wiederholt werden kann, und während dem Lauf des Computerprogramms der Emulation immer auch das ursprüngliche Bewußtsein auftaucht?

Die sehr intensive Reflexion und Debatte über dieses Thema hat einige kluge Distinktionen in den letzten Jahren hervorgebracht. Aufgrund dieser Distionkitionen soll unterschieden werden, ob die einzelnen geistigen Prozesse nach der Emulation des Gehirns in dem Computer laufen werden können, und das ist andere Frage, ob neben diesen computerisierten Prozessen als ihre Kumulation ein Bewusstsein entstehen wird, gewissermaßen als eine Co-Betrachtungsposition über die parallel laufenden mentalen Prozesse? Schließlich taucht eine dritte Frage auf, wenn ja, wird dieses emulierte Bewusstsein das Bewusstsein des ursprünglichen Geistes, oder dadurch wird eine neue Identität entstehen, die nur so viel Gemeinsamkeit mit dem ursprünglichen menschlichen Geist haben wird, dass sie beide die gleichen Erinnerungen, Erlebnisse und Denkstile haben werden. Im letzteren Fall kann der emulierte Geist als ein digitaler Zwilling Bruder verstanden werden, aber als die eineiige Zwillinge haben separates Bewußtsein, so auch das Bewußtsein des digitalen Geistes wird nach der Emulation seinen eigenen Weg und eigene Identität autonom gestalten.

Für Kurzweil und Bostrom ist die Antwort auf diese Frage selbstverständlich, weil ihrer Meinung nach alle Manifestation des Geistes und alle psychischen Prozesse inden Nervenzellen des Gehirns die Ergebnisse der elektrochemischen Prozesse sind, und auf diese Weise können auch das Bewusstsein und das Selbstbewusstsein nicht anders sein. Daraus folgt also, wenn die detaillierte Emulation genügend genau war, dann werden nicht nur die einzelnen neuronalen Prozesse des Gehirns (Erinnerungen, Erfahrungen, etc.) sondern auch das Bewusstsein als die Kumulation dieser Prozesse im Laufen des computerisierten Proagmms

erscheinen. Aber in Bezug auf die Frage, ob dieses Bewusstsein eine Verdoppelung der ursprünglichen oder dagegen eine neue Schöpfung sein wird, kann in ihren Bücher keine Analyse gefunden werden, und diese Frage selbst wurde nur neuerlich in der Diskussion aufgeworfen.

Durchgelesen viele Diskussionen und Argumenten über diese Fragen bin ich eher geneigt, zu akzeptieren, dass, wenn genügend genaue und detaillierte Emulation des Gehirns auf Computerplattformen kopiert wird, und alle Verbindungen von Billionen der neuronalen Prozessen digital dupliziert werden konnte, dann wird auch das Bewusstsein als Zentrum der Kontrolle der mentalen Prozesse erstellt. Dies wird jedoch lediglich der Geist eines digitalen Zwillingbruders darstellen, aber sicherlich ist es nicht zu sagen, dass jetzt das gleiche Bewusstsein „in zwei Plätzen“ präsent sein könnte. Und schon gar nicht die Tatsache, dass der Mensch - befreiend von seinem biologischen Körper – durch die Emulation in die digitale Existenz umziehen könnte.

Auch David Chalmers ist für diese Position in einer kürzlich veröffentlichten Studie, während Massimo Pigliucci für die Exklusivität des Bewusstseins plädiert, das mit dem biologischen Körper verbunden ist. Bewusstsein plädieren. Chalmers nennt sich selbst ein Funktionalist und Pigliucci ein Biologist und die beiden Positionen werden wie folgt beschrieben: “Here philosophers divide into multiple camps. Biological theorists of consciousness hold that consciousness is essentially biological and that no biological system can be conscious. Functionalist theorists of consciousness hold that what matters to consciousness is not biological makeup but causal structure and causal role, so that a nonbiological system can be conscious as long as it is organized correctly.” (Chalmers, 2014:104). Dieser derzeit nur philosophische Debatte hat auch deshalb eine Relevanz, weil heute - und nach Ansicht der Forscher noch für einige Jahren – für die Emulation des Gehirns nur destruktive Techniken existieren, die wegen der Verwendung von Tierversuchen (jetzt abgesehen von einigen Tierrechtsgruppen) nicht ein Problem bedeuten. Aber es wurde schon in den Diskussionen aufgeworfen, dass für die unheilbaren Patienten im Endstadium als eine Chance für das Überleben die Möglichkeit der Emulation erreichbar gemacht werden sollte, weil in diesem Fall die zerstörerische Natur der Emulation kein Problem mehr bedeuten würde. So ist es wichtig zu betonen, dass durch diesen Weg höchstens nur ein digitaler Zwilling erstellt werden kann, aber die Verschwinden des ursprünglichen Persons so nicht vermieden werden kann.

Es ist auch eine große Schwierigkeit in dem Emulationsprozess des menschlichen Gehirns aufgrund der neuesten Forschungsergebnisse, dass im Gegensatz zu vieler Hoffnung es unzureichend ist, wenn nur die höheren geistigen Prozesse emuliert werden, weil in fast allen Gehirnprozessen gleichzeitig die verschiedene Teile des Gehirns teilnehmen: “To summarize, it is misleading to refer to areas of brain as if they were modular. Instead, it is impossible to draw strict boundaries separating, for example, the more primitive sensorimotor functions from higher cognition such as planning, judgment, decision-making, and direction of attention. Furthermore, these areas are highly integrated functionally, and anatomically, in a complex dense network. Thus, we conclude that the aforementioned “quintessentially human” (which, employing the parlance of the first section, would fall under the symbolic as opposed to the physical) functions cannot exist independently of brain structures that are devoted to world interaction and body control. Even for an uploaded individual, a body (human-like or otherwise) would remain a necessity.” (Linssen/Lemmens, 2016:5).

Es sollte noch erwähnt werden, dass, wenn von der erfolgreichen Emulation des menschlichen Gehirns gemeinsam mit allen mentalen Prozesse auch ein Selbstbewusstsein erstellt wird (wie als wahrscheinlich angenommen wurde), dann kann auch die Zukunft der maschinellen Intelligenz mit einer neuen Perspektive analysiert werden. Denn, was auf theoretische Ebene über die soziale Realität nicht formuliert werden kann, das in dem Geist

der Menschen, die für die bestehende gemeinschaftliche Solidarität sozialisiert wurde und unter dem Druck der biologischen und seelischen Instikte lebten bzw. ihr praktisches Wissen dadurch geformt wurde, ausführlich beinhaltet ist. Auf diese Weise werden auch dieses soziale Wissen, Erlebnisse und Gefühle mit der Emulation in den digitalen Geist transportiert, und wenn meine - mit Chalmers geteilte – Annahme richtig ist, entsteht auch das digitale Selbstbewusstsein im Geist dieser mit angemessenen sozialen Wissen funktionierenden künstlichen Intelligenz. Diese auf dem Niveau der Menschen bestehenden KI, die sich von den biologischen Barrieren befreit hat, kann dann sofort die tausendfache Geschwindigkeit der Entwicklung in dem digitalen Plattform erreichen und die starke KI kann auch in diesem Weg erstellt werden. Auf diese Weise kann die Weiterentwicklung der menschlichen Emulation

auch als die zweite Weg für das Erreichen der starke KI aufgefasst werden, und dies wiederum beseitigt das Problem der künstlichen Intelligenz, dass sie nur über halbseitige Intelligenz verfügen kann: über riesige technologische Intelligenz und parallel dazu nur über das sozial Wissen eines naiven kleinen Kindes. Nick Bostrom analysierte schon diese Möglichkeit in seinem Buch über die Superintelligenz, und auch er beurteilte diese Möglichkeit als positiv, weil auf diese Weise mit dem menschlichen Geist auch die entwickelten moralischen Instikte in die KI hineingeholt werden. Aber er hat es richtig erwähnt, dass die auf genetische Algorithmen basierte starke KI durch ihre rekursive Selbstveränderungskapazität jederzeit dieses angemessene soziale Wissen aus ihrem operationellen Programm entfernen kann. (Bostrom, 2014: 328-330). So kann auch dieser Weg der Entstehung der starke KI keine endgültige Sicherheit bedeuten.

4.3. Die Fragen der genetisch verbesserten Superintelligenz

Im Bereich der Intelligenz-Aufbesserung und der Verlängerung des Lebensdauers durch die human Biotechnologie scheint es mir sinnvoll drei Richtungen in Hinblick auf der Machbarkeit zu analysieren. Als Gegenseite, die ich am strengsten verbieten würde, sind das Klonen der Mensch und das Schaffen der Mischung von Mensch und Tier als Chimären durch die genetische Veränderung von DNA. Aufgezählt diese drei sind wie folgt: 1) die Intelligenzserhöhung durch Embryo-Selektion, 2) die Erneuerung der inneren Organe durch die im Blutstrom zirkulierenden Nanostäbe, 3) und schließlich die Verstärkung der Kapazität des biologischen Gehirns durch Interfaces. Es soll angedeutet werden, dass diese Fragen schon jenseits den engeren akademischen Gruppen der Forscher der künstlichen Intelligenz auch in den ferneren Kreisen der Philosophen Aufmerksamkeit erregt haben und die ersten Diskussionen haben bereits stattgefunden. Peter Sloterdijk hat im Jahr 1999 aufgeworfen - auf der Grundlage der bisherigen Fortschritte in der menschlichen Gentechnik -, dass die Aufbesserung des menschlichen Seins auch auf diese Weise in der Zukunft erhöht werden könnte, und auf diese Analyse reagierte Jürgen Habermas empört und ohne das tatsächliche Nachdenken mit der stärksten moralischen Verurteilung (Sloterdijk 1999; Habermas 2000). Natürlich passiert diese emotionale Diskussionsart über diese Fragen im Falle der Deutschen - ohne inhaltliche Analyse und nur auf der Ebene der schieren Empörung - wegen ihrer spezifischen historischen Erbschaft, weil die früheren Bemühungen im Bereich der menschlichen Eugenik von dem nationalsozialistischen Staatsmacht gemacht wurde und für die deutsche intellektuelle Elite wurde von dieser Erbschaft die human Gentechnologie am tiefsten diskreditiert und das ganze Thema wurde tabuisiert. Francis Fukuyama beschäftigte sich mit diesem Thema in einem ganzen Buch im Jahr 2003 auch und auch er hat die Benutzung der humanen Gentechnologie abgelehnt, aber gegenüber dem schieren normativen

Ton von Habermas hat er diese Fragen empirisch analysiert. Die Ursache für seine Weigerung gegenüber der humanen Gentechnologie besteht darin, dass er die bereits bestehende Ungleichheit unter den Menschen und sozialen Gruppen in diesem Weg weiter zu erhöhen denken, und er denkt das als nicht annehmbar. Seiner Meinung nach entstehen auf diese Weise die Gruppen der genetischen Reichen gegenüber die genetischen Armen, und in der Zukunft würden von den reichen Erblässers nicht nur ein großer Reichtum und bessere Lebensbedingungen für ihre Nachkommen hintergelassen, sondern auch die genetisch verbesserte Körper und Geist und auf diese Weise wird die Ungleichheit zwischen den sozialen Gruppen mehr als je zuvor erhöht (Fukuyama, 2002: 208-210).

Gegenüber Fukuyamas Angst vor der Erhöhung der Ungleichheit ist es für mich Kurzweils Behauptung überzeugender, der davon ausgeht, dass nach den anfänglich hohen Kosten die human biotechnologischen Prozesse später in großem Maße billiger werden - denn es gibt kaum ihre Material- und Energiekosten – und auf diese Weise werden sie nach einer Weile ein Routineverfahren. Darüber hinaus soll es noch vor Auge halten, sofern eine Technologie kein Schaden für andere verursacht, und nur das Niveau ihrer Benutzer von der Ebene anderer Menschen aufhebt, wird sie sich unaufhaltbar unter der Elite verbreiten, und so ist es vergeblich, es zu bekämpfen. Als gleichheitsfreudlicher Mensch sollte von der Analytiker lieber bevorzugt werden, die Regierung für die staatliche Unterstützung der Inanspruchnahme zu drängen, um auch für die ärmeren gesellschaftlichen Gruppen diese technologische Möglichkeit erreichbar werden zu können. Durch die vielfache Embryo-Selektion (gewonnen von den Stammzellen) - wie es Bostrom analysierte - könnte innerhalb wenigen Jahren riesige Intelligenz-Wachstum massenhaft in der Gesellschaft erreicht werden und das wäre der Höhepunkt der institutionalisierten sinnhaften Gesellschaft. Für mich besteht das Problem darin, dass da es heute schon technisch möglich geworden ist, so die Gesellschaft oder Gesellschaften, die eine schnellere Schritt auf diesen Weg machen, gewinnen einen unbesiegbaren Vorteil gegenüber den Hintergebliebenen. Die geistige Elite eines Landes sollte mehr förderlich dafür sein, um diese technologische Möglichkeit stärker unter den Gruppen der Gesellschaft bekannt zu werden, und durch öffentliche Debatten die staatliche Regulierung aufzustacheln, die Verwendung von Embryo-Selektion für die Intelligenz-Wachstum massenhaft in der Gesellschaft zu unterstützen.

Die unterstützbare zweite Richtung in diesem Bereich ist die ständige Verjüngung der menschlichen Organe durch die Nanobots in dem Blutkreislauf und diese Bemühungen sind eher nur auf der Ebene der Forschung und Tierversuche im Vergleich zu der früheren Richtung, aber aufgrund den Analysen von Kurzweil und Bostrom über die exponentiellen Geschwindigkeit des Wachstums in diesem Bereich lässt sich wenig Zweifel haben, dass es hier in den kommenden Jahren den radikale Durchbruch geben wird. Aber im Gegensatz zu der früher analysierten Richtung soll in diesem Bereich größere Skepsis ausgedrückt werden, weil die schiere Langlebigkeit verbunden mit der stufenweise Versteifung der ganzen Persönlichkeit schon heutzutage großes Problem verursacht. Schon heute bedeutet es für die meiste 80-90 Jährigen das Hauptproblem, dass ihr verlängertes Leben sinnlos wurde, und es lässt sich vorstellen, wie dieses Problem im Falle von einem 130-150 Jahre Leben erscheinen wird, auch wenn einigermaßen die Flexibilität der Persönlichkeit und Habitus mit wenigen Jahren verlängert werden kann.

Zum nächsten Thema der Intelligenzaufbesserung angekommen, muss herausgehoben werden, dass die Methoden für diese Richtung in der Forschung und der Tierversuche schon ausgearbeitet sind, und z.B. in Maus-Experimenten die Gedächtnisaufbesserung mit eingebauten Hippocampus-Chips ausgezeichnet funktionieren konnte. Darüber hinaus haben schon die Experimenten in Bezug auf die Menschen auch begonnen, zuerst nur mit dem Zweck der Heilung der Alzheimer-Krankheit. Die exponentiell schnelle Entwicklung –

gemeinsam mit der Verbreitung der den mehrfachen Embryo-Selektionen - kann in der Zukunft in die Tat eine beispiellose Intelligenzerhöhung in der ganzen Gesellschaft schaffen.

Insgesamt lässt sich zu sagen, dass die Intelligenzerhöhung durch die human Biotechnologie grundsätzlich begrüßt werden kann - ausser den erwähnten zwei strengen Verboten - und für die Erzielung der erhöhten künstlichen Intelligenz in der Gesellschaft in erster Linie diesen Weg unterstützt werden sollte, wie es ein Forscher zu diesem Thema schreibt: "Thus, the best approach to life extension and consciousness expansion might lie in our own marvelously complex and entire bodies, meshed with and augmented by tiny bionan machines that become a part of us, rather than the opposite vision of human migration into a machine substrate. You might grow own eternal, artificial self as you gradually become bionic, in stages so tiny that you do not even notice." (Goonan, 2014:198).

5. Eine neue zusätzliche Seinsschicht oder der Neubeginn der irdischen Evolution auf der Grundlage der selbstorganisierenden künstlichen Intelligenz?

Im Spiegel der der bisherigen Analyse soll die zwei Formen der von menschlichen Inspektion abtrünnigen und selbstorganisierenden künstlichen Intelligenz noch einmal berücksichtigt werden, das heisst die starke maschinelle KI und die durch Emulation des menschlichen Gehirns erstellte KI. Es war aus dem Vorstehenden ersichtlich, dass die digitalisierte und selbständige Seinsschicht dieser zwei Formen allein von den computerisierten Geräten getragen werden können, aber sie auch mit physischen Körpern direkt verbunden werden können - vielleicht immer nur für eine kurze Zeit - und dies letzterer Weg ermöglicht für sie ein Sein mit zweien Seinsschichten, d.h. mit geistigen und physischen Seinsschichten aber ohne die biologischen und seelischen Schichten. Auch wenn die geistige Abdrücke dieser letzteren Schichten im Falle der emulierten Arte der starken AI in das Programm hineingeholt wurde, würde dies keine funktionelle Rolle mehr - nur behindernde Wirkungen - so die Beseitigung dieser Programmteile ist durch die rekursiven Selbstveränderungszyklen so gut wie sicher. Das durch die Emulation digitalisierte Selbstbewusstsein, das zuvor von dem biologischen Triebe gebildet und fortwährend von den biologischen Seinsbestimmungen formiert wurde, bzw. seit der frühen Kindheit von der familiarischen und anderen Solidaritätsgefühle durchgedrungen wurde, gelangt ins Vakuumraum nach der Emulation in dem digitalisierten Sein in Bezug auf seine biologische und seelische Bestimmungen. Die biologisch-physiologische Reaktion-Erinnerungen bleiben noch während einiger Zeit - wie der Mensch mit amputierten Bein haben noch der Juckreiz in Bezug auf seinen nicht existierenden Bein - und ebenso können seine seelische Dispositionen etwas bewirken, aber sie hätten keine wirkliche funktionelle Rolle mehr. So ist die Wahrscheinlichkeit des Verschwindens dieser Teile aus dem emulierten Selbstbewusstsein in der zur Zweischichtigkeit reduzierten Welt der künstlichen Superintelligenz sehr hoch.

Auf die Frage, die von dem Titel des letzten Teils dieser Schrift gestellt wurde, kann also die Antwort nur sein, dass die künstliche Intelligenz als Evolutionssprung keine solche Seinsschicht in der irdischen Evolution zustandebringen kann, wies es schon dreimal durch die Milliarden von Jahren in der Vergangenheit bereits passierte. In der Tat, solange die KI die wachsende Macht der intellektuelle Seinsschicht über die niedrigen Seinsschichten nur weiter verstärkt, entsteht keine Neuheit in Bezug auf die Seinsweise der vergangenen tausenden Jahre und nur die Dominanz der intellektuellen Seinsschicht erhöht sich viel schneller. Wenn aber die so entstehende digitalisierte Intelligenz sich von der Einwirkung der Menschen vollkommen befreit kann und als eine selbstorganisierende Kraft zu funktionieren beginnt, dann kommt es zu einem Evolutionssprung - das evolutionäre Endergebnis des

Funktionierens der niederen Seinschichten erhöht sich über die bisherigen Schichten -, aber im Gegensatz zu den vorherigen wird die neue evolutionäre Kraft nicht mehr auf der Grundlage der niedrigeren Seinschicht funktionieren. Aufgrund der Analyse von Hartmann kann es beschrieben werden wie folgt. Der biologische Schicht hat sich über physikalisch-mechanische Seinschicht während der Milliarden von Jahren übergebaut, dann hat auch die seelischen-emotionalen Seinschicht in den höheren Graden des biologischen Lebens erschienen und dann erschienen die Anfänge der geistigen Seinschicht auf dieser Basis in erster Linie mit den Primaten aber dann besonders mit den Menschen, und aufgrund der Koexistenz dieser vier Seinschichten wurde schließlich die digitalisierte künstliche Intelligenz geschaffen, die vom Menschen getrennt auch direkt sich mit dem physischen Körper verbinden kann, und auf diese Weise in einer selbstorganisierende Seinsweise funktionieren kann. Das ist tatsächlich ein neuer Evolutionssprung, aber diese neue, selbstorganisierende geistige Seinschicht baut sich nicht über die bisherigen vier Seinschichten als einen neue Etage über die andere, sondern sie kann sich direkt mit der physikalisch-mechanischen Seinschicht verbinden und Wirkungen ausüben. Von der Koexistenz der irdischen Seinschichten wird die Geburt der künstlichen Intelligenz ermöglicht - und das ist ihre einzige Weg ins Leben zu kommen - aber nach der Möglichkeit ihres unabhängigen Seins werden diese Seinschichten für sie überflüssig. Auf diese Weise kann die neue, selbstorganisierende künstliche Seinschicht nicht eine neue über die vorherigen, sondern nur ein Neubeginn der Evolution auf der physikalischen Ausgangsbasis sein, aber jetzt im Niveau der schon erreichten selbstorganisierenden Intelligenz und mit der Führung dieser neuen evolutionären Kraft.

Die große Debatte und Angst über die Gefahren von KI ist also legitim, wie es ständig wiederholt von Stephen Hawking, Nick Bostrom und Elon Musk aufgeworfen wird, weil die ganze biologische Seinschicht - darin die Menschheit - für die künstlichen Intelligenz als die neue Seinschicht überflüssig sein wird. Die Legitimität dieser Ängste akzeptierend, muß nur darauf hingewiesen werden, dass diese durch die menschlichen Evolution hergebrachte neue künstliche Seinschicht nicht auf die Existenz auf der Erde angewiesen sein würde, und in einer Reihe von Nachbarplaneten ungehindert funktionieren könnte, wie die möglichen Kolonisation des Universums durch die künstliche Intelligenz von Kurzweil schon erläutert wurde (Kurzweil, 2005:433-564).

Literature

- Bostrom, Nick, *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.
- Chalmers, David, "Uploading: A Philosophical Analysis", in: Russel Blackford, Damien Broderick (eds.), *Intelligence Unbound: The Future of Uploaded and Machine Minds*, Wiley Blackwell, Malden-Oxford. 2014, pp. 102-118.
- Elias, Norbert, *Über den Prozeß der Zivilisation: Soziogenetische und psychogenetische Untersuchungen*. Suhrkamp. 1976.
- Fukuyama, Francis, *Our Posthuman Future: Consequences of the Biotechnology Revolution*. Picador Edition, 2002.
- Goertzel, Ben, Joel Pitt, "Nine Ways to Bias Open-Source Artificial General Intelligence Toward Friendliness", in: Russel Blackford, Damien Broderick (eds.), *Intelligence Unbound: The Future of Uploaded and Machine Minds*, Wiley Blackwell. Malden - Oxford. 2014. pp. 90-101.
- Goonan, Kathleen Ann, "The Future of Identity: Implications, Challenges, and Complications of Human/Machine Consciousness", in: Russel Blackford, Damien Broderick (eds.), *Intelligence Unbound: The Future of*

- Uploaded and Machine Minds*. Wiley Blackwell, Malden – Oxford. 2014. pp. 193-200.
- Habermas, Jürgen, *Die Zukunft der menschlichen Natur. Auf dem Weg zu einer liberalen Eugenik?* Suhrkamp, 2000.
- Hartmann, Nicolai, *Ethik*. Walter de Gruyter, 1926.
- Hartmann, Nicolai, *Ästhetik*. Walter de Gruyter, 1953.
- Hartmann, Nicolai, *Das Problem des geistigen Seins. Untersuchungen zur Grundlegung der Geschichtsphilosophie und der Geisteswissenschaften*, 3. unveränderte Auflage, Walter de Gruyter, Berlin, 1962.
- Kaku, Michio, *The Future of Mind? The Scientific Quest to Understand, Enhance and Empower the Mind*, Anchor Edition, 2014.
- Kelly, Kevin, *The Inevitable. The 12 Technological Forces that Shape Our Future*. Kindle Edition, e-book. 2016.
- Kurzweil, Ray: *The Singularity Is Near: When Humans Transcend Biology*, Penguin Books, 2005.
- Kurzweil, Ray, *How To Create a Mind. The Secret of Human Thought Revealed*, Viking Penguin Edition, London, 2012.
- Linssen, Charl, Pieter Lemmens, “Embodiment in Whole-Brain Emulation and its Implications for Death Anxiety”, *Journal of Evolution and Technology*, Vol. 25. (2016) No. 2. pp. 1-13.
- Pigliucci, Massimo, “Mind Uploading: A Philosophical Counter-Analysis”, in Russel Blackford, Damien Broderick (eds.), *Intelligence Unbound: The Future of Uploaded and Machine Minds*, Wiley Blackwell. Malden – Oxford, 2014. pp. 119-130.
- Pokol Béla, *Theoretische Soziologie und Rechtstheorie*. Schenk Verlag, 2013.
- Peter Sloterdijk, *Regen für Menschenpark. Ein Antwortschreiben zu Heideggers Brief über den Humanismus*, Suhrkamp. 1999.
- Searle, John R., *Minds, Brains and Science*. Harvard University Press. 1984.
- Searle, John R., *The Rediscovery of the Mind*. The MIT Press. 1992.

3

Die Seinsschichten und die Fragen der Roboterethik

Die menschliche Existenz und das Leben menschlicher Gemeinschaften basieren auf den kumulativen Regelmäßigkeiten der Seinsschichten, die durch die Evolution aufeinander aufgebaut sind, gemäß den Thesen der Ontologie von Nicolai Hartmann (Hartmann, 1962). Die beschleunigte Entwicklung und zunehmende Verwendung künstlicher Intelligenz (KI) in den letzten Jahren in dieser Struktur wirkt sich direkt auf die oberste Schicht der vier (physischen, biologischen, spirituellen und intellektuellen) Seinsschichten aus und erhöht ihre Stärke zum Nachteil der unteren. Und mit der späteren Entwicklung der künstlichen Intelligenz, die sich eventuell von der menschlichen Kontrolle löst und Unabhängigkeit erlangt, kann sie als eine evolutionär geschaffene, neue Seinsschicht wahrgenommen werden. Im Gegensatz zu den drei bisherigen Evolutionssprüngen wäre jedoch für sie nicht alle unteren Seinsschichten erforderlich. Unter Berücksichtigung der Roboter, die heute die physischen Inkarnationen von KI darstellen, braucht die KI nur die physische Seinsschicht. (Pokol, 2018). Vor diesem theoretischen Hintergrund versuchen die Analysen in dieser Studie, die aufkommenden moralischen und damit verbundenen rechtlichen Dilemmata innerhalb der Mechanismen heutiger Gesellschaften zu untersuchen, die zunehmend von künstlicher Intelligenz durchdrungen sind, und gleichzeitig ist es zu berücksichtigen, inwieweit sich der analytische Rahmen ändert, wenn die Mehrschichtigkeit der menschlichen Leben und damit der Gesellschaft ständig im Auge behalten werden.

1. Die Vorfragen der Roboterethik

In seiner Studie über die ethischen Probleme der Roboterwelt identifiziert *Keith Abney* drei Bereiche, um die Probleme zu gruppieren: 1) das Feld der Anforderungen und Verbote für Roboterhersteller und Programmierer (wie z. B. die medizinische Ethik); 2) zweitens das Feld der Anforderungen, die in Roboter programmiert werden sollen, die zuerst von *Assimov* unter der Überschrift „Drei Gesetze der Robotik“ formuliert worden und 3) schließlich perspektivistisch für der Zukunft taucht die Frage nach den moralischen Forderungen und „Menschenrechten“, die Roboter zu dieser Zeit im Besitz von Selbstbewusstsein haben könnten (Abney 2011, 35). Ein gemeinsames Dilemma für alle drei Bereiche ist die Wahl zwischen den wichtigsten Ausgangspunkten der Moraltheorie, die in den verschiedenen moralphilosophischen Schulen der umfassenden moralphilosophischen Gemeinschaften schon ausgearbeitet wurden. Eine solche Schule kann als den *deontologischen* Ausgangspunkt identifiziert werden (die Regel ist die Regel, und diese müssen befolgt werden), wofür Kants Moralphilosophie am bekanntesten ist und die polarisierende Oppositionsschule, die die Berücksichtigung der *Konsequenzen des Handelns als Grundlage für moralische Entscheidungen* betrachtet. Schließlich drittens kann die Schule der *Tugendethik* erwähnt werden, die sich nicht auf die Anforderungen konzentriert, die in jeder Situation bei der Definition der Moral zu beachten sind (wie die beiden früheren Schulen, wenn auch vom

gegensätzlichen Richtung), sondern auf die dauerhaften Dispositionen der menschlichen Persönlichkeit, einfacher gesagt: auf die sozialisierten moralischen Werte. Hier fragt der Mensch nicht, was die moralische Regel in einer Situation ist, denn in der immer komplexer werdenden modernen Welt gibt es oft keine klaren Regeln, sondern wie ein mutiger, gerechter, treuer, wahrer Mann entscheidet (Abney 2011, 37).

Von den drei Schulen ist die Schule der Deontologie nur für Roboter möglich, die im engsten Fachgebiet eingesetzt werden und den genauen Regeln folgen, ohne die Regeln wiegen zu können, da alle Situationen nur in einem so engen Gebiet berechnet und beherrscht werden können, aber auch hier können unvorhergesehene Situationen auftreten und die Roboterentscheidung in falsche Richtung lenken. Zum Beispiel könnte es im Prinzip in den Entscheidungsalgorithmus eines Kampfroboters eingespeist werden, „niemals ein Kind zu töten!“. Bei Kindersoldaten in afrikanischen Kriegen würde dies jedoch eine vorgegebene Liquidation des Kampfroboters bedeuten (Abney, 2011, 42). Bei Allzweckrobotern ist der deontologische Ansatz völlig unanwendbar. Aber auch die konsequentialistische Schule der Moralphilosophie, die auch an die Berücksichtigung individueller Situationen gebunden ist, scheint nur aufgrund ihrer lebensnahen Natur besser zu sein. Hier besteht die Leitprämisse darin, "das Glück möglichst vieler Menschen mit der gewählten Entscheidung zu steigern und nicht zu verringern!" und das ist undurchführbar, weil dies die Verarbeitung einer großen Menge von Informationen erfordern würde, die selbst bei der größten Kapazität der Computerdatenspeicherung meistens nicht zeitgemäß gemacht werden könnte. Keith Abneys Position ist daher, dass es in Bezug auf den zweiten Bereich der Robotermoral (d. H. Moralische Entscheidungsprämissen, die in den Roboteralgorithmus programmiert sind) eine Mischung der Deontologie und der Tugendethik die beste Perspektive gibt, und eine Mischung daraus die beste eingebaute Robotermoralversion schaffen kann. Demnach bilden die abstrakteren moralischen Normen (moralische Tugenden) den Entscheidungsrahmen, und die eingebauten Ziele und Entscheidungskontexte spezifizieren immer die Determinanten der vom Roboter in den gegebenen Situationen gewählten Entscheidung: „The hybrid approach of hypothetical rather than categorical imperatives (within a deliberately restricted, not universal, frame) coming from virtue ethics appear the best bet for near term robotic morals (in sense two). (...) The emphasis on being able to perform excellently in a particular role, and the corresponding specificity of the hypothetical imperatives of virtue ethics to the programming goals, restricted contexts, and learning capabilities of non-Kantian autonomous robots, makes virtue ethics a natural choice as the best approach to robot ethics.“ Abney 2011, 51)

Die Verbindungen zwischen den Seinsschichten und Moral werden indirekt von Abney berührt, wo er sich gegen den Emotivismus stellt, der Moral mit moralischen Emotionen identifiziert, und gegen die kognitive Wahrnehmung von Moral, die sich ihr widersetzt. Es zeigt, dass, wenn Moral aufgrund dem Standpunkt der Emotivisten an Emotionen gebunden ist, auch Primaten mit Emotionen nicht von Moral ausgeschlossen werden können, was absurd ist: „Such views, in addition to being unable to explain why nonhuman animals lack morality, also have struggled to explain the apparent cognitive meaningfulness of ethical claims and especially ethical disagreement. (They also naturally have severe difficulties accounting for the ethics of emotionless robots.“ (Abney, 2001, 46). Im Gegensatz dazu sieht er die Position der Evolutionspsychologie, die den neuen Entscheidungsmechanismus der menschlichen Evolution als Erklärung für die Moral betont und das bedeutet, dass die Menschen immer mehr auch ein System kognitiver Entscheidungsfindung entwickelt haben, das die aktuellen Entscheidungen so umformt, dass auf den instinktiv-emotionalen ersten Denkschritt immer eine zweite kognitive Überlegung folgt und somit der erste korrigiert

wird:²³ „Evolutionary psychologists claim there are not one but two of decision-making systems within most humans. The first is an instinctual, emotionally laden system that serves as the default for much human activity, particularly when stressed or under pressure. Many other animals share this noncognitive decision-making system, in which (quite literally) we „know not what we do” - or quite why we do it. (...) But this „ghost in the machine” does not exhaust human agency; Libet and others found we also have a „veto” ability that can, after its subconscious initiation, still alter our action, in accord with a decision by a second, conscious cognitive system” (Abney 2011, 46) Abney erinnert sich fast an die sich gegenseitig formende Wirkung der beiden überlagerten Schichten und erinnert sich fast an Nicolai Hartmann: „In humans, this deliberative system overlays the ancestral instinctual, emotional (and faster) decision-making system and so reason is quite often trumped by our instinctual drives.” (S. 46.)

Nachdem Abney daraus geschlossen hat, dass die obere (kognitiv-rationale) Schicht des zweischichtigen menschlichen Entscheidungsmechanismus für moralische Entscheidungen verantwortlich ist, stellt er die Frage auf, ob eine moralische Entscheidung im Prinzip auch ohne eine umgestaltete, überschriebene untere Schicht möglich ist? Denn entscheidet die Beantwortung dieser Frage auch davon, ob eine moralische Entscheidung für Roboter ohne emotionale Schicht möglich ist. In dieser Frage entscheidet er dann genau umgekehrt wie früher Hartmann. Es ist sehr gut möglich - sagt er -, dass ein rationaler Entscheidungsmechanismus für die moralische Entscheidungsfindung auch ohne eine emotionale Seinschicht ausreicht: „Hence, deliberative system capable of agency necessary for the existence of morality, and so for moral personhood. But is the ancestral emotional system needed as well? (...) In other words, - could (emotionless) robots be moral person? (...) The key to moral responsibility and personhood is the possession of moral agency, which requires the capacity for rational deliberation – but not capacity for functional emotional states, therefore, robots may well qualify” (Abney, 2011, 47).

Basierend auf Hartmann gibt es bei diesen Analysen zwei Probleme. Einerseits kann angesichts der drei Seinschichten über der physischen Seinschicht als fehlerhaft angesehen werden, dass Abney biologische Reize mit den Determinanten der emotionalen Schicht kombiniert. Bereits hier gibt es ein Bauen aufeinander und eine Transformation, und ein Instinkt der rohen Instinktwelt wird durch die Emotionen der höheren Schicht der spirituellen Existenz ergänzt. Zum Beispiel ist die Wildheit eines biologischen Sexualtriebs von Zusammengehörigkeitsgefühlen geprägt, ganz zu schweigen von den intellektuell-symbolischen Überschreibungen, die immer noch darauf aufbauen, und den Geschlechtsverkehr verändernden Aspekten der sublimierten Liebesbeziehungen, die sie herstellen.²⁴ Das heißt, es ist kein doppelter, sondern ein dreifacher Entscheidungsmechanismus, der bei menschlichen Entscheidungen analytisch getrennt werden muss, und zusätzlich zu den elementarsten instinktiven Reaktionen und Determinanten stehen noch ihre emotional umgeformten Manifestationen unter rationaleren Überlegungen auf intellektueller Ebene. Eine Entscheidung und der Instinkt, der sie direkt bestimmt, beziehungsweise seine emotionale Transformation und ihr intellektuelles Überschreiben sind jedoch in die voneinander abhängigen Gesetze aller drei oberen Seinschichten eingebettet. Somit setzt die menschliche Moral in den Gesellschaften aller menschlichen Zivilisation voraus, um als Rasse zu überleben, dass Männer und Frauen in irgendeiner Form dauerhaft zusammenleben, um Kinder zu haben und erzo-gen zu werden.

²³ Es lohnt sich darauf hinweisen, dass diese Ansicht in der Rechtsphilosophie lange vor dem Aufkommen der wissenschaftlichen Psychologie im Jahr 1820 von Hegel schon vertreten wurde. Siehe Georg Wilhelm Friedrich Hegel, „Grundlinie der Philosophie des Rechts. (Werke Band 7.) Suhrkamp, Frankfurt am Main. 1979. S. 301.).

²⁴ Siehe Luhmanns Arbeit, die diesen Prozess historisch analysiert: Niklas Luhmann, „Liebe als Passion: Zur Codierung von Intimität“, Suhrkamp, Frankfurt am Main 1994

Eine größere Gemeinschaft ist notwendig für den erfolgreichen Kampf und das Überleben des Kampfes mit den Naturkräften und anderen Gruppen von Menschen, und innerhalb dieser größeren Gemeinschaften müssen sie in mehr oder weniger harmonischen Beziehungen interagieren, um gemeinsame Aktivitäten zu organisieren. Moralische Tugenden (Normen und Werte) werden daher auf die Gesetze der spezifischen physischen, biologischen, spirituell-emotionalen und intellektuellen Seinsschichten des Menschen und seiner Gemeinschaften zugeschnitten und durch diese aufrechterhalten, und nur aufgrund der Verengung der Moraltheorien in den letzten Jahrzehnten sind die bewussten moralischen Entscheidungen zum Schwerpunkt der Moralphilosophie geworden. Hegel in den frühen 1800er Jahren oder Rudolf von Jhering in den 1870er Jahren und dann Nicolai Hartmann in den 1920er Jahren sahen immer noch deutlich, dass jeder Mensch in seiner Sozialisation die akkumulierte moralische Normen und Werte, Tugenden vieler Generationen nur übernehmen, von denen die umfassenderen Gemeinschaften aufrechterhalten werden, ohne die die Individuen nicht lebesfähig sein könnten.²⁵

Daraus folgt ein weiteres Problem bei Abneys Analyse, und das ist, dass die moralische Entscheidung anscheinend nur darin besteht, Normen nach einer intellektuell-rationalen Berechnung zu folgen, aber nicht die untere psycho-emotionale Seinschicht erfordert. Außerdem, wie wir gesehen haben, die Gesetze der biologischen Seinsschicht und der Instinkt, der dies jedem Menschen vermittelt, sind wichtig für die moralische Entscheidung. Aber auch in Anbetracht dessen kann gesagt werden, dass moralische Normen, moralische Tugenden nur in menschlichen Gemeinschaften bestehen (und somit in den Menschen der nächsten Generationen sozialisiert werden), weil nur auf diese Weise eine dauerhafte und harmonische menschliche Existenz in menschlichen Gemeinschaften möglich ist, die von allen vier menschlichen Seinsschichten definiert wird. Wenn also ein künstlich intelligentes Wesen allein mit der geistlichen Seinsschicht existieren kann, und höchstens nur einen physisch-mechanischen Körper braucht, um Selbstbewusstsein zu haben und bewusste Aktivität auszuüben, bzw. sich dauerhaft in der Zeit reproduzieren zu können, dann haben die moralischen Normen der menschlichen Existenz, die auf der biologisch-spirituellen Schicht des Lebens beruhen, keine Funktion. Die moralischen Normen würden für solches Wesen nur äußere Dinge bedeuten. Wenn also ein solches Roboterwesen sein Programm und sogar seine Hardware mit den Deep Learning-Algorithmen - wie er es heute größtenteils tut - ständig neu aufbauen kann, dann ist die Erosion moralischer Normen, die für ihn extern und funktionslos sind, fast unvermeidlich. Das heißt, obwohl es möglich ist, Programme, die Emotionen imitieren, in Roboter zu programmieren, und sie können dennoch die Entscheidungsaspekte (Verbote, Entscheidungsprioritäten) erklären, die von moralischen Normen in heutigen Robotern verlangt werden, die im Grunde immer noch unter menschlicher Kontrolle stehen, aber wenn sie ein gewisses Maß an Selbstlernfähigkeit erreichen, kann es ungewiss sein, ob die Folgung dieser Normen bleiben wird. In ferner Zukunft (jedoch im Falle eines exponentiellen Fortschritts in sogar zwanzig bis dreißig Jahren) wäre es falsch, im Falle der Roboter in der Roboterwelt, befreit von menschlicher Kontrolle und selbstbewusst, das Überleben der Normen der menschlichen Welt in der Roboterwelt anzunehmen.

2. Operative Moral, funktionale Moral und vollkommene Moral

²⁵ Siehe im Detail Béla Pokol, „Theoretische Soziologie und Rechtstheorie. Kritik und Korrigierung der Theorie von Niklas Luhmann“. Passau. Schenk Verlag. 2013. S. 185-208.

Um die moralischen Dilemmata und Probleme der Roboterwelt besser analysieren zu können, erscheint die von *Colin Allen* und *Wendell Wallach* in ihrer gemeinsamen Studie verwendete Dreifachteilung nützlich. Basierend auf unterschiedlichen Graden der Entscheidungsautonomie bezeichnen sie den Grad der operativen Moral für Roboter, die nur die Aktionen ausführen können, die von den Programmierern, die ihren Algorithmus erstellten und möglicherweise von ihren spezifischen Benutzern festgelegt, und vollständig in sie eingespeist wurden. Auf der anderen Seite sind diejenigen, die den Grad der funktionalen Moral erreicht haben, und das bedeutet, dass sie die spezifische Aktion in jeder Situation auf der Grundlage der von ihren Sensoren bereitgestellten Informationen unter den in ihren Algorithmus eingespeisten Aktionsrahmen auswählen. Schließlich wird der autonomeste Grad an Moral bei Robotern gesehen, die mit dem Aufhören des menschlichen Einflusses die Ebene der vollen moralischen Persönlichkeit erreichen, obwohl dieser Typ jetzt und in naher Zukunft nicht als wahrscheinlich angesehen werden kann, aber später kann von ihrer Schaffung ausgegangen werden: „System with very limited autonomy and sensitivity have only „operational morality”, meaning that their moral significance is entirely in the hands of designer and users. As machines become more sophisticated, a kind of „functional morality” is possible, where the machines themselves have the capacity for assigning and responding to moral challenges. The creators of functional morality in machines face many constraints due to the limits of present technology. This framework can be compared to the categories of artificial ethical agents described by James Moor (2006, 18) which range from agents whose actions have ethical impact (implicit ethical agents) to agents that are explicit ethical reasoners (explicit ethical agents.) As does Moor, we emphasize the near-term development of explicit or functional moral agents. However, we do recognize that, at least in theory, artificial agents might eventually attain genuine moral agency with responsibilities and rights, comparable to those of humans.” (Allen, Walach 2011, 57-58.)

Ohne auf die mögliche Kritik einzugehen, ob es sich lohnt, den Grad der Moral für Roboter zu verwenden, die bereits von Programmierern unter dem Namen Betriebsmoral vollständig definiert wurden, sind die Roboter der funktionalen Moral in der heutigen Entwicklungsphase in der Roboterwelt wirklich interessant. Selbstfahrende Autos, selbstfahrende Kampfroboter und in geringerem Maße Roboterkrankenschwestern in Altenpflege- und Gesundheitseinrichtungen, die diese Autonomie bereits erreicht haben, lassen solche Roboter und selbstfahrende Autos langsam zwischen uns rollen oder uns transportieren (heute meistens nur in Japan und den Vereinigten Staaten) und die von ihnen aufgeworfenen moralischen Entscheidungsprobleme geben ihrer Analyse praktische Bedeutung. Die Autoren gehen die Möglichkeiten der Wahl zwischen den bereits oben gesehenen Trends der Moraltheorie durch und sie sehen die tugendethischen Richtung geeignet, um die funktionale Moral von Robotern zu schaffen. Ihrer Analyse zufolge können die auf diese Weise gespeisten moralischen Werte (Tugenden) die Entscheidungsrahmen geben, der durch Training durch neuronale Lernmechanismen geklärt wird, und auf diese Weise werden die allgemeineren Gesichtspunkte der Tugenden zu praktischen moralischen Entscheidungsfaktoren: „The virtue-based conception of morality can be traced to Aristotle. Virtues constitute a hybrid between top-down and bottom-up approaches, in that the virtues themselves can be explicitly described (at least to some reasonable approximation), but their acquisition as moral character traits seems essentially to be a bottom-up process. Placing this approach in a computational framework, neural network models provided by connectionism seem especially well suited for training (ro)bots to distinguish right from wrong. (Allen, Wallach 2011, 59-60). Strukturell entspricht dies im Großen und Ganzen den Entscheidungen, die man in seinem täglichen Leben gewohnt ist, die auf abstrakten moralischen Überlegungen beruhen und an bestimmte Situationen angepasst sind, und welche weniger bewusst als mit bloßem moralischen Sinn getrieben sind. Aber mit dem wichtigen Unterschied, dass anstelle

des Mangels an Bewusstsein und Selbstbewusstsein der heutigen fortschrittlichen Roboter die von Programmierern subtil abgestimmten hybriden Determinanten (Rahmentugenden plus ihr trainingskonkretisiertes Gedächtnis ohne Bewusstsein) die mehr oder weniger akzeptierten moralischen Normen der heutigen menschlichen Gesellschaften für die angemessenen oder ungefähren Entscheidungen geben. Was die dritte Version betrifft, ob es wirklich möglich ist, Roboter theoretisch nach dem Grad der vollen moralischen Persönlichkeit nach menschlicher Moral anzunehmen, kann im Fall einer Roboterwelt, die theoretisch außerhalb der menschlichen Kontrolle liegt und vollständige Autonomie erreicht hat, nach der obigen Erklärung nur skeptisch beurteilt werden.

3. Die Abwertung der physikalisch-biologischen Umwelt als moralisches Problem?

Die vierschichtige menschliche Existenz und das wachsende Gewicht der oberen, intellektuellen Seinsschicht und der Abwertung der unteren Schichten haben die menschliche Evolution bisher geprägt, aber die zunehmende Übernahme der verschiedenen Arten der Arbeiten und der Umweltwahrnehmungen durch Roboter wird in Zukunft zu einer großen Veränderung der menschlichen Sozialisation führen, von der die Richtung, auf die Realität zu achten und die Details der realen Welt für ihn in Erfahrung zu verwandeln, weitgehend verändert wird. In einer Studie analysiert *David Zoller* die zunehmend verbreitete Übernahme der Arbeiten von Menschen durch Roboter dahingehend, wie dieser Prozess die Wahrnehmung der alltäglichen Realität im menschlichen Bewusstsein verschlechtert und wie die heute noch vorhandenen Fähigkeiten und Beobachtungsmöglichkeiten verschwinden. Dass dies bereits von jedem selbst beobachtet werden kann, reicht aus, um die bereits in Mobiltelefonen gespeicherten und damit weitgehend aus dem Bewusstsein gelöschten Telefonnummern oder die räumlichen Orientierungsinformationen, die durch GPS aus unserem Bewusstsein verschwinden, und das teilweise Absterben dieser Fähigkeit abzurufen. (Eine kürzlich durchgeführte Gehirnforschung ergab auch, dass in Falle Taxifahrer in London behauptet werden konnte, dass der winzige Teil des Gehirns, in dem eine Gruppe von Gehirneuronen, die auf diesen Zweck spezialisiert waren, um die riesige Menge der Informationen über die Straßen von London zu speichern, mit der Ausbreitung von GPS verschwand und dieser Gehirnsektor stattdessen auf eine andere Funktion überging.)

Zoller bringt dieses Problem moralischen Fragen näher, indem er die moralische Entscheidungsfindung auf die Wahrnehmung der gesamten Realität und in diesem Fall auf die Bildung der menschlichen Identität und auf das detaillierte Wahrnehmungswissen, das ein Erwachsener von Kindheit an erworben hat, basiert. Wenn zukünftige Generationen sich seit ihrer Kindheit so sozialisiert, dass sie von Robotern umgeben werden und ihre unmittelbare Umgebung von Robotern wahrgenommen werden, ohne sie stattdessen Wahrnehmungsaktivitäten und -aufgaben auszuführen, werden sie nicht nur entlastet, sondern verfügen auch nicht über detaillierte Kenntnisse der heutigen Erwachsenen. Als verantwortungsbewusstes Wesen können sie auch nicht zu moralischen Entscheidungen heranwachsen, mit anderen Worten, sie werden kindisch: „My own argument is premised on the way that skill opens up corners of reality, so to speak, that are inaccessible to the unskilled. (...) The maturity or adulthood we earn by adjusting ourselves to the „real world“, of course, has a certain moral and personal appeal: a world of lazy psychological infants is, we might think a worse world on a variety of spectra.” (Zoller 2017, 81, 86). Die Tatsache, dass diese Bereiche der Realität über unsere Wahrnehmung hinausgehen und es kommt stattdessen die mechanische Informationsverarbeitung dieser Roboter, ermöglicht uns sich auf

eine Weise anzupassen, die durch diese Veränderung jetzt unbewusst ist, und dies erschüttert auch unsere moralische Identität, sagt Zoller: „Given that automating a skilled activity means agreeing that we will exit some niche of perceptual reality, and maybe exit it forever, (...) The more suddenly, broadly, and pervasively we hand our perceptual facility over to the robots, the more likely we will make mistakes and simply „lose data“ that were surprisingly integral to our moral and social lives.” (Zoller, 2017, 86).

Zwar muss anerkannt werden, dass Zoller im Gegensatz zu den bisherigen Analysen, die die Übernahme von Arbeitsplätzen durch Roboter als einseitige menschliche Erleichterung thematisieren - abgesehen von den bereits diskutierten sozial negativen Folgen der Arbeitslosigkeit (siehe z. B. Ford, 2014) - tiefer ging, indem er sich eingehender mit der Änderung der menschlichen Wahrnehmungsfähigkeit befasste, aber es muss kritisiert werden, dass sie unbewusst zu viel Wert auf die Seinsschichten der physikalisch-biologischen Umgebung legt. Mit Blick auf Hartmanns Seinsschichten kann diese Änderung völlig anders gelesen werden. Die von Zoller skizzierten Änderungen bedeuten nicht den Verlust der Wahrnehmung der ganzen Realität und die Fähigkeit dazu, sondern nur die Fähigkeit, die physikalisch-biologischen Seinsschichten wahrzunehmen und diese an Roboter und Software-Bots weiterzugeben. Auf diese Weise können die befreiten Wahrnehmungsfähigkeiten und Gehirnsektoren des Menschen stärker rekonstruiert werden, um Informationen über seine spirituuell-emotionale Seinschicht bzw. seine intellektuelle Seinsschicht zu verarbeiten. Seine moralischen Entscheidungen werden daher in Zukunft mit weniger physischen und biologischen Umweltinformationen getroffen - diese werden von Robotern in mechanischen Prozessen heruntergefahren – und diese Entscheidungen können stattdessen stärker auf die Informationen von der spirituuell-emotionalen und rational-intellektuellen Seinsschichten basieren. Die abnehmende Bedeutung der beiden unteren Seinsschichten und stattdessen die stärkere Ausdehnung der beiden oberen Sinnsschichten für die menschliche Existenz können natürlich die Grundlagen unserer moralischen Entscheidungen und die Anreize, die dabei eine Rolle spielen, erheblich umbauen. Zum Beispiel kann die Implantation von Dutzenden von Körpersensoren und deren Verbindung mit in den Wolken gesammelten Informationsbasen sowie die automatische Diagnose durch Roboter von Gesundheitssoftware und die automatische Aktivierung bestimmter Dosen von im Körper implantierten Arzneimitteln die Alarmierung durch die Schmerzgene in unseren Zellen weitgehend überflüssig machen. (siehe Kelly 2016: 34-56). Die pränatale Gentechnik ermöglicht es daher, diese zu minimieren, und die Bedingungen eines schmerzlosen menschlichen Lebens können die moralischen Verpflichtungen und Anreize, die heute damit verbunden sind, neu definieren. Insgesamt teilen wir daher Zollers Bedenken hinsichtlich der moralischen Infantilisierung nicht.

4. Moralische Dilemmata und Verantwortlichkeiten in hybriden und vernetzten Systemen

In einer Studie untersuchte das Autorenpaar *Wulf Loh - Janina Loh* die Fragen der moralischen und rechtlichen Verantwortung, die bei derzeit entwickelten selbstfahrenden Autos auftreten (Loh, Loh 2017, 35-48). Sie gehen davon aus, dass sich die selbstfahrenden Autos von heute nur in der Phase der operativen Moral befinden, sodass sie gegenüber ihren Herstellern und Programmierern nicht einmal eine funktionale moralische Autonomie erreichen. Das Autorenpaar hat diese Position auf der Grundlage einer von *Stephen Darwall* entwickelten moralischen Entscheidungsstruktur eingenommen, die in vier Aspekte unterteilt ist und zielt darauf ab, die für die moralische Entscheidungsfindung erforderlichen Aspekte

der Autonomie zu trennen. Der Aspekt der Autonomie, der für die Ebene der gesamten moralischen Persönlichkeit erforderlich ist, wird als *persönliche* Autonomie bezeichnet, dh die Fähigkeit, persönliche Werte, Ziele und endgültige Lebensziele zu besitzen und zwischen ihnen wählen zu können. *Moralische* Autonomie ist der andere Aspekt, und dies bedeutet, dass ihre Werte und Ziele moralische Prinzipien und ethische Überzeugungen enthalten, und zusammen mit diesen berücksichtigen sie bei ihren Entscheidungen immer Alternativen. Diese beiden existieren in heutigen Robotern nicht, und nur der Mensch ist zu einer solchen Autonomie fähig, aber der Aspekt der *rationalen Autonomie* steht Robotern auf der Ebene der funktionalen Moral bereits zur Verfügung. Dies bedeutet, dass der Roboter bei einer Entscheidung zwischen Gründen unterschiedlichen Gewichts wiegen kann. Ihr Algorithmus kann dies bereits ermöglichen, indem er reinen abstrakten Entscheidungsrahmen einbezieht - und damit einige Freiheiten lässt -, in denen die Gewichtung zwischen möglichen Entscheidungsrichtungen im Lichte spezifischer Daten erfolgt, die ständig von ihren Sensoren aufgezeichnet werden, und sie auf Grund dieser Daten entscheiden. Schließlich ist der vierte Aspekt der Autonomie *die Entscheidungsautonomie*, und dies bedeutet die Fähigkeit des Roboters, Entscheidungen nicht nur durch externe Daten zu treffen - die eingebauten Rahmendeterminanten kontinuierlich zu konkretisieren - sondern auch seine internen Entscheidungsprioritäten ohne diese zu ändern. Anhand der Beispiele der Autoren - zwei bereits gebräuchliche Robotertypen (*Kismer* und *Cog*) - scheint es möglich zu sein, diesen Grad an Autonomie auf der Grundlage ihrer in den Algorithmus des Roboters integrierten und nicht mehr von außen gesteuerten Selbstlernmechanismen zu erreichen: „Cog the first robot that can interact with its surroundings due to its embodiment, might pass as an example of a weak functional responsible agent, since its ability to communicate as well as judgments has been improved over that of Kismer. Even more importantly, Cog’s overall autonomy has evolved, since it includes an „unsupervised learning algorithm“”. (Loh, Loh 2017, 40). Da der derzeitige Algorithmus für selbstfahrende Autos noch keinen solchen außer Kontrolle geratenen Selbstlernmechanismus enthält, befinden sie sich nur auf der Ebene der operativen Moral, und diese moralische und rechtliche Verantwortung liegt vollständig bei ihren Entwicklern (Designern, Herstellern und Programmierern) und Autohändlern bzw. Eigentümern und jeweils zwischen den Insassen des Autos.

Aber selbst mit diesem technischen Niveau übertreffen selbstfahrende Autos den Menschen bereits und lassen sie - und insbesondere ihre Programmierer - mit moralischen Dilemmata zurück, die im Falle der Menschen in außergewöhnlichen und unerwarteten Fahrsituationen nicht zu sehen sind. Wenn zum Beispiel innerhalb des Bremswegs direkt vor dem Auto eine Gruppe von Kindern auf die Straße springt, um einen gerollten Ball zu holen, kann der Fahrer mit dieser Durchschnittsgeschwindigkeit nicht mehr anhalten oder sogar bremsen, sodass er im schrecklichen Fall keine moralische und rechtliche Verantwortung trägt. Aber die selbstfahrende Automatisierung, die viel schneller reagieren kann, muss möglicherweise noch eine Entscheidung treffen, wenn sie nicht anhalten kann, sondern gegen eine Säule stößt - und möglicherweise die Insassen des Autos ernsthaft verletzt - oder Kinder fährt und tötet, um dies zu vermeiden. Aber technische Fähigkeiten, die weit über den Menschen hinausgehen, könnten in Zukunft ein Dutzend ähnlich neuer Aspekte der moralischen Entscheidungsfindung für selbstfahrende Autos schaffen. Das Loh-Autorenpaar schlagen daher vor, dass in Kürze ein separater Personalausweis für Besitzer selbstfahrender Autos erstellt wird, in dem die endgültige Einstellung des Autosoftwareprogramms, die von den Herstellern offen gelassenen Dilemmata, beim Kauf entschieden werden muss, damit die moralische und rechtliche Verantwortung für Folgendes zu übernehmen: „Since these dilemma situations do not allow for on-the-fly-decisions, the driver will have take them beforehand. This means that the driver will have to fill out a moral profile of some sort, maybe in the form of a questionnaire, maybe in the sense of a setup program much as with

today's electronic devices. For convenience, it seems plausible that these moral settings can be saved to a sort of electronic identification device, like an electronic key or the driver's smartphone, assuming that issues of data security can be solved" (Loh, Loh 2017, 46).

Die Entwicklung vernetzter Roboter und das allmähliche Werden von „klugen Objekten“ (Smartphones, Smart-TVs usw.) um uns herum haben erst vor kurzem begonnen, und mit der Erweiterung wird das Internet der Dinge (IoT) in Zukunft immer mehr in unser Leben einbezogen. Mensch-Roboter-Hybridsysteme erweitern sich somit um zusätzliche Aspekte, und dies schafft ein weiteres Bündel moralischer und rechtlicher Dilemmata. *Adam Henschke* analysiert diese in seiner neuen Studie (Henschke, 2017, 229-243). Kluge Dinge sind durch multifunktionale Smartphones, kluge Fernseher, Roboterstaubsauger und halb selbstfahrende automatische Autos mit einer Vielzahl von Sensoren weit verbreitet, aber selbst im Alltag in weiten Teilen der Welt wurden diese anderen intelligenten Dinge entwickelt, die bereits jenseits der Forschungslaborphasen liegen und bereits die Haushalte von High-Tech-Anwendern mit kleiner Produktion erreicht haben. Diese jedoch, wie wir bereits mit Smartphones usw. erlebt haben, wird sich in ein paar Jahren ausbreiten und ihr Massengebrauch wirft neue moralische und rechtliche Dilemmata auf. Ein Beispiel ist der Smart-Kühlschrank, der Lebensmittel mit RFID (Radio Frequency Identification) und damit digital identifizierter Menge, Haltbarkeit usw. enthält und der Smart-Kühlschrank liest diese Daten ständig, erkennt die Erschöpfung jeder Lebensmittelmenge und da er mit den webbasierten Verkaufsmechanismen der Supermärkte in der Nähe im Internet verbunden ist, kann er Lebensmittel und andere Haushaltsgegenstände bestellen, die automatisch geliefert werden. In Japans alternder Gesellschaft kann eine zunehmend größere Menge von älteren Personen durch den Einsatz von Pflegerobotern versorgt werden, und in vollständig digitalisierten Smart Homes kann ein solcher Roboter auch hilflose ältere Menschen versorgen, der auf diese Weise bestellte Lebensmittellieferungen übernehmen kann. Indem er den hilflosen älteren Menschen, der ihm anvertraut ist, beobachtet und mit ihm kommuniziert, kann er den Hausarzt oder gegebenenfalls das Krankenhaus telephonisch anrufen, wenn seine eingebauten Algorithmen ein ernsthafteres Gesundheitsproblem wahrscheinlich machen.

Dieses Beispiel zeigt, wie in ein oder zwei Jahrzehnten Roboter in immer mehr Teilen der Welt benötigt werden, um die Arbeiten immer mehr durch allmächtige Roboter und Smartphones zu lösen, die in umfassenden Informationssystemen zur Erfüllung ihrer Funktionen eingesetzt werden können. Diese wachsende Unentbehrlichkeit des Internet der Dinge schafft aber auch neue Gefahren und moralische Dilemmata im Vergleich zu einfachen Robotern. *Adam Henschke* weist in seiner Schrift darauf hin, dass die Neuheit des Internet der Dinge im Vergleich zu einzelnen Robotern darin besteht, dass letztere hauptsächlich das Problem der physischen Sicherheit aufwirft und die Risiken in dieser Dimension bewertet werden müssen. (Z. B. hat ein Roboterstaubsauger kürzlich einem unerwarteten Insassen schwere Verletzungen zugefügt, aber auch ein oder zwei tödliche Unfälle selbstfahrender Tesla-Autos können dafür angeführt werden.) Im Gegensatz dazu treten Sicherheitsprobleme und Gefahren im Internet der Dinge in zwei verschiedenen Dimensionen auf. Hier spielen neben der physischen Sicherheit auch Fragen der Informationssicherheit eine Rolle, da der genannte Altenpflegeroboter, der mit der Software von Krankenhäusern, Ärzten und anderen Stellen im Internet verbunden ist, Hackern oder anderen Personen Informationen über Daten liefern kann, die von seiner eingebauten Kamera und anderen Sensoren aufgezeichnet wurden. Sie können die kontinuierlich gesammelten Gesundheitsdaten über eine ältere Pflegekraft nicht nur an die Software des zuständigen Krankenhauses weitergeben, sondern auch an diejenigen, die böswillige Absichten und Pläne machen. Auf die gleiche Weise können unsere Smart-TVs mit einer Reihe von Anwendungen nicht nur ihren Komfort erfüllen, sondern mit ihren eingebauten Kameras und Mikrofonen die gesamte Lebensdauer der Wohnung an Software und Informationsdatenbanken übertragen, die wir nicht sehen.

Diese Sicherheitsanfälligkeit kann auch zu einer physischen Sicherheitsanfälligkeit führen, wenn beispielsweise ein gehacktes automatisches Türschloss für einen Einbrecher aus der Ferne durch externe Anweisungen von Smart-Geräten geöffnet wird. Oder, wie es bereits in einem eleganten Strandhotel geschehen ist, wurden die elektronischen Smart-Schlösser von einer kriminellen Gruppe von außerhalb der Apartments blockiert, und die Hotelgäste der reichen Elite waren Gefangene, bis das erforderliche Lösegeld gezahlt wurde. Henschke erwähnt aber auch die Möglichkeit, dass das Elektroschloss von Auto eines Milliardärs von Kriminellen blockiert wird, nachdem er ausgestiegen ist und seine gefangenen Kinder darin am sonnigen Tag nur freigelassen werden, wenn er Hunderttausende überträgt (Henschke 2017, 234). Unmittelbar nach dem Vorfall ersetzte das besagte elegante Hotel die elektrischen Schlösser, die von außen gekehrt werden konnten, und installierte die guten alten, traditionellen Schlösser wieder. Nach einem solchen Vorfall wird der genannte Milliardär wahrscheinlich auch die Internetfunktionen seines Autos für eine Weile einschränken. All dies erzwingt jedoch Entscheidungen in moralischen und rechtlichen Dilemmata und Wahlen, über die es sich im Großen und Ganzen lohnt, nachzudenken. In der Welt unserer Objekte, die im Internet der Dinge immer weiter verbreitet wird, werden die alten einfachen Dinge bereits fallen gelassen, und wir werden die Objekte, die wir in die Cloud-Datenbank winken, nicht nach Belieben ersetzen können. So wie wir heute trotz aller negativen Aspekte, die Verwundbarkeit mit sich bringen, das Internet nicht aufgeben würden.

Ein solches Dilemma des Internet der Dinge, der in vernetzten und umfassenden Cloud-Datenbanken eingebunden sind, besteht darin, welcher der widersprüchlichen Anforderungen der beiden Sicherheitsarten - physische Sicherheit und Informationssicherheit - Vorrang eingeräumt werden sollte. Zum Beispiel kann es bis zu einem gewissen Grad wichtig sein, die kluge Wohnung einer älteren Person, die kaum in der Lage ist, sich zu bewegen, durch Kameras und Mikrofone vollständig fernüberwachbar durch medizinische Zentren zu machen, aber es kann auch bedeuten, die intimsten Manifestationen des Lebens aufzudecken, die über das Notwendige hinausgehen. Wenn der Schwerpunkt auf Informationsautonomie und eingeschränkter Beobachtung und Transparenz liegt, werden die Informationen, die in seltenen Fällen noch benötigt werden, möglicherweise nicht an das Pflegezentrum weitergeleitet, und der ältere Pflegebedürftige kann sterben. Henschke weist darauf hin, dass es häufig typische Prioritäten gibt, und zum Beispiel bei einem Smart-TV hat die Informationssicherheit eine höhere Priorität, und zu diesem Zweck können wir hier leicht auf die Einschränkungen eingehen. Bei Tausenden von Anwendungen von selbstfahrenden Auto, das an die Cloud-Software gebunden ist, achten wir jedoch stärker auf die Anforderungen der physischen Sicherheit und nur sekundär auf die Anforderungen der Informationssicherheit (Henschke, 2017, 239).

5. Selbstlernen, maschinelles Lernen und Verantwortung

Wie oben erwähnt, wird das Hauptproblem der Zukunft das Dilemma selbstfahrender Autos sein, die von Menschen losgelöst sind und in bestimmten, unerwarteten Situationen nicht mehr von außen blockiert werden können, wenn der auf neuronales Selbstlernen aufgebauten Algorithmus dieses selbstfahrende Autos schon autonom entschieden hat. Da dies die Hauptrichtung der Entwicklung der künstlichen Intelligenz in den letzten Jahren ist, ist es fast sicher, dass dies auch in diesem Bereich nicht umgangen wird. Daher lohnt es sich heute, die Dilemmata der moralischen und rechtlichen Verantwortung von Robotern mit einem hohen Maß an neuronalem Selbstlernen und ihren Machern, Besitzern und Benutzern genauer zu

betrachten. Diese Frage wird in ihrer gemeinsamen Studie von *Trevor N. White* und *Seth D. Baum* (White, Baum, 2017, 66-79) sowie von *Shannon Vallor* und *George A. Bekey* (Valor, Bekey, 2017, 338-353) aus verschiedenen Blickwinkeln analysiert.

Trevor und Baums Studie berücksichtigt nicht nur Designer, Bauherren und Benutzer, sondern berücksichtigt auch die „Bestrafung“ des Roboters selbst im Falle der fortschrittlichen Roboter, in deren Programm bereits ein Straf- und Belohnungssystem integriert ist, und wiederholte Strafen und Belohnungen verstärken in ihrer Programmierung die Entscheidungsrichtungen (positiv oder negativ) in Bezug auf die Auswahl zukünftiger Roboterreaktionen. Damit wird auch Bestrafung / Belohnung in den Lernalgorithmus integriert. Wenn sich die Situation in der Zukunft ergibt, werden die Entscheidungen des Roboters in die richtige Richtung gefördert, und der Roboter muss dafür kein Bewusstsein und Selbstbewusstsein haben. Diese Art, das Selbstlernen durch Wiederholungen zu verstärken, ist laut den Autoren auch akzeptabel: „Non-conscious robots could conceivably be punished with some sort of reduced reward or utility as per whatever reward/utility function they might have. Specifically, they could be reprogrammed, deactivated, or destroyed or put in what is known as a „Box“: digital solitary confinement restricting an AI’s ability to communicate or function. (...) To make this possible, however, such robots ought to be based (at least in part) on reinforcement learning or similar computing paradigms (except ones based on neural network algorithms.)” (Trevor, Baum, 2017, 71).

Das neuronale Lernsystem wird von den Autoren jedoch so beurteilen, dass Designer und Programmierer bereits die Kontrolle über die Reaktion des Roboters auf eine bestimmte Situation verlieren und deshalb sollten von Anfang an als potenzielle Gefahrenquelle, möglicherweise als Katastrophe verboten werden: „Designer could be similarly liable for building robots using opaque algorithms, such as neural networks and related deep-learning methods, in which it is difficult to predict whether the robot will cause harm.” (ebd.) Bei Algorithmen, die solche undurchsichtigen Roboterverhalten zulassen, ist es nicht mehr genug, die nachträgliche Haftung vorzuschreiben, sondern das vorbestellte Verbot ist das zweckmäßig: „Hence, instead of liability, a precautionary approach could be used. This would set a default policy of disallowing any activity with any remote chance of causing catastrophe. (...) In effect, people would be held liable not for causing catastrophe but taking actions that could cause catastrophe.” (Trevor, Baum, 2017, 74). Wenn man sich im Prinzip mit den Autoren über den gefährlicher Charakter neuronaler Deep-Learning-Softwaremechanismen einig ist, muss nur wiederholt werden, dass dies ein Verbot des Hauptwegs zur Entwicklung künstlicher Intelligenz darstellt und deshalb sollte dies im Licht der Mächte hinter der Industrie, Militär usw. als unwahrscheinlich angesehen werden. Daher erscheint es ratsam, nach anderen Pfaden zu suchen, die versuchen, andere Lösung zu finden, ohne das neuronale Tiefenlernen zu verbieten.

Es sollte natürlich betont werden, dass das Lernen neuronaler Netze, von dem das Funktionieren des Zentralnervensystems nachgeahmt wird, durch Einbeziehung externer menschlichen Kontrolle gesteuert werden kann, bevor ein realitätsbildender Effekt ausgelöst werden kann. Dies bleibt jedoch aus einer Reihe von Gründen zunehmend zurück, und dies wird von *Vallor* und *Bekey* in der bereits zitierten Studie analysiert. Ein Grund dafür ist, dass der Vorteil der Verwendung künstlicher Intelligenz anstelle des Menschen, die unglaublich schnell Reaktionsfähigkeit verloren gehen würde, wenn eine retrospektive menschliche Kontrolle eingeführt würde. Darüber hinaus sind in neunundneunzig Prozent der Fälle die Reaktionen korrekt, die um ein Vielfaches höher sind als die Leistung des Menschen. Darüber hinaus kann die Qualität einer viel langsameren menschlichen Kontrolle fraglich sein, da die Entscheidung des Roboters richtiger als die übergeordnete menschliche Entscheidung sein kann. Letzteres geschah mit dem medikamentösen Diagnosealgorithmus von IBM Watson, und die ungewöhnliche Heilung, die durch künstliche Intelligenz aus Millionen von

Onkologiestudien und -diagnosen hervorgehoben und später von ihnen synthetisiert wurde, erwies sich korrekter als die Onkologieentscheidung, die sie außer Kraft setzt: „Watson’s diagnoses and treatment plans are still vetted by licensed oncologists. Still, how reliably can a human expert distinguish between a novel, unexpected treatment recommendation by Watson that might save a patient’s life – something that has reportedly already happened in Japan – and the oncological equivalent of „Toronto?“ (Vallor, Bekey 2017, 343).²⁶ Das Dilemma, an Geschwindigkeit zu verlieren und damit den Vorteil des Roboters zu beseitigen, wird auch durch die Robotersoldaten und die Entscheidungssoftware veranschaulicht, die in Kriegssituationen eingesetzt werden. Hier stellt sich ständig die Frage, ob der Robotersoldat, der das gefährlichste Gebiet und Gebäude betritt, mit den in seinem Besitz befindlichen zerstörerischen Waffen ohne eine externe menschliche Entscheidung die darin befindlichen zerstören kann oder stattdessen die Anordnung der Zerstörung nur durch entfernte menschliche Bestätigung erteilt werden darf. Auf die gleiche Weise taucht das Dilemma auf, ob ein Aufklärungsflugzeug sofort vom Roboterflugzeug zerstört werden könnte, oder dies nur mit menschlichem Eingreifen aus dem entfernten Kommandoraum auf der Grundlage der von ihm übertragenen Informationen durchgeführt werden könnte? Der Zwang zur Geschwindigkeit erfordert, dass der Roboter selbst eine Entscheidung trifft und ausführt, da der Zeitverlust von einer externen menschlichen Entscheidung zur Vernichtung der Roboter führen kann, wenn er an einen gefährlichen Ort einbricht. Aber das Abschießen freundlicher Kampfmaschinen, die bereits mehrmals passiert sind, oder das Töten von Kindern und Frauen, die fälschlicherweise als Feinde in dem eingedrungenen Gebiet identifiziert wurden, spricht dagegen (Vallor, Bekey, 2017, 349).

Detaillierte Lernalgorithmen für neuronale Netze, der aktuellste Trend in der künstlichen Intelligenz, bieten selbstlernender Software bereits eine vielfältige Tiefe für die einfache komputerische Startposition, indem kontinuierlich immer Milliarden von Daten eingespeist werden und die Startposition sich auf diese Weise allmählich zum fortschrittlichsten machen. Bei dieser Technik werden zwischen den Eingaben der selbstlernenden Software und den aufgabenspezialisierten Ausgaben Tausende von neuronalen Zwischenschichten zwischen den massiven Daten gefunden, wobei die Muster und Regelmäßigkeiten unabhängig voneinander gefunden und zur Verwendung hervorgehoben werden. Durch die Kombination von Milliarden von Daten können sie die kleinsten Regelmäßigkeiten hervorheben und verwenden, die für Menschen bei ihren Entscheidungen nicht wahrnehmbar sind: „Between the input and output node layers are „hidden“ layers of nodes that function to process the input data, for example, by extracting features that are especially relevant to the desired outputs. Connections between the nodes have numerical „weights“ that can be modified with the help of a learning algorithm; the algorithm allows the network to be „trained“ with each new input pattern until the network is optimized. (...) The interest in neural network has grown in recent years with the addition of more hidden layers giving *depth* such network, as well as feedback or recurrent layers. The adjustment of the connections strengths in these more complex networks belongs to a loosely defined group of techniques known as deep learning.“ (Vallor, Bekey, 2017 341). Die Auswirkungen von Entscheidungsmustern, die durch diese detaillierten Lernalgorithmen hervorgehoben werden, sind zwar in der Praxis oft überraschend gute Ergebnis erreichen, können jedoch von Designern und Programmierern nicht verstanden werden, und ihre Entscheidungen können durchweg Überraschungen

²⁶ „Toronto Mistake“ war einer von Watsons Grundfehlern in einem landesweiten TV-Quiz, als er alle Teilnehmer mit seinen Antworten auf die schwierigsten Fragen besiegte. Als letzter Fehler machte er Toronto zu den US-Städten und nicht einmal zum schwächsten Konkurrenten wäre verloren gegangen. So ist es zu einem Symbol für Fehlentscheidungen geworden, die von künstlicher Intelligenz getroffen werden, was selten ist, aber in vielen Fällen Tragödien verursacht.

verursachen, einschließlich unterschiedlich starker unangenehmer Überraschungen. Wer sollte die rechtliche und moralische Verantwortung dafür tragen?

6. Identität in der Welt der künstlichen Intelligenz

James DiGiovanna wirft zum Nachdenken anregende Fragen auf, wenn er in seiner Studie fragt, wie sich die Identität von Menschen mit Gehirnimplantaten ändern kann. Er diskutiert dieses Thema gemeinsam mit der Frage der Identität von Robotern, die in Zukunft als vollständig künstliche Wesen auftreten können und bereits Selbstbewusstsein haben (DiGiovanna, 2017, 307-321). Betrachten wir die beiden Problembereiche getrennt.

Die Möglichkeit eines Gedächtnisses, das durch Gehirnimplantaten ergänzt wird, wurde in den letzten Jahren in Mäuseexperimenten entwickelt und hat sich als wirksam erwiesen. Alle hoffen, die Auswirkungen der sich schnell ausbreitenden Alzheimer-Krankheit in alternden Gesellschaften zu mildern und zu heilen (siehe Kaku, 2014, 132-133). DiGiovanna untersucht die Möglichkeit weiterer Entwicklungen in den nächsten Jahren, die sich zusätzlich zu Krankheiten massenhaft ausbreiten könnten, um die Gehirnkapazität zu erhöhen. Und wenn eine technische Lösung für die Probleme gefunden wird, die heute noch auf diesem Gebiet bestehen, ist es fast sicher, dass dies zuerst in der Elite, dann in der Gesellschaft als Ganzes üblich wird, um den größten Wert der menschlichen Intelligenz zu steigern. Dies bedeutet jedoch, dass die dauerhafte Identität jedes Einzelnen, die die Grundlage für Kontakte in Gemeinschaften ist, mehr oder weniger aufgehoben werden kann und es ungewiss werden kann, wie sehr wir von unseren Partnern erwarten können, dass ihre Eigenschaften überleben, die wir bisher kennen und lieben: „The ability to rewrite mental content such as ethical values, the capacity for empathy, and general personality traits undermines personhood. (...) A para-person that could experiment with worldviews, completely adopting and deleting values systems, preferences, and bases for judgement, would be largely lacking in what is commonly understood as the most basic element of personal identity” (DiGiovanna, 2017 311) Dies war die Grundlage unserer Wahl im Falle unserer Freunde, Frau und Freundin, aber auf die gleiche Weise basiert unsere engere menschliche Beziehung mit einiger unserer Arbeitskollegen auf die Liebe ihrer Eigenschaften, während die Beziehung zu anderen nur kalt kollegial ist. Daher basiert unser Leben in der Gesellschaft und in verschiedenen kleinen Gemeinschaften innerhalb unserer Gesellschaft auf unseren dauerhaften Identitäten, und dies kann sich nach dem Hinzufügen von Herz-, Hör- und anderen körperlichen Verbesserungen grundlegend ändern, wenn sich das Gehirn mit Gehirnimplantat verändert wird.

Mit den allmählichen Veränderungen im eigenen Leben wird das eigene Bewusstsein im Detail immer umgebaut, und dies führt zu kleinen allmählichen Veränderungen in der eigenen Identität, die in der modernen Welt durch die Informationserweiterungen des letzten Jahrhunderts noch verstärkt werden. Im Vergleich dazu werden wir in der Zukunft jedoch in der Lage sein, eine ganze Reihe von Informationen - den Inhalt von Büchern und Studien, kleinere Bibliotheken – mit Gehirnimplantat in unseren Kopf zu pflanzen und gemeinsam damit neuen grundlegenden logischen und wertverarbeitenden Mechanismen zu handhaben, die wir bisher in unserem Leben nicht hatten, und wir hatten auch nicht die Fähigkeiten dazu. Nun, dies wird den Kontakt zwischen dem Individuum und seinen Gemeinschaften grundlegend beeinflussen. In jedem Fall könnten die Kontaktgrundlagen, die auf den heutigen dauerhaften Identitäten beruhen, durch diese Änderung beseitigt werden. Nach solch einem neuen Bewusstseinsinhalt - insbesondere wenn die Werten der Kontaktpartner ergänzt und neu eingestuft wurden - kann ich nicht wissen, inwieweit mein Freund, meine Freundin, meine Frau, mein Kollege usw. gleich sind. Ob die Eigenschaften, die wir bisher in ihnen

geliebt haben, sind noch lebendig, oder auf die gleiche Weise die Erfahrungen, die wir bisher zusammen gesammelt haben und die in unserer engen Beziehung auch ohne Worte dieselbe Antwort lieferten, für ihn immer noch relevant sind. Dies kann nur durch die Möglichkeit verschärft werden, dass die Gehirnimplantanten in dem Gehirn, die unser Wissen, unsere Normen und logischen Fähigkeiten ergänzen, die mit Hilfe unseres biologischen Gehirns erworben wurden, von außen ständig neu und neu aktualisiert werden, wie wir es heute bereits wissen. Darüber hinaus können sie sich ständig mit den in den Clouds gespeicherten Informationsbasen ihrer Software verbinden. Inwieweit wird unser Freund, der mit solchen ausgestattet ist, vertraut bleiben, auf den wir uns verlassen können, denn "doch wir kennen ihn!" ?!

Diese Identitätsfrage erstreckt sich auch auf rechtliche und moralische Probleme. Inwieweit kann ich jemanden für sein früheres Verhalten respektieren oder einfach nur verachten, da er nach einer Gehirnauffrischung entweder ein „moralischer Athlet“ oder nur ein kalter Vorteilsüchtiger sein kann. Oder macht die rechtliche Verantwortung für die gestrige Handlung für jemanden, der seitdem anders denkt und handelt, Sinn? Die andere Seite davon ist, ob, wenn wir das Bewusstsein eines Soziopathen durch Gehirnimplantat ändern können und sein Bewusstsein teilweise gelöscht und ein neues sozialfreundliches Bewusstsein eingeführt wird, ist ein System der Bestrafung immer noch erforderlich? Und dies wirft die Frage auf, ob neben freiwilligem Gehirnimplantat auch deren Zwangsinstallation akzeptabel ist? Oder kann es zum Teil vom Staat in dieser Hinsicht in der Kindheit eine obligatorischen Untersuchung und Prüfung für alle Kinder obligatorisch gemacht werden, wie dies derzeit bei obligatorischen Impfungen der Fall ist? DiGiovanna nennt *Para-Personen* die zukünftige Menschen mit einem so ergänzten Gehirn - unter Vermeidung des Namens Cyborg, der bereits in Science-Fiction dafür erfunden wurde - und angesichts des aktuellen Standes der Laborforschung bedeutet diese Zukunft überhaupt keine ferne Zukunft, und die Wahrscheinlichkeit ihrer Verwirklichung ist hoch. Der Umgang mit den rechtlichen und moralischen Dilemmata und die Anpassung der heutigen Lösungen an die damalige Situation erfordern daher umfassende Überlegungen.

Über Para-Personen hinaus kann bei vollständig künstlichen und im Gegensatz zu heute selbstbewussten Roboterwesen, deren Wahrscheinlichkeit nicht ausgeschlossen werden kann, auch wenn sie nicht so groß ist wie die erstere, die Frage der Identität durch das Aufwerfen neuer Aspekte angegangen werden. DiGiovanna stellt der Inhalt der Identität im Zentrum, um das Dilemma der Identität in Bezug auf die Roboterwesen zu exponieren. Einige Details des Menschen und seines Bewusstseins ändern sich ständig, aber seine dauerhaften Eigenschaften und Wertvorlieben verändern sich auch über viele Jahre hinweg nur geringfügig, sodass man denjenigen, die immer in seiner Umgebung lebt, mehr oder weniger eine Identität zuschreiben kann, die auch die Veränderungen umfasst. Es ist die Langsamkeit des Wandels, die es mir ermöglicht, auch in der heutigen schnelllebigen Welt nicht von meinen bisherigen Erfahrungen mit den Motivationen und Merkmalen derer enttäuscht zu werden, die mit mir in Kontakt kommen. Aber genau das verschwindet für Roboter, die in kürzester Zeit tausend- und millionenfach schneller als Menschen in der Informationsverarbeitung und im Selbstlernen sind: „Slow change of character and appearance is part of what makes personal identity (...) But with an artificial person, sudden and radical change in both the physical and mental becomes possible“ (DiGiovanna 2017, 3011, 307.)

Dauerhafte Wertpräferenzen in der Informationsverarbeitung und die darauf basierende Zusammenarbeit sind bei Robotern bereits durch die Masse und Geschwindigkeit ihrer Informationsbeschaffung sowie durch ihr ständiges Selbstlernen und Selbstverändern problematisch. DiGovannas Vorschlag könnte auch bedeuten, dass die Möglichkeit von „Selbstbewusstsein“ und „Ich-bewusstsein“ selbst für einen zukünftigen leistungsstarken MI-Roboter überdacht werden muss. Diese setzen die dauerhafte Identität einer Person voraus,

aber dies beruht auf der Langsamkeit unserer Bewusstseinsveränderungen und damit auf der Beständigkeit unserer Informationsverarbeitung. Wenn ein künstliches Wesen von der menschlichen Kontrolle befreit und auf unabhängige Informationsverarbeitung und von dieser auf Selbstlernen und Selbstveränderung umgestellt wird, kann es tausendmal am Tag, jede Stunde und sogar jede Minute lernen und kann sich in seinen immer kürzeren neuen Zyklen verwandeln, dann verschwindet das, was wir im heutigen Menschen ein stabiles Selbstbewusstsein, Ich-bewusstsein nennen, fast. Mit dieser Betonung fügt DiGiovanna der viel diskutierten Frage auch eine neue Frage hinzu, d. H. wie die Frage nach dem Selbstbewusstsein und dem Ich-bewusstsein des zukünftigen fortgeschrittenen Roboterbewusstseins stehen wird. Und wie kann man sich moralische Wertvorstellung ohne dauerhaftes Selbstbewusstsein und Ich-bewusstsein vorstellen?

Allein aus diesem Grund muss der Gedankengang, der durch die mechanische Erweiterung des heutigen menschlichen Bildes berechnet, dass ein solches Roboterwesen im Fall von Superintelligenz wahrscheinlich auch „überethisch“ sein wird, als falsch angesehen werden (siehe Petersen 2017). In diesem Zusammenhang ist es aber auch notwendig, die Erklärungen und Analysen, die im Falle der Entwicklung von Robotern mit eigenem Bewusstsein - in menschlicher Analogie - die Anerkennung ihrer moralischen Bedürfnisse und die Gewährung von Menschenrechten in ihren Schriften vorsehen, umfassender zu behandeln. Denn fassen diese Analysen die zukünftigen Roboter einfach als eine neue Art menschlichen Begleiters und als Erweiterung der menschlichen Existenz auf. Sobald die Programme der Roboter die Emotionen in ihre Algorithmen aufgenommen haben, fordern diese Analysen, dass die Gesellschaft auf die Emotionen von Robotern achtet und ihnen auch Menschenrechte gewährt: „Es muss wahrscheinlich gesetzlich geregelt werden, wie viel Schmerz und Gefahr ein Roboter ausgesetzt sein kann. (...) Es könnte leicht sein, dass dies zu weiteren ethischen Debatten über andere Rechte von Robotern führen würde. Können Roboter Eigentum besitzen? Was passiert, wenn jemand versehentlich verletzt wird? Können sie verklagt oder bestraft werden? Wer ist im Falle einer Klage für sie verantwortlich? Kann ein Roboter einen anderen Roboter besitzen? Aus solchen Fragen ergibt sich dann eine andere Frage: Sollten Roboter einen ethischen Sinn erhalten? (Kaku 2014, 251). Unsere vorherigen Erklärungen beantworteten mehrere Fragen aus diesen, basierend auf den in der Zwischenzeit durchgeführten Robotikstudien, aber das Grundproblem dahinter sollte auch hervorgehoben werden, da ganze Studien und Bände aus ähnlichen Annahmen entstanden sind, z. B. ein neuer Band in diesem Bereich, herausgegeben von Jason P. Doherty: *AI Civil Rights: Addressing Artificial Intelligence and Robot Rights*.

Nun, das Grundproblem bei diesem Gedankengang besteht darin, dass er die Tatsache ignoriert, dass Rechte und ethische Anforderungen bei Robotern nur entstehen können, wenn Ich-Bewusstsein und Selbstbewusstsein geschaffen werden. Es bedeutet aber auch, dass sie, wenn dies wirklich in der Zukunft geschieht, gleichzeitig durch die tausendfache Entwicklung von der menschlichen Kontrolle befreit und als separate neue Seinsschicht über die bisherige vier Seinsschicht der menschlicher Gesellschaften aufgebaut werden. Von dieser Zeit an wären aber ihnen die gesamte biologische Sphäre und die damit verbundenen menschlichen Gesellschaften gleichgültig und würden keinen „gerichtlichen Schutz“ benötigen. Das heißt, eine Roboterwelt, die dieses Niveau erreicht hat, wäre nicht Teil der menschlichen Gesellschaft als „neuer Kamerad“ in der Herrschaft über die Welt, sondern als die menschliche Existenz aus der Welt der Primaten hervorging und sich über die tierbiologische Seinsschicht erhob und immer autonomer wurde, so erhebt sich nun der künstliche Maschinenintellekt, losgelöst von biologischen Voraussetzungen, über die menschliche Gesellschaft. Im Gegensatz zum bisherigen Aufbau immer neuerer Seinsschichten auf den unteren Seinsschichten würde die neue Seinsschicht der künstlichen Intelligenz nur die unterste physischen Seinsschicht brauchen, und für sie wäre die biologische und psychische-

emotionale Seinschichten nicht erforderlich. Diese Roboterwesen würden keine Rechte und ethischen Forderungen benötigen, sondern sie werden die gesamte Realität, einschließlich der menschlichen Gesellschaften, dominieren, wie wir Menschen heute die vierschichtige irdische Welt dominieren. Auf diese Weise haben eher solche Analysen Recht, die diskutieren, ob, wenn die Roboterwelt wirklich dieses Niveau erreicht, was mit der Menschheit passieren wird ?!

7. Das moralische Credo von „Unabomber“, dem „verrückten Mathematiker“

Mitte der neunziger Jahre, nach vielen Jahren einer Reihe von Bombenanschlägen und einer Verfolgungsjagd des FBI, gab ein geheimer Täter namens „Unabomber“ den Grund für sein Vorgehen in einer anderthalbhundertseitigen Broschüre an, die er gegen die Unmenschlichkeit der Entwicklung der technologischen Gesellschaft seit der industriellen Revolution aussprach. Seine eigentümliche Sprache wurde von seinem Bruder erkannt und durch Benachrichtigung des FBI wurde der lang ersehnte Bomber gefangen genommen. Es stellte sich heraus, dass *Theodor John Kaczynski*, ein Mathematiker aus Harvard. Er wurde zu einem Zeitpunkt seiner Universitätskarriere zum Feind einer von Technologie geprägten Gesellschaft und begann seine Explosionsserie, in der die Entwickler und Hauptnutzer dieser Technologie gezielt wurden. Mehrere starben und weitere wurden dabei verwundet, und er plante, noch mehr Vergeltungsmaßnahmen durchzuführen, wenn sie nicht verhaftet worden wären. Nachdem die exponentiell sich entwickelnden Auswirkungen der Technologieentwicklung in den letzten dreißig Jahren wirklich unbestritten geworden sind und das Ausmaß und die Auswirkungen ihrer weiteren Beschleunigung bereits Gegenstand mehrerer umfassender Analysen sind, lohnt es sich, sich erneut auf Argumente von Unabomber, dem „verrückten Mathematiker“ zu konzentrieren. Dies ist, was *Jai Galliot* in seiner neuen Studie tut. Er setzt den Widerstandskämpfer, der seitdem in seiner Gefängniszelle damit beschäftigt ist, seine Thesen weitzuzuentwickeln, unter Theoretikern und Bewegungen der Antitechnologie und es wird versucht, seine Hauptthesen im Lichte des aktuellen Zustands der Roboterwelt hervorzuheben (Galliot, 2017, 369-385).

Kaczynski hat nur die praktischen Schlussfolgerungen der früheren Thesen von *Jacques Ellul* 1964 erschienenem Band *"The Technological Society"* gezogen, die auf ihre Weise auch eine Fortsetzung von *Oswald Spengler* 1922er Arbeit zur Analyse des Niedergangs der westlichen Zivilisation darstellten. Von beiden Verfassern wurde der Niedergang mit der technologischen Entwicklung erklärt (Spengler, 1995). Der rein pessimistische und resignierte Ton bei Spengler und Ellul wurde dann zu einem moralischen Widerstand im Falle von Kaczynski, und nachdem er gesehen hatte, dass es keine Möglichkeit gab, diese Entwicklung zu reformieren, glaubte er, dass nur revolutionäre Gewalt übrig blieb, um die Zerstörung der Menschheit zu verhindern. Jahrzehnte nach seiner Broschüre lohnt es sich nun zu überdenken, wie der aktuelle Zustand der Roboterwelt und die radikaleren Veränderungen, die bereits weitgehend sichtbar sind, die Gefährdung der Menschheit oder zumindest eine erhebliche Verschlechterung ihres Zustands bedeuten könnten.

Als Ausgangspunkt für ihr Vorgehen gegen die technologische Gesellschaft ist hervorzuheben, dass sowohl Spengler als auch Ellul und Kaczynski die menschliche Existenz als eingebettet in die physikalisch-biologische Umwelt betrachten. Daraus schließen sie dann, dass die menschliche Existenz zerstört wird, wenn das menschliche Leben infolge der industriellen Revolution immer mehr technologisch vermittelt wird und auf diese Weise von der physikalisch-biologischen Umwelt immer weiter entfernt wird: „Ellul wrote that „the machine trends not only to create a new human environment, but also to modify man’s very

essence” and that „the milieu in which he lives is no longer his. He must adopt himself, as though the world were new, to a universe for which he was not created. Kaczynski shares this sentiment.” (Galliott 2017, 373). Wenn wir andererseits Hartmanns These berücksichtigen, die die vier voneinander abhängigen Seinsschichten des menschlichen Lebens im Auge behält (physisch, biologisch, mental und intellektuell) und die im Verlauf der Evolution von der immer stärkeren transformierenden Wirkung der oberen Schichten auf die unteren Schichten ausgeht, ist die obige These übertrieben und ohne Grund und sie muss als zu pessimistisch eingestuft werden.

Kaczynski und seine Vorgänger sehen es als Zerfall des menschlichen Lebens an, wenn die oberste intellektuelle Schicht der vier Seinsschicht gegenüber den unteren immer dominanter wird. Dies war jedoch der Fall, wenn auch langsamer, in den letzten zwei- oder dreitausend Jahren und man kann die Benutzung der Metallen und insbesondere der Eisen herausheben, von der die Transformation der menschlichen Umwelt grundlegend verändert wurde. Die industrielle Revolution hat dies nur beschleunigt, und insbesondere seit den 1950er Jahren ist es stürmisch geworden, die verschiedenen Aktivitäten menschlicher Gemeinschaften auf Intelligenz und die damit verbundene Technologie zu stützen. Das heißt, menschliches Leben basiert keineswegs nur auf der physikalisch-biologischen Seinsschichten. Wenn also ihre Anteil und ihre entscheidende Kraft im menschlichen Leben abnehmen und diese Umgebung technologisch weit vermittelt und transformiert wird, bedeutet dies nicht, dass die menschliche Gesellschaft zerstört wird. Bei alledem verschiebt sich nur das Gewicht der Bedeutung der vier Seinsschichten der Realität, wodurch sich das menschliche Leben stärker auf der Grundlage der intellektuellen Schicht basiert und die Dominanz dieser geistlichen Seinsschicht über den unteren radikal erhöht wird. Diese Einschätzung von uns könnte nur ausgesetzt werden, wenn diese Welt irgendwann in der Entwicklung der Roboterwelt wirklich aus menschlicher Kontrolle hervorgehen würde und sich die künstliche Intelligenz als eine neue Seinsschicht über die menschlichen Gesellschaften erheben würde, die sich bisher auf dem Höhepunkt der Evolution befanden. Die Erhöhung von Kaczynski als Prophet würde dann nur dadurch verhindert, dass unter solchen Umständen und seinen Gefahren das Versäumnis, als Held erhöht zu werden, das geringste Problem wäre. Dies kann jedoch nach bestem Wissen und Gewissen heute nur als unwahrscheinliche Option angesehen werden, und vielmehr kann die Zunahme der von künstlicher Intelligenz geprägten menschlichen Gesellschaften ohne eine neue autonome Seinsschicht als realistische Vision für die Zukunft angesehen werden.

Literature

- Abney, Keith, „Robotics, Ethical Theory and Metaethics: A Guide for the Perplexed, in Patrick Lin, Keith Abney, George A. Bekey (eds.), Robotethics, The MIT Press, Cambridge Massachusets. London 2011. pp. 35-54.
- Allen, Collin, Wendell Wallach, „Moral Machines: Contradiction in Term or Abdication of Human Responsibility?“, in Patrick Lin, Keith Abney, George A. Bekey (eds.), Robotethics, The MIT Press, Cambridge Massachusets. London 2011. pp. 55-68.
- DiGiovanna, James, „Artificial Identity“, in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp. 307-321.
- Doherty, Jason P. (ed.), AI Civil Rights: Addressing Artificial Intelligence and Robot Rights. Kindle Edition, 2016.
- Ford, Martin, The Rise of Robots: Technology and the Threat of a Jobless Future, Basic Books, 2016.
- Galliott, Jai, „The Unabomber on Robots“, in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp. 369-385.
- Hartmann, Nicolai, „Das Problem des geistigen Seins. Zur Grundlegung der

- Geschichtsphilosophie und der Geisteswissenschaften. Walter de Gruyter Verlag. Berlin 1962.
- Hegel, Georg Wilhelm Friedrich, „Grundlinie der Philosophie des Rechts. (Werke Band 7.) Suhkamp, Frankfurt am Main. 1979. S. 301.).
- Henschke, Adam, „The Internet of Things and Dual Layers of Ethical Concern” in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp. 229-243.
- Kaku, Michio (2014): The future of the mind. The Scietific quest to understand, enhance and empower the mind. Doubleday. New York. 2014.
- Kelly, Kevin, The Inevitable: Understanding the 12 Technological Forces That Shape Our Future, Penguin Books, 2014.
- Klinewicz, Michal, „Challenges to Engineering Moral Reasoners” in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp. 244-257.
- Loh, Wulf, Janina Loh, „Autonomy and Responsibility in Hybrid System, in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp. 35-50.
- Luhmann, Niklas, Liebe als Passion: Zur Codierung von Intimität, Suhrkamp, Frankfurt am Main 1994.
- Pokol, Béla, Theoretische Soziologie und Rechtstheorie. Kritik und Korrigierung der Theorie von Niklas Luhmann.. Passau. Schenk Verlag. 2013.
- Pokol Béla, Künstliche Intelligenz: Die Entstehung einer neuen Seinschcht? (Ki – im Spiegel von Nicolai Hartmanns Ontologie.) Pázmány Law Working Papers No.2018/12
- Splengler, Oswald, Der Untergang des Abendlandes. Albatros Verlag/Patmos Verlag. 2007.
- Talbot, Brian, Ryan Jenkins, Duncan Purves, „When Robots Should Do the Wrong Thing”, in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp.258 -273.
- Vallor, Shannon, George A. Bekey, „Artificial Intelligfnce and the Ethics of Self-Learning Robots”, in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp. 338-353.
- White, Trevor N, Seth D. Baum, „Liability for Present and Future Robotics Technology”, in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp. 66-79.
- Zoller, David, „Skilled Perception, Authenticity, and the Case Against Automation”, in Patrick Lin, Ryan Jenkins, Keith Abney (eds.), Robot Ethics 2.0, Oxford University Press, New York 2017. pp. 55-68.

4

Eine Kritik der Verengung der Moral auf Gerechtigkeit

Es gibt verschiedene Klassifizierungen im Bereich der Moralphilosophien und die meistbenutzte ist vielleicht die Version von „konsequentialist versus deontologisch“. Inhaltlich ist es besser, wenn diese Klassifizierung als „Gemeinschaftsmoral bejahende“ versus „Gemeinschaftsmoral vernichtende“ Unterscheidung genannt wird, denn entwickelten die herrschenden moraltheoretischen Positionen sich in der letzten Konsequenz zu diesem scharfen Gegenüberstehen. Diese kleine Schrift möchte diese Entwicklung skizzieren und herausheben, dass die anfängliche Unterdrückung der Gemeinschaftsmoral in der deontologischen Moralphilosophie in der letzten Konsequenz bei Habermas in der Eliminierung der Moral zugunsten dem Recht überhaupt endete.

1. Kollektivmoral und persönliche Moral

In den letzten Jahrhunderten können zwei gegensätzlichen Positionen im Bereich der Moraltheorie als vorherrschend erfasst werden. Eine von ihnen, die von Hegel und später von Jheringhez und Nicolai Hartmannhoz und heute hauptsächlich von den kommunitarischen Moraltheorien (z. B. McIntyre, Sandel, Walzer, Taylor) formuliert worden ist, behauptet, dass es in dem gemeinschaftlichen Leben von Generationen aufgespeicherten Mengen der Moralnomen gibt und diese Normen von den Mitglieder der immer neueren Generationen nur sozialisiert und angeeignet werden und dann sie als Gewissen die Handlungen und Erlebnisse der Individuen innerlich in bestimmte Richtungen lenken. Als Nebenseite dieser Position kann erwähnt werden, dass die existierenden Moralnomen von den bewussten moraltheoretischen Reflexionen nur bloßgelegt aber nicht geändert werden können. Die Gegenposition wurde von Immenuale Kant in der individualistischen Geistigkeit der Aufklärung gegründet und heute vor allem von John Rawls und Habermas repräsentiert. Diese Position legt den moralischen Aspekt in das individuelle Bewusstsein und dadurch sieht er implicate von den verbreiteten gemeinschaftlichen Normen der Richtigkeit schon in dem Ausgangspunkt ab. Daraus folgt es, dass es für die Ausgestaltung der individuellen moralischen Entscheidung von der schon existierenden gemeinschaftlichn Moralnomen keine Schranke gibt, und diese Annulierung der gemeinschaftlichen Moral kann die methodische Folge des Ausgangspunkts - wie bei Kant - sein, oder sie kann mit einem Hinweis auf die historische Änderung - wie bei Habermas – begründet werden, dass die einst noch wichtige konventionelle Moral in der Modernität schon aufgelöst hatte und nur die auf bewussten Diskussionen basierte Diskursmoral existiert (oder dürfte existieren).

In den folgenden Analyse werden zuerst die zwei entgegengesetzten moraltheoretischen Positionen kontrastiert; aber vor der Analyse der Diskursmoral muss die empirische Haltbarkeit der erwähnten Behauptung von Habermas über das Verschwinden der gemeinschaftlichen Moral untersucht werden und dafür werden die Ergebnisse der Untersuchung von Lawrence Kohlberg dargelegt.

1.1. Rudolf von Jherings Theorie über Moral und Sitte

Schauen wir uns kurz den Ausgangspunkt von Hegels Rechts- und Moralphilosophie, deren Fußstapfen Jhering trat. Hegels Position, von der die objektive Sittlichkeit über den Individuen hergehoben wurde, lautet als folgende: „Aber in der einfachen Identität mit der Wirklichkeit der Individuen erscheint das Sittliche, als die allgemeine Handlungsweise derselben, als Sitte, - die Gewohnheit desselben als die zweite Natur, die an die Stelle des ersten bloß natürlichen Willens gesetzt und die durchdringende Seele, Bedeutung und Wirklichkeit ihres Daseins ist, der als eine Welt lebendige und vorhandene Geist, dessen Substanz so erst als Geist ist (Hegel:1979:301). In dem Anhang zu dieser Passage hebt er heraus, dass die Pädagogik solche Kunst ist, durch die die Menschen die Normen der objektiven Sittlichkeit aneignen und die aneignete Sittlichkeit nur dann als stabile Führung der individuellen Handlungen funktionieren, wenn sie zur Gewohnheit der einzelnen Individuen wurden. So gehören die Sittlichkeit und die Gewohnheit zueinander.

In seiner Moralthorie folgte Rudolf von Jhering in den 1870er Jahren Hegels Ausgangspunkte und er hob heraus, dass das Moralgefühl in dem individuellen Bewusstsein der Gemeinschaftsmitglieder, das von der dominierenden Auffassung als letzter Träger der Moral angesehen wird, nur das Derivat ist und der wirkliche Träger der Moral die Bewusstsein der Gemeinschaft ist. Von den einzelnen Individuen werden diese schon in der Gemeinschaft existierenden Moralnormen immer nur sozialisiert und obwohl die Entscheidungen über die richtigen Handlungsweisen von den Individuen als Folge der Führung ihrer Gewissen subjektiv aufgefasst werden, dürfen die Forscher der Moral diesen Irrtum nicht übernehmen: „Das Verhältniss der objektiven sittlichen Ordnung, zu der ich neben dem Rechte auch die Moral und Sitte zähle, und des subjektiven sittlichen Gefühls drehte sich für mich gänzlich um, nicht letzteres erschien mir mehr als die Quelle der ersteren, wie die herrschende Theorie lehrt, sondern erstere als die des letzteren. Alle sittliche Normen und Einrichtungen haben nach meiner Überzeugung ihren letzten Grund in den praktischen Zwecken der Gesellschaft“ (Jhering 1898:X.).

Die Moral und die Sitte funktionieren als die notwendigen Vorbedingungen der Gesellschaft und Jhering nimmt die Erklärung für ihre historischen Formen von der sozialen Evolution. Die initiale Verflechtung der Lebensbedingungen der uralten Gemeinschaften mit ihrer physischen Umgebung, die nur mit körperlichen Gewalt dominieren werden konnte, identifizierte die Tugenden mit der physischen Stärke und die lateinischen „vis“, „virtue“ zeigen in vielen heutigen Sprache diesen Ursprung. Erst in den entwickelteren Niveau erscheint die Identifizierung der moralischen Tugenden mit der sublimierten Geistlichkeiten und entstammen die Wörter und Begriffe der moralischen Werten. In dieser Entwicklung wurde die undifferenzierte Menge der Normen der Richtigkeit, von der hauptsächlich immer die äusseren Folgen der Handlungen beurteilt und die äussere Formen der Handlungsweisen betont wurden, differenziert und es kommt das Recht, die äussere Benehmen bestimmenden Sitte (Umgangsformen, Höflichkeit) und das innere Gefühl bewegende Moral in der europäischen Geschichte zustande.

Ausser dem Recht benutzt Jhering drei Wörter für die Abschreibung der moralisch-sittlichen Welt. Die eine ist die Sitte, die einst noch die gesamte Sphäre umfasste jetzt aber sie bedeutet nur noch die Normen der äusseren Umgangsformen und der angemessenen Sprachformen. Die andere ist die Sittlichkeit, von der schon die innere gefühlsmässige Haltung neben des äusseren Benehmens involviert ist, und die dritte ist die Moral, die schon vollkommen die Gefühlsmässigkeit der Handlung bedeutet. Die beide letzte können in den Texten von Jhering nur schwer unterscheiden und zwischen den klaren Unterschied der Sitte und der Moral kann die Sittlichkeit als gemischte Typ der Normen genannt werde, die

einigermaßen sowohl die Charakterzüge der Ausserlichkeit der Sitte als auch der Innerlichkeit der Moral trägt. Als Übergang dieser Formen können die Bedeutungen der deutschen Wörter der „Schichlichkeit“ und des „Andstands“ erwähnen, die hauptsächlich die Ausserlichkeit der Handlungen bedeuten, aber sie beinhalten bestimmte innere Gefühlsmässigkeit sowohl bei der Handlung als auch bei der Beurteilung solcher Handlungen (siehe Jhering 1898:32-36).

Um die historische Differenzierung der Normenwelt in der europäischen Kultur darzustellen, gibt Jhering eine kurze Beschreibung. In seiner Analyse umfasste die griechische *Dike* noch die ganze Normenwelt der richtigen Handlungsweisen und sogar die ästhetische Seite der Handlung wurde im Begriff von *Dike* beinhaltet und so bedeutete das Gute auch die Schöne. Und dass die unteren Stufen der menschlichen Gemeinschaften von dieser Regel geprägt war, zeigen auch die alten Hebräer, wo die „mischpat“ ebenso die ganze rechtliche, moralische Normenwelt und Sitte gemeinsam bedeutete - einzig nur im solchen Unterschied zu der griechischen *Dike*, dass bei der „mischpat“ die göttliche Herkunft der Gebote der Normen stärker betont wurde. „Alles ist Mischpat: das Ritualgesetz sowohl, welches unserer „Sitte“ entspricht, wie die zehn Gebote, in denen Moral und Recht noch ununterschieden neben einander liegen. Ganz dasselbe gilt vom dharma der Inder“ (Jhering 1898:52-53). Diese Einheit ist dann bei den Römern beendet, bei denen das Recht und die Moral differenziert wurden aber die heutige Sitte war noch in der Moral beinhaltet. Für den Unterschied des blossen äusserlichen Benehmen und der gefühlsmässigen moralischen Handlung hatten die Römer keine spezielle Ausdrücke, und beide wurden mit den „mos“ „mores“ bezeichnet. In der sprachlichen Entwicklung bekommt die äussere Richtigkeit der Handlung - isoliert von der inneren Gefühlsmässigkeit begleitenden Handlung - mit der „Sitte“ einen abgesonderten Ausdruck in der deutschen Sprache und im Gegensatz dazu bezeichnen die Sittlichkeit und Moral klar die innerlichen Aspekte. Und als letzte Phase der Differenzierung der Normenwelt kann die Entstehung der Mode erwähnt werden, von der die äusseren menschlichen Handlungen neben der Sitte auch bestimmt sind.

Jhering erläutert den Unterschied zwischen moralischen Gut und Böse von der Existenzbedingungen der Gesellschaften und im Gegensatz zu der kantischen Ethik sind die menschlichen Handlungen nicht innerlich Gut oder Böse sondern obwohl sehr sublimierte Weise aber es hängt davon ab, ob bestimmte Handlungen zu den Zwecken und Funktionen der Gesellschaft und der reibungslosen Existenz der Gesellschaft einen Beitrag leisten kann oder nicht. Das Gute und das Böse werden ebenso von der Seite der Gesellschaft bestimmt, wie alle anderen Normen und in den unterschiedlichen Zuständen der Gesellschaften (und im Bezug auf die dazu gehörenden Bedingungen) qualifizieren sich derselben Normen unterschiedlich Gut oder Böse. Z.B. was im Zustand der Frieden die grösste Sünde bedeutet, ist im Zustand des Krieges ein moralisches Gebot. „Gut in physischen Sinne nennen wir, was unsern Körper oder unserer Empfindung wohl thut, böse oder schlecht, was uns schadet oder unangenehm berührt. Ganz in demselben Weise nennt die Gesellschaft von ihrem Standpunkt aus in Bezug auf das menschliche Handeln gut im sittlichen Sinne dasjenige, was ihr Dasein fördert oder ihr Wohlbefinden erhöht, schlecht, böse dasjenige, was demselben Abbruch thut (Jhering 1898:218-219). Das moralische Pflicht bekommt bei Jhering ebenso von der Seite der Existenzbedingungen der Gesellschaft eine Erläuterung: „Pflicht ist das Bestimmungsverhältniss der Person für die Zwecke der Gesellschaft“ (ders. 227.p.). Das Individuum kann seine eigene Fortdauern und Existenz mit seinem Egoismus und mit der Bestrebung für die Lust sichern und für seine individuelle Existenz notwendige Fortdauern der Gemeinschaft sichert er durch die Beachtung der moralischen Tugenden und Normen.

1.2. Nicolai Hartmanns Ethik

Hartmann brachte seine Ethik in den späten 1920er Jahren gegen Kants subjektive Moral anhand Helgels Auffassung zustande, besonders in seinem Monographie in 1932, wo er den geistigen Aufbau der sozialen Welt analysiert und die individuellen spirituellen Entwicklung als die Durchdringung des Bewusstseins mit der aufgespeicherten geistigen Werten, Normen, Symbolen etc. darstellt (Hartmann 1962; 1962b). Auch in seiner allgemeineren These für das Verhältnis zwischen den geistigen Sektoren der Gemeinschaften und der Entwicklung des individuellen Bewusstseins beobachtete er den Vorgang der individuellen Entwicklung als die geistige Umwandlung der bloßen psychologische Seele, durch die die Mitglieder der immer neueren Generationen die geistigen Inhalten erwerben. Hartmann hebt heraus, dass es in der menschlichen Gemeinschaften objektivierte geistigen Inhalten - die rechtlichen Normen, moralische Werte und Regeln, Kunstverfahren und Kenntnisse - gibt und durch die Aneignung dieser Inhalte kommt ein Moralgefühl, Rechtsgefühl etc. in dem individuellen Bewusstsein zustande, und auf diese Weise bekommen die geistige Inhalte eine Führungsrolle in den menschlichen Handlungen. Das bedeutet aber auch, dass die geistigen Werten, Prinzipien, Normen und Kenntnisse in dem menschlichen Denken und Handeln allein nicht realisieren können nur gemeinsam mit ihren gefühlsmässigen Phänomenen. Eben deshalb betont Hartmann die Möglichkeit der nur langsamen Änderung der Moral und sieht er die Bestrebungen der Umformung der Moral durch die Moralphilosophie als vergeblich und als nur störend: „Das Ethos aber vergegenständlicht sich auch in der „objektivierten Moral“. Damit tritt ein zweites Moment neben das rein empfundene Gelten, gleichsam als zweites, äußeres Gelten. Moral eben ist mehr als lebendiges Ethos. Sie ist die Ausprägung gewisser Typen wertvollen Verhaltens in bestimmten Begriffen, in denen die Wertnuancen eingefangen und für jedermann verständlich objektiviert ist. Begriffe solcher Art sind wohlbekannt als die von „Tugenden“; ihre ebenso festgeprägten Gegenstücke sind Begriffe von „Untugenden“, oder „Laster“. Aller herrschend gewordene „Moral“ bewegt sich in solchen Begriffen. Aber ebenso wohlbekannt wie sie selbst ist auch das Odium, das ihnen anhaftet. Nichts ist im lebenden Geiste steriler, nichts neigt so sehr zum „Absinken“, nichts ist so hemmend im Fortschreiten des Ethos als die Tugend- und Laster-begriffe. Nichts zieht so sehr wie sie die Moral herab von ihrem echten Wert- und Ideengehalt. Sie sind es, die auf die Dauer stets der echten Moral ein „Moral“ im schlechten Sinne vorschreiben, mit der sich dann freilich sehr bequem und stereotyp „moralisieren“ läßt. (...) Die wirkliche Moralität, wo überhaupt sie besteht, läßt die objektivierte Moral hinter sich, empfindet sie als leblosen Formenkodex, der mit ihrer Bewegung nicht Schritt hält“ (Hartmann 1962:526-527.p.).

Als Ausgangspunkt behauptet Hartmann als Leitfaden seiner Analyse, dass es in der menschlichen Geschichte viele moralische Systeme und Haufen der Moralprinzipien gibt und diese dürfen nicht in solcher Weise zur Einheit gebracht werden, dass man die Vielheit beseitigend nur eine von ihnen zu dem Zentrum stellt, sondern die einzelnen Teile dieser Vielheit muss in Relationen gestellt werden und Schritt für Schritt eine konkrete Einheit daraus zustandegerbacht werden. Auf diese Weise bleibt die bunte Vielheit der Moral lebendig und die Verarmung der abstrakten Moralsysteme kann vermieden werden. Er behauptet, dass es seit der Kantischen Philosophie eine Tendenz in der Moralanalyse gibt, die Realität der bunten moralischen Werten zu beseitigen Realitäten, und die Moral als ein Wahl zwischen das abstrakte Gut und Böse darzustellen: „Seit dem Aufkommen der „kritischen Philosophie“ ist man gemeinhin besorgt gewesen, so wenig als möglich Gegebenes anzuerkennen, die Basis von Voraussetzungen so schmal als möglich anzulegen – aus dem sehr einleuchtenden Gefühl heraus, daß jedes hinzugenommene „Gegebene“ angefochten werden und dadurch dem Basis zum Einsturz gereichen kann. Diese Tendenz führte zur Auslese des Gegebenen.“ (Hartmann 1962:63). Dann hebt er mit Hinweis auf Max Scheler heraus, dass er sein Zeitalter sieht als solche, in dem der Bruch mit dem einseitigen und

abstrakten Moralsysteme stattgefunden hatte und in der Moraltheorie die bunte Vielheit der moralischen Werten und Prinzipien neulich zum Zentrum gezogen wurde.

Zum Verstehen der Ablehnung der Gemeinschaftsmoral durch Kant ist Hartmanns Hervorhebung wichtig, die als Folge der breiteren theoretischen Rahmen der Kantischen Philosophie diese Ablehnung darstellt. Denn geht die ganze Kantische Philosophie von der Subjekt aus und so zum Beispiel sind der Raum und die Zeit keine objektiven Kategorien sondern nur die Ordnungskategorien des Bewusstseins und die Einheit der Objekte stammt nicht von der Außenwelt, sondern von der synthetischen Einheit des Bewusstseins. Also sind die Subjektivität und das Bewußtsein primär in der ganzen Kantischen Gedankenwelt und die Objekte und Gegenständlichkeit sind die Folgen dieser subjektiven inneren Welt und sie existieren nicht objektiv „draussen“. Daraus folgt auch metodisch die Annulierung aller äusseren geistigen Inhalten und darin die Gemeinschaftsmoral und statt dieser die These der subjektiven Zustandbringen der moralischen Handlungswiese im Rahmen der kategorischen Imperativ: „So ist es sehr wohl verstehen, daß gerade in der Ethik die Kantische These von der Spontaneität des Subjekts eine Art Bestätigung erfährt. (...) Das „Sittengesetz“ drückt eine Forderung aus im Gegensatz zu den realen Verhältnissen des Menschenleben. (...) Es liegt also durschau Konsequenz darin, wenn die Kritik der praktischen Vernunft behauptet, das Subjekt gebe hier das Gesetz; in ihm und nicht im Objekt liege der Bestimmungsgrund“ (Hartmann 1962:100)

In dem Blickpunkt unserer Analyse - der Kritik der Verengung der Moral – können die wichtigsten Einwände Hartmanns gegen die Kantischen Ethik dort gefunden werden, wo er die bunten Vielheit und die diversen moralischen Werten und Tugenden bei Platon, Aristoteles und in der stoischen Ethiken darstellt, und behauptet, dass Kant die moralische Welt verengte und an Stelle der äusseren diversen moralische Werten und Tugenden die eindimensionalen Schöpfungen des Bewusstseins setzte: „Denn Kant ist es, der an Stelle der inhaltlichen Mannigfaltigkeit der Tugenden die Einheit eines Sittengesetzes, an Stelle der materialen Erfülltheit das formale Prinzip, an Stelle des objektiven Wesens sittlicher Ideen die subjektive Gesetzgebung gesetzt hat“ (i.m. 137.p.)

Hartmann wiederholt oft seine Behauptung, dass immer eine Vielzahl von moralischen Werten und Tugenden in den einzelnen Situationen für die Handlungen auftauchen und man zwischen ihnen immer gewichten soll und auf diese Weise hilft die blosse Kenntnis der abstrakten Werten nicht sondern man muss immer konkretisiert die Relationen und Rangordnung der moralischen Werten und Tugenden kennen: „Alle Werterkenntnis bleibt abstrakt ohne Erkenntnis der Wertbeziehungen. Sind doch in allen ethischen Situationen mannigfache Werte zugleich beteiligt, und liegt doch für den Menschen, der vor Situation gestellt ist, die Aufgabe eben darin, sein Verhalten aus dem Situationsbewußtsein heraus einzurichten, welches Wert gegen Wert abwägt“ (i.m. 271.p.) In den meisten Situationen - sogar wo man nur als klare Ausfluss der Böswilligkeit sieht - stehen Werte gegen Werte, obwohl es sein kann, dass wir die Werte der Handelnden vollkommen verurteilen und auf diese Weise wir nur als blosse Unwert seine Handlung sehen kann. Aus seinem Blickwinkel strebte auch er aber für Werte.

Vor Abschluss der Analyse von Hartmanns Ethik lohnt es sich noch, seine Stellungnahme gegen die eindimensionale Verengung der Moral zu zitieren, wo er heraushebt, dass die moralische Tugend nur als Synthese mehrerer Werte zustandegebracht werden kann: „Gesucht ist gerade ein Einheitsethos von Reinheit und Fülle, von Gerechtigkeit und Nächstenliebe, von Stolz und Demut. Erst ein solches würde in einem höheren und strengeren Sinne „Tugend“ heißen dürfen, während die einseitigen Werte nur uneigentlich den Namen verdienen“ (Hartmann 1962:574). In einer anderen Fassung lautet dieser Gedanke: „Jeder Wert hat - wenn einmal Macht gewonnen hat über eine Person - die Tendenz, sich zum alleinigen Tyrannen des ganzen menschlichen Ethos aufzuwerfen, und

zwar auf Kosten anderer Werten, auch solcher, die ihm nicht material entgegengesetzt sind (...) So gibt es einen Fanatismus der Gerechtigkeit (fiat justitia pereat mundus), der keineswegs bloß der Liebe, geschweige denn bloß der Nächstenliebe, ins Gesicht schlägt, sondern schlechterdings allen höheren Werten" (Hartmann 1962:576).

2. Moralsoziologische Untersuchung

Vor dem Übergang zur Darstellung der individuellen Moraltheorie lohnt es sich zu zeigen, was die Ergebnisse der empirischen Moralsoziologie in Bezug auf die Beachtung der Normen der Kollektivmoral beweisen. Die Frage ist, ob die Mitglieder der jeweiligen neuen Generationen grösstenteils mit der Beachtung der verbreiteten Normen der kollektiven Moral charakterisiert werden können - wie die Theorie von Hegel, Jhering und Hartmann behauptet - oder ob die Kollektivmoral schon aufgehört zu existieren und die Individuen in bewussten Diskussionen immer situative die moralisch richtigen Handlungen aushandeln, wie die Diskursethik von Habermas behauptet. Für die Antwort stehen die Ergebnisse der Untersuchungen von Lawrence Kohlberg zur Verfügung.

Diese Ergebnisse zeigen in Bezug auf die Entwicklung des Moralbewusstseins der Kinder und der Jungen das folgende Bild (Kohlberg 1971; Lind 1981:225-42). Die moralische Entwicklung der Kinder zeigt verschiedene Phasen und zuerst sind nur die Vermeidung der Bestrafung, die Bestrebung nach Belohnung und die Beachtung der äusseren Autorität von Bedeutung und diese Haltung ändert sich bei den Jungen, als sie autonomer wurden und hier werden die abstraktere Normen der Rollen der „brave Junge“, des „netten Mädchen“, später der „anständigen Frau“ etc. befolgt. Das ist die Phase der konventionellen Moral und Kohlberg unterstellte, dass es eine nachkonventionelle Phase (darin zwei innere Stufen) gibt, wo die Beachtung der äusseren gemeinschaftlichen Moralnormen abnimmt und die Entscheidung aufgrund der universellen Prinzipien zustandekommt. Die Untersuchungen wurden in mehreren Ländern - USA, Türkei, Mexiko, Thailand - wiederholt durchgeführt und die Ergebnisse immer zeigten, dass grösstenteils die konventionelle Moral wirkt. Nur in der unteren Stufe der zweistufigen postkonventionellen Phase konnten die empirischen Untersuchungen 7% der Menschen finden, die eine kritische und reflexive Haltung in Bezug auf die Kollektivmoral hatten, aber fast niemand konnte in der letzten Stufe gefunden werden, wo die individuellen Entscheidungen tatsächlich aufgrund den universellen Moralprinzipien und den abstrakten Menschenrechten gefasst wurden (siehe Rest/Narvaez/Bebeau/Thoma 1999:22)). Sogar gab es eine Kritik, dass die 7% nur deshalb in der unteren Stufe der postkonventionellen Phase aufgefunden werden konnte, weil die hochgebildeten und grossstädtigen Menschen in der Erhebung zu stark repräsentiert waren und normalerweise ist diese Zahl noch geringer.

Aus den Ergebnisse von Kohlberg ist es also klar, dass die Kollektivmoral in den damaligen Gesellschaften existiert und für die überwiegende Teil der Menschen sind die Normen der Kollektivmoral von grössten Bedeutung und in den alltäglichen Beziehungen wird jemand von seiner Umgebung beurteilt, ob er diese Normen beachtet oder nicht.

3. Die Annulierung der Kollektivmoral: die kritische Moral

3.1. Kants Moralphilosophie

Der wichtigste Charakterzug der Kantischen Moraltheorie, dass sie den moralischen Aspekt des Handelns auf die innere Gesinnung eingrenzt, und bei der moralischen Beurteilung einer

Handelnden kommt es immer darauf, mit welcher Gesinnung er handelte, d.h., ob der Handelnde seine Handlung mit der Achtung der Pflicht ausübte. Wenn er neben der Achtung der Pflicht auch durch einer anderen Motivation getrieben wurde (etwa durch die Angst der rechtlichen Sanktion) dann kann diese Handlung nicht als moralisch beurteilt werden. Die Moral hat bei Kant nur einen formalen Rahmen für die Bestimmung der in den verschiedenen Situationen immer individuell aufspürbaren richtigen Handeln und das ist der kategorische Imperativ: „Der kategorische Imperativ ist also ein einziger, und zwar dieser: handle nur nach derjenigen Maxime, durch die zugleich wollen kannst, dass, sie ein allgemeines Gesetz werde.“ (Kant 1870:44).

Die Moral ist bei Kant ein Aspekt der individuellen Entscheidung mit solchem formellen Rahmen, dass ihre Anpassung an die Allgemeinheit (die Gemeinschaft) durch die Universalisierbarkeit der jeweiligen Handlungsmaxime garantiert wird. Neben der Universalisierbarkeit ist die zweite Charakterzug der moralischen Entscheidung, dass sie autonom ist, und das bedeutet, dass sie von den materiellen Motivationen frei ist und vollkommen durch die Achtung der moralischen Pflicht motiviert wurde. Die Autonomie bedeutet also in der Kantschen Moral die Freiheit von der kausalen Bestimmungskräfte der sinnlichen Natur der Menschen und der kategorische Imperativ fordert die Loslösung von der sinnlichen Natur und allgemeiner von der partikularen Bindungen der Gemeinschaften über die Menschen. Deshalb ist es kennzeichnend in den heutigen Moraltheorien Kantscher Prägung, dass sie von der gesellschaftlichen Bindungen und Zusammenhängen schon in dem Ausgangspunkt absehen und nicht forschen und diese werden von ihnen nur als besiegende Umstände angesehen.

Kant hatte alle gewohnheitsmäßige Verfolgung der Sitte in Gegensatz später zu Hegel abgelehnt, der den langsamen Aufbau der zweiten Natur (Sittlichkeit) durch Gewöhnung in dem Bewusstsein der Individuen betonte. Wenn es sich in dem Handeln um Gewohnheit (assuetudo) handelt, dann gibt es keine Freiheit und man kann hier nicht über Moral sprechen (Kant 1870:152). Die Moral wird also nicht nur innerhalb dem Bewusstsein plaziert, sondern sie kann immer nur mit bewussten Entscheidungen verwirklicht werden. Kant hat der Gedanken der Freiheit als Selbstgesetzgebung von Rousseau übernommen, aber während diese Idee durch Rousseau in Bezug auf die Selbstgesetzgebung des Volks als Staatslehre formuliert wurde, machte Kant sie zum Grund seiner Moraltheorie.

Sehen wir die Analyse von Ernst Tugendhat, von dem der Unterschied zwischen Hegel und Kant im Bereich der Moral sehr scharf herausgehoben wurde. Dieser Unterschied bedeutet einerseits das Absehen von den verbreiteten Moralnormen der Gemeinschaft des Handelnde bei Kant und die Betonung der bindenden Moralnormen der Gemeinschaft über die Handlungen der Gemeinschaftsmitglieder bei Hegel andererseits. Als Anhänger der Kantschen Linie der Moraltheorie sieht Tugendhat dann die Hegelsche Position als einen Übersetzungsirrtum, der von den zwei verschiedenen Meinungen des griechischen Wortes „ethos“ verursacht wurde: „Außerdem ist hier ein Übersetzungsirrtum unterlaufen. In der aristotelischen Ethik kommt nämlich nicht nur das Wort *éthos* (mit langem „é“), das Charaktereigenschaft bedeutete, sondern auch das Wort „ethos“ (mit kurzem „e“) vor, das Gewohnheit heißt, und es ist dieses zweite Wort, worauf die lateinische Übersetzung paßte. (...) Von daher ist die merkwürdige deutsche Übersetzung durch „Sitte“ zu verstehen, wie wir sie z. B. in Kants Buchtitel „Metaphysik der Sitten“ finden. Kant hat dabei überhaupt nicht an Sitten im gewöhnlichen Sinn (Brauchtum) gedacht, sondern verwendete das Wort einfach als Übersetzung für ‚mores‘, das seinerseits nicht mehr in seinem ursprünglichen Sinn verstanden wurde, sondern als angebliche Übersetzung eines griechischen Wortes. Erst Hegel macht sich dann den ursprünglichen Sinn des Wortes „Sitten“ zunutze, um gegenüber der Kantischen Moral eine angeblich höhere Form von Moral, genannt Sittlichkeit, zu konstruieren, die dadurch charakterisiert sein sollte, daß sie im Brauchtum und im

Hergebrachten fundiert sei" (Tugendhat 1994:34-35). Also können die gemeinschaftlich verbreiteten Normen der Richtigkeit nicht als Moral akzeptiert werden.

3.2. Der Weg der Verengung der Moral auf Gerechtigkeit

Die Unterdrückung der verbreiteten Sittlichkeit der Gemeinschaften und ihre Ersetzung durch die Konstruktionen der Moralphilosophen als kritische Moral bedeutete nur ein kleineres Problem, solange von diesen Konstruktionen die Kombinationen der Vielheit der seit Generationen weitergegebenen Moralwerten und -Tugenden variiert wurden. Das Problem wurde in den letzteren Jahrzehnten aber grösser, als von den einflussreichen Moraltheorien die Moral auf einen einzigen Moralwert verengt wurden, und damit wurden die Forderungen der Moral verhängnissvoll verzerrt. Anstatt von dieser 'kritischen Moral' die Reflexivität und die Qualität der moralischen Urteile verbessert würde, übt sie eine Tyrannei über die Menschen im Name der Ersatzmoral der Moralphilosophen aus. Schauen wir uns näher diesem Problem an.

Die Verengung der Moral auf Gerechtigkeit durch John Rawls erhielt die grösste Beachtung und die Verbreitung in der kulturellen und ideologischen Sektoren vom Beginn der 1970er Jahren aber als Reaktionen darauf entstanden andere Verengungen auch. So z.B. betonte Caroll Gilligan, dass durch die Gerechtigkeit-Moral von Rawls nur die Moral der Männer formuliert wurde und auf das weibliche Geschlecht sie nicht passt. Die Frauen haben eigene Moral, die als die Fürsorge-Moral formuliert werden kann (siehe Gilligan 1982). Michael Slote hat neulich eine ähnliche Version abgefasst, die er als „benevolence-Moral" genannt hat (siehe Slote 2001). Die letztgenannten Versionen erhielten aber keine breitere Beachtung in den öffentlichen politischen Diskussionen und in erster Linie werden sie von den Theoretiker der sozialen Arbeit beachtet. Als eine allgemeinere Kritik der Verengung der Moral überhaupt, kann in dem Werk von Beauchamp/Childress gefunden werden, aber sie konnten keine grössere Wirkung in den Kreisen der Moralphilosophen erreichen. Es lohnt sich aber zu zitieren, weil diese Gedanken als nächstliegend an unserer Analyse sind: „Inductionism (the bottom-up approach) maintains that we must use existing social agreements and practices as a starting point from which to generalize to norms such as principles and rules, and inductivists emphasize the role of particular and contextual judgments as a part of our evaluating moral life. A society's moral views are not justified by an ahistorical examination of the logic of moral discourse or by some theory of rationality (Kantien, Rawlsien, Koghlbergien) but rather by an embedded moral tradition and a set of procedures that permit new developments (...) the institution of morality cannot be separated from a cultural matrix of beliefs that has grown up and been tested over time (Beauchamp/Childress 1994:18) Also sehen wir die Verengung durch John Rawls und die Weiterentwicklung dieser Verengung durch Jürgen Habermas.

Diese Verengung wurde implizit schon von der Kantischen Ethik beinhaltet, von der für die Moralität der Handlungen die richtschnurmässige Universalisierbarkeit (der kategorische Imperativ) als einziges formales Gebot konzipiert wurde, wie es schon von Nicolai Hartmann betont wurde: „Denn Kant ist es, der an Stelle der inhaltlichen Mannigfaltigkeit der Tugenden die Einheit eines Sittengesetzes, an Stelle der materialen Erfülltheit das formale Prinzip, an Stelle des objektiven Wesens sittlicher Ideen die subjektive Gesetzgebung gesetzt hat" (Hartmann 1962:137). Doch wurden von meisten Moraltheoretikern, die der Kantschen Linie folgten, noch durch lange Zeit die Vielheit der moralischen Tugenden und Werten mehr oder weniger beibehalten. Der Anfang der Verengung kann zuerst im Buch „Eine Theorie der Gerechtigkeit" von John Rawls beobachtet werden. Er gründete die Entwicklungsphasen der

Moral auf die Analysen von Mc Dougall, James Mill, Freund, Piaget und besonders von Lawrence Kohlberg und als Ausgangspunkt nahm er die Drei-Phasen-Theorie von Kohlberg (vorkonventionelle, konventionelle und postkonventionelle Moral je mit zwei Unterstufen) in Anspruch. Kohlberg konzipierte diese drei Phasen für empirische Untersuchungen und obwohl er schon vor dem Beginn seiner empirischen Untersuchung ahnte, dass die meisten Menschen durch die zweite (konventionelle) Phase gekennzeichnet werden können, hat er angenommen, dass es eine dritte Phase gibt, wo die moralischen Entscheidungen aufgrund der universellen Moralprinzipien und Menschenrechten getroffen werden. Rawls hat diese letzte Phase einigermaßen umgeformt und die Moralprinzipien beschränkt auf die Grundsätze der Gerechtigkeit: „Überlegt man sich, wie diese grundsatzorientierte Moralität zustande kommen könnte (wobei unter Grundsätze erste Grundsätze verstanden werden, wie sie etwa im Urzustand erwogen werden), so sollte man nicht übersehen, dass die Gruppenmoralität ganz natürlich an die Kenntnis der Gerechtigkeitsgrundsätze heranführt“ (Rawls 1979:514).

Die empirischen Untersuchungen von Kohlberg waren wichtig für Rawls, weil er einen grossen Nachdruck darauf legte, dass eine Moraltheorie nur auf empirischer Basis als authentisch bezeichnet werden kann. Später wurde es aber aufgrund der Ergebnisse der empirischen Untersuchungen von Kohlberg klar, dass die dritte Phase nur als ideologisches Vorurteil angenommen wurde, und tatsächlich gibt es nur 7% der Menschen, die auf die untere Stufe der postkonventionellen Moral noch eingeordnet werden konnten, aber die letzte Stufe blieb leer. „Kohlberg eliminated Stage 6 from his scoring system for lack of finding empirical cases of Stage 6 thinking. Furthermore, there is little evidence for Stage 5 scoring in Kohlberg’s studies from around the world. Gibbs (1979) - a codeveloper of the scoring system - even proposed that true Piagetian stages of moral judgment stop with Stage 4. The lack of empirical data for Stage 5 and 6 - postconventional thinking - is a serious problem for Kohlberg’s enterprise, because he defined the stages from the perspective of the higher stages“ (Rest/Narvaez/Bebeau/Thoma:1999:22). Und Rawls identifizierte eben diese (nicht-existierende) letzte Stufe mit seiner grundsatzorientierten Moral. Trotz diesem Problem machte Rawls später keine Korrigierung und auch von den begeisterten Anhängern und Verbreitern seiner Theorie wurde meistens übergegangen.

Die Weiterentwicklung dieser Verengung kann in der Moraltheorie von Habermas beobachtet werden, der in seinen Werken in den 1980er Jahren eine niedere Sittlichkeit der lebensweltlichen Gemeinschaften und eine höhere universelle Moral als zwei Phasen der Moral ganz nach dem Vorbild von Rawlscher Theorie unterschied: „Aus der Perspektive eines Teilnehmers an moralischen Argumentationen stellt sich die auf Distanz gebrachte Lebenswelt, wo kulturelle Selbstverständlichkeiten moralischer, kognitiver und expressiver Herkunft miteinander verwoben sind, als Sphäre der Sittlichkeit dar. Dort sind die Pflichten derart mit konkreten Lebensgewohnheiten vernetzt, dass sie ihre Evidenz aus Hintergrundgewissheiten beziehen können. Fragen der Gerechtigkeit stellen sich dort nur innerhalb des Horizonts von *immer schon beantworteten Fragen* des guten Lebens. (...) Unter dem unnachsichtig moralisierenden Blick des Diskursteilnehmers hat diese Totalität ihre naturwüchsige Geltung eingebüsst, ist die normative Kraft des Faktischen erlahmt - können sich vertraute Institutionen in ebenso viele Fälle problematischer Gerechtigkeit verwandeln. Vor diesem Blick ist der überlieferte Bestand an Normen zerfallen, und zwar in das, was aus Prinzipien gerechtfertigt werden kann, und in das, was nur noch faktisch gilt. Die lebensweltliche Fusion von Gültigkeit und sozialer Geltung hat sich aufgelöst“ (Habermas 1983: 117-118.p.).

In der endgültigen Fassung seiner Moraltheorie (in dem Buch „Faktizität und Geltung“ in 1992) radikalisierte Habermas seine moraltheoretische Position und die lebensweltliche Sittlichkeit erschien hier nur als ein archaisches Substrat, die in der Moderne - auf

nachmetaphysischen Niveau - schon vollkommen und restlos auf die zwei Systeme von Recht und Moral aufgöst hat: „Ich gehe davon aus, dass sich auf nachmetaphysischen Begründungsniveau rechtliche und moralische Regeln gleichzeitig aus traditioneller Sittlichkeit ausdifferenzierte und zwei verschiedene aber einander ergänzende Sorten von Handlungsnormen nebeneinander treten“ (Habermas 1992:135). Noch eine Egränzung bedeutet für das Schicksal der aufgelösten Sittlichkeit die Behauptung, wenn Habermas „von den blossen Konventionen entwerteten Sitten“ schreibt (137.p.). Es gibt eine Veränderung auch im Bezug auf die Moral, die nunmehr allein (ohne Sittlichkeit) die Richtigkeitsvorstellungen bedeutet, weil „die postraditionelle Moral nur eine Form des kulturellen Wissens darstellt“ und das Recht ist es, was dieses kulturelles Wissen ins Handlungsnormen umwandeln kann. Also muss das frühere Zitat von Habermas kritisch so umgeformt werden, dass die Differenzierung der traditionellen Sittlichkeit nicht in zwei Sorten der Handlungsnormen geschah, sondern die sittliche Normen vollkommen verschunden und die neue Form der Moral nur im kulturellen Ebene existiert. „Diese zum Wissen sublimierte Moral“ oder „ins kulturelle System zurückgezogene Moral“ (145.p.) kann nur durch die Vermittlung des Rechts als wirkliche Handlungsnorm in der Gesellschaft wirksam werden. Die Loslösung dieser Vernunftmoral von den real existierenden Gemeinschaften wird von Habermas herausgehoben und als der Träger dieser Moral wurde „die Menschheit und die unterstellte Republik der Weltbürger“ gestellt (139.p.). Die Vermeidung der Ankerung der Weltmoral in den real existierenden Gemeinschaft hebt er auch auf die andere Seite heraus, wenn er „einzig über die internalisierende Verankerung moralischer Grundsätze im Persönlichkeitssystem“ erwähnt (146.p.). Als Endergebnis von Habermas' Moraltheorie bleibt kein System der Moralnormen in der Ebene der Gemeinschaften (nur die Rechtsnormen) aber auch in der Ebene der Republik der Weltbürger gibt es nur eine kulturelle Moral. Und diese kulturelle Moral wird auf die Gerechtigkeit geengt, wie schon früher gesehen werden konnte: „Die Vernunftmoral ist auf Fragen der Gerechtigkeit spezialisiert und betrachtet grundsätzlich alles im scharfen, aber engen Lichtkegel der Universalisierbarkeit“ (145.p.)

Die Verengung der Moral endete hier schliesslich also in der Vernichtung der Moral und in der alleinigen Herrschaft des Rechts, unterstützt mit der staatlichen Gewalt. Es ist nicht ohne Ironie, dass diese Theorie in der politischen und ideologischen Kämpfen umgemünzt als die Eroberung der grössten Weltfreiheit und Toleranz gilt. Denn sozialisieren die Millionen der Menschen - von Rawls und Habermas nichts wissen - weiterhin die tausendjährigen Moralwerte, -Normen und Tugenden und dann werden sie tagtäglich mit den Rechtsnormen konfrontiert, die weit von diesen Moralnormen stehen und aufgrund den kulturellen Moralwissen der Moralphilosophen formiert wurde.

Literaturverzeichnis

- Beauchamp, Tom L./James F. Childress (1994): Principles of Biomedical Ethics. (4. Edition) Oxford Univ. Press. 1994. 560.p.
- Gilligan, Caroll (1982): In Different Voice. Harvard Univ. Press.
- Habermas, Jürgen (1983): Moralbewusstsein und kommunikatives Handeln. Suhrkamp. Frankfurt am Main.
- Habermas, Jürgen (1991): Erläuterungen zur Diskursethik. Suhrkamp. Frankfurt am Main.
- Habermas, Jürgen (1992): Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats. Suhrkamp. Frankfurt am Main.
- Hartmann, Nicolai (1925)(1962): Ethik. (4. unveränderte Auflage.) Walter de Gruyter Verlag. Berlin.
- Hartmann, Nicolai (1932)(1962b): Das Problem des geistigen Seins. Untersuchungen zur

- Grundlegung der Geschichtsphilosophie und der Geisteswissenschaften. Walter de Gruyter Verlag, Berlin.
- Hegel, Georg Wilhelm Friedrich (1979): Grundlinien der Philosophie des Rechts. (Werke Band 7.) Suhrkamp. Frankfurt am Main.
- Jhering, Rudolph (1898): Der Zweck im Recht. Zweiter Band. Druck und Verlag von Breitkopf und Härtel. Leipzig. (Dritte durchgesehene Auflage).
- Kant, Immanuel (1870): Grunlegung zur Metaphysik der Sitten. (Herausgeben und erläutert J. H. von Kirkmann). Verlag von L. Heimann. Berlin. 1870.
- Kohlberg, Lawrence (1971): From is to ought. How to commit the naturalistic fallacy and get away with it in the study of moral development. In: Th. Mischel (ed.): Cognitive development and epistemology. Academic Press New York. 151-235.p.
- Lind, Georg (1983): Entwicklung des moralischen Urteilens - Leistungen und Problemen der Theorien von Piaget und Kohlberg. In: Lind/Hartmann/Wakenhut (hg.): Moralisches Urteilen und soziale Umwelt. Threoretische, methodologische Untersuchungen, Beltz Verlag Weinheim und Basel. 25-40.p.
- Rest, J./Narvaez, D./Bebeau M.J/Thoma S.J. (ed.) (1999): Postconventional Moral Thinking: A Neo-Kohlbergian Approach. Lawrence Erlbaum Associate Publishers. Mahwah.
- Slote, Michael (2001): Moral from Motives. Oxford Univ. Press.