Digitalisierung & KI

Ulrich Hemel

Vom Defizitmodell des Menschen zur digitalen Humanität

Was unterscheidet Menschen von Künstlicher Intelligenz?

Institut für Sozialstrategie
Laichingen – Jena – Berlin

Bleichwiese 3, 89150 Laichingen
www.institut-fuer-sozialstrategie.de
kontakt@institut-fuer-sozialstrategie.org

Berlin, Dezember 2021.

Abstract [en]:

Humanity is currently facing a huge transformation towards a digital-analogue, hybrid form of existence. While the difference between humans and animals used to be the subject of philosophical reflection, nowadays, the distinction between humans and artificial intelligence is the fundamental question. In comparison to the widespread deficit model of humans, according to which artificial intelligence is superior to human intelligence in every respect, the author proposes a new model that appreciates humans in their individual and cultural existence. Accordingly, humans can be characterized by special qualities and abilities that artificial intelligence does not have. These include corporeality, contradictoriness, human self-awareness and social iteration in the symbolic universe, as well as symbolic, cooperative and cultural capacity. The challenges posed by the digital transformation should be solved at the level of global civil society. In order to enable peaceful and sustainable coexistence in the digital world, digital humanity must be the standard in the future. The author understands the term digital humanity in two senses. On the one hand, as a regulatory idea with the aim of minimizing risks from AI applications and other technologies. On the other hand, it is the realization of humanistic ideals in the digital world that promotes the development of the best human abilities.

Keywords: *Artificial Intelligence, digital ethics, digital anthropology, digital humanity*

Abstract [de]:

Gegenwärtig steht die Menschheit vor der gewaltigen Transformation hin zu einer digital-analogen, hybriden Form der Existenz. Während früher der Unterschied zwischen Menschen und Tieren zum Gegenstand der philosophischen Reflektionen wurde, steht aktuell die Abgrenzungsfrage zwischen Menschen und Künstlicher Intelligenz im Vordergrund. Im Vergleich zu dem weit verbreiteten Defizitmodell des Menschen, nach dem Künstliche Intelligenz der menschlichen in jeglicher Hinsicht überlegen sei, schlägt der Autor ein neues Modell vor, das Menschen in ihrer individuellen und kulturellen Existenz aufwertet. Dementsprechend zeichnen sich Menschen durch besondere Eigenschaften und Fähigkeiten aus, über die Künstliche Intelligenz nicht verfügt. Dazu zählen Leiblichkeit, Widersprüchlichkeit, menschliches Selbstbewusstsein und soziale Iteration im symbolischen Universum sowie Symbol-, Kooperations- und Kulturfähigkeit. Die Herausforderungen, die der digitale Wandel mit sich bringt, sollten auf der Ebene der globalen Zivilgesellschaft gelöst werden. Um ein friedliches und nachhaltiges Zusammenleben in der digitalen Welt zu ermöglichen, muss dabei künftig digitale Humanität als Maßstab gelten. Der Begriff digitale Humanität wird in dem Beitrag im doppelten Sinne verstanden. Zum einen als eine regulative Idee mit dem Ziel, Risiken aus KI-Anwendungen und aus anderen Technologien zu minimieren. Zum anderen als Verwirklichung humanistischer Ideale in der digitalen Welt, die die Entfaltung der besten menschlichen Fähigkeiten befördert.

Schlagworte: *Künstliche Intelligenz, digitale Ethik, digitale Anthropologie, digitale Humanität*

Anwendungen Künstlicher Intelligenz (KI) prägen längst unseren Alltag, häufig unerkannt. KI erleichtert den Alltag, ermöglicht aber auch Machtmissbrauch und Manipulation. Wir haben als Gesellschaft darüber hinaus längst gelernt, schwache KI von starker KI zu unterscheiden. Auch in der Politik ist die Regelbedürftigkeit von KI-Anwendungen angekommen. Beispielsweise hat die EU hat in der Zwischenzeit Risikoklassen von KI-Anwendungen formuliert (European Commission, 2019). Grundsätzlich ist die Erfahrung, dass neue technologische Entwicklungen auch Risiken und Gefährdungen nach sich ziehen, nicht neu.

Über technische und politische Fragen hinaus wird ein entscheidender Punkt häufig unterschätzt. Dabei geht es um die Auswirkungen einer neuen Technologie **auf das menschliche Selbstbild**. Die Frage einer selbstreflexiven Vorstellung des Menschen von sich selbst in Reaktion auf technische Entwicklungen ist bislang jedoch eher ein blinder Fleck der Philosophie. Aus diesem Grund soll hier die Frage nach der Abgrenzung zwischen Menschen und Künstlicher Intelligenz im Vordergrund stehen.

Bis in die jüngere Vergangenheit hinein gab es immer wieder Versuche zur Abgrenzung zwischen Menschen und Tieren (vgl. U. Lüke, G. Souvignier 2020). Neu ist nun eine zweite Abgrenzungslinie in Gestalt der Frage nach dem Unterschied zwischen Menschen und Künstlicher Intelligenz. Dies gilt vor allem mit Blick auf humanoide Roboter und Androide, wie sie schon heute in Filmen und Science Fiction thematisiert werden (vgl. J. Nida-Rümelin, N. Weidenfeld 2018), sowie für die Diskussion um Zukunftsideen wie der einer Superintelligenz, die der menschlichen in jeglicher Hinsicht überlegen wäre (vgl. N. Bostrom 2018).

Da Menschen zur Selbstreflexion fähig sind, denken sie gelegentlich über sich selbst nach. Wie weit dieses Nachdenken einer Generalisierung über Zeit und Raum oder gar einem übergreifenden anthropologischen Theorieanspruch standhält, ist innerhalb der aktuellen Wissenschaftskultur durchaus fraglich (E. Bohlken, C. Thies 2009).

Kritische Grenzziehungen und die Sorge vor nicht gerechtfertigter Generalisierbarkeit setzen den Wert menschlicher Selbstreflexion nicht außer Kraft. Noch mehr: Menschen brauchen für ihre Weltorientierung alltagspraktisch verfügbare Denk- und Verhaltensmuster, die in ihrer Gesamtheit eine Art Weltbild oder eine „**mentale Architektur**" darstellen und die den Blickwinkel auf die Welt bestimmen (vgl. U. Hemel 2019, 335-350).

Interessanterweise gilt dies auch dann, wenn solche Weltbilder weder hinterfragt noch diskutiert werden. Schon wissenschaftspragmatisch ließe sich im Fall von wirksamen, aber nicht explizit ausgewiesenen, generellen Annahmen über Menschen die Forderung nach „Transparenz" knüpfen. Denn nur wenn ich meine eigenen Voraussetzungen ausweise, kann ich sie zum Gegenstand rationaler Diskussionen machen.

Diese Aussage gilt auch für die digitale Welt. So ist bei der Forschung rund um das Autonome Fahren der Begriff des „Weltmodells" längst zum Alltag geworden (F. Lütteke 2014). Die digitale Erfassung der Umgebung eines Fahrzeugs muss ja so geleistet werden, dass das System eine Plakatwand von einer LKW-Plane und eine Skulptur im öffentlichen Raum von einem spielenden Kind unterscheiden kann. Für die Parametrisierung des entsprechenden KI-Systems hat sich dabei der Begriff des Weltmodells durchgesetzt. Dabei wird unterstellt, dass die technische Parametrisierung eines Weltmodells sich gleichwohl von den

komplexen Weltsichten von Menschen, also humanen „Weltmodellen", in Komplexität und Reichweite unterscheiden.

Menschen haben aber auch ihrerseits spezifische Weltmodelle beim Umgang mit technischen, autonomen oder teilautonomen Systemen. Ein Klassiker des Umgangs mit uns wichtigen Gerätschaften ist die Namensgebung, etwa wenn jemand sein Auto „Mister Spock" oder „Schnauferle" nennt. Eine solche Personalisierung ist auch im KI-Umfeld häufig, nicht nur über die Avatare, also Alias-Personen im digitalen Raum, sondern auch über die emotionale Erfahrung, die sich im Verhalten technischer Systeme spiegelt. So höre ich oft, „mein Computer denkt noch" oder „er will gerade nicht", so als ob es sich um Personen handelte.

In solchen Fällen handelt es sich um Pseudopersonen (vgl. U. Hemel 2020a, 386), weil Menschen im Grunde wissen, dass es sich nicht um ein menschliches Gegenüber handelt. Mit Blick auf die immer besseren kognitiven Leistungen von Computern zieht aber weit verbreitet ein ganz anderes „Weltmodell" in die Diskussion ein. Im Vergleich mit Rechnern und gut programmierten, selbst lernenden Anwendungen ziehen Menschen aus funktionaler Sicht den Kürzeren, etwa weil sie ermüden, die Lust verlieren, unkonzentriert sind und mit einer Reihe weiterer Mängel versehen sind. Ein Lager- oder Produktionsroboter hingegen braucht keinen Schlaf, sucht und findet keine Ablenkung und funktioniert stets klaglos.

Aus der Beobachtung einer Reihe von Ethik-Workshops mit Personen aus den Bereichen Technik, Maschinenbau, Informatik und Künstliche Intelligenz, insbesondere im Umfeld des CyberValley Tübingen/Stuttgart, dem größten europäischen KI-Forschungsverbund, scheint das intuitiv plausible mentale Modell des Menschen durch genau solche kognitiven Defizite geprägt zu sein. Zugespitzt ließe sich formulieren: **Menschen sind Computer, nur eben schlechtere**. Folgefragen zu einer solchen Auffassung sind die nach dem Verschwinden der Arbeit (C. B. Frey, M. Osborne 2017), nach Superintelligenzen (N. Bostrom 2018), nach der Beherrschung der Welt durch Supercomputer und dergleichen (vgl. J. Nida-Rümelin, N. Weidenfeld 2018).

Anthropologisch mag die Sichtweise des Menschen als Defizitmodell im Vergleich mit überlegenen digitalen Anwendungen problematisch sein. Wirkmächtig ist sie dennoch. Sie bestimmt beispielsweise die Perspektive des Buchs von Armin Grunwald mit dem bezeichnenden Titel „Der unterlegene Mensch" (A. Grunwald 2019). Zumindest aus dem genannten Blickwinkel heraus lohnt sich folglich die Auseinandersetzung mit der These des „Defizitmodells", zugespitzt in der Frage, wie sich nun Menschen von KI tatsächlich unterscheiden.

Genauer gesagt geht es um drei Folgefragen, die sich aus der Entwicklung Künstlicher Intelligenz ergeben:

- Welches Bild haben wir Menschen von uns selbst? Welches „Weltmodell" wenden wir an, wenn wir uns selbst beschreiben? Wie stimmig ist es und wo liegen Grenzen seiner Generalisierbarkeit?
- Ausgehend von einer humanen Selbstbeschreibung bleibt die klassische Frage bestehen, wie Menschen sich von Tieren unterscheiden. Diese Frage wird besonders dringend dann, wenn Vorrechte von Menschen gegenüber anderen Spezies mit dem

Vorwurf des „Speziesismus" versehen und folglich bestritten werden (vgl. B. Steinbock 1978, L. Caffo u.a. 2015)

- Neu ist die Frage nach der Unterscheidung des Menschen von noch weiter entwickelten Formen Künstlicher Intelligenz als einem interaktionsfähigen Gegenüber von Menschen (vgl. U. Hemel 2020a, 356-365). Aus dieser Konstellation entstehen weitere, hier nicht behandelte Folgefragen rund um Themen wie Maschinenwürde, und zwar in Analogie zur Menschenwürde und (neuerdings) zur Tierwürde (vgl. J. S. Ach 2018, U. Hemel 2020a, 339-345).

Keine dieser Fragen ist trivial. Ihre Beantwortung hat massive Auswirkungen auf die Zuschreibung ethischen Verhaltens von Menschen und womöglich von anderen intelligenten Akteuren. Sie hat ethische, soziale, pädagogische und politische Folgen. Vor dem Versuch zu einer anthropologischen Antwort ist es daher sinnvoll, diese Folgewirkungen von anthropologischen Weltmodellen und der mit ihnen verbundenen Frage einer geeigneten mentalen Architektur noch einmal näher zu beleuchten.

1. Die massiven sozialen und politischen Folgen anthropologischer Weltmodelle

Durch mentale Rückkopplungsschleifen sind Selbstbilder von Menschen im individuellen ebenso wie im kollektiven Leben von handlungsleitender Bedeutung.

Ein aktuelles Beispiel ist die Diskussion um genderfluide Identitäten, Trans-Identitäten und nicht-binäre Formen von Sexualität (J. Sunderland, L. Litoselliti 2002, M. P. Galupo u.a. 2017). Wenn sich beispielsweise eine Person einem anderen Geschlecht zugehörig gefühlt, als es ihr Körper und ihre Geburtsurkunde nahelegen, so gilt heute ein sich formierender Konsens in der Art und Weise, dass das eigene Selbstbild Vorrang vor der biologischen Konstitution hat. Handlungsleitend wird ein Transgender-Selbstbild dann, wenn Auftreten und Kleidungsstil verändert werden oder sogar entsprechende medizinische Eingriffe folgen. Sozial wirksam ist ein Transgender-Selbstbild heute deshalb, weil die Gesellschaft zunehmend bereit ist, auf die Unterdrückung abweichenden Verhaltens mit Blick auf eine unterstellte zweigeschlechtliche Norm zu verzichten.

Schon dieses scheinbar individuelle Beispiel zeigt die Verwobenheit persönlicher Selbstbilder mit dem sozialen Kontext auf. Wenn wir die fraglich gewordene Überlegenheit des Menschen über andere Spezies auf die Spitze treiben und eine Art „Gleichberechtigung" aller Spezies fordern würden, dann ergeben sich ebenfalls massive soziale und politische Folgen. „Tierwohl" (vgl. J. S. Ach 2018, D. Wawrzyniak 2019) und „Menschenwohl" können dann im Extremfall auf eine Stufe gestellt werden. Das Essen von Tieren (vgl. Hirschfelder, G., Lahoda, K. 2012) wird dann moralisch bewertet, nicht nur aufgrund der Folgen für den Ausstoß oder „Verbrauch" von CO^2, sondern auch aus der Ableitung weitgehender ethischer Gleichwertigkeitsansprüche zwischen Menschen und Tieren, zumindest Wirbeltieren.

Tiere werden in einem Äquivalenzmodell für Wert und Würde von Menschen und Tieren tendenziell personalisiert. Es entstehen folglich nicht nur hochpersönliche Bindungen an Haustiere, sondern auch Bestattungsriten, Verhaltensregeln für die Zufügung von Schmerz und vieles mehr. Erörtert wird dann allenfalls die Grenzlinie zwischen Wirbeltieren und anderen Tieren wie etwa Spinnen, Bienen und Regenwürmern.

So wie die menschlichen Selbstbilder bei der freien Wahl des eigenen Geschlechts und bei der äquivalenten Würde von Menschen und Tieren in persönliches Verhalten und in den sozialen Raum ausgreift, so wirksam ist die Sichtweise des Menschen in einem kognitiven Weltmodell der Selbstbeschreibung. Der Mensch wird dann als effizientes System der Informationsverarbeitung beschrieben und sieht sich auch selbst so. Das Verhältnis von Inputgrößen zu Outputgrößen beschreibt dann ein Spektrum menschlicher Leistungsfähigkeit, das auch Auswirkungen auf persönliche Selbstbilder hat. Selbst im wirtschaftlichen Bereich ist aber eine Einschränkung des Menschen auf seine funktionalen, zweckorientierten Handlungen nicht ausreichend (vgl. C. Dierksmeier u.a. 2015, U. Hemel 2015, U. Hemel 2020b).

Naheliegend ist beispielsweise die Bemessung des Wertes eines Menschen anhand seiner kognitiven Leistungsfähigkeit. Kognitive Höchstleister könnten in einem solchen „kognitivistischen" Modell gesellschaftlich höher bewertet werden als kognitive Schwachleister. Das hört sich unangenehm an, ist aber in mehrfacher Hinsicht Teil der sozialen Realität: Erstens beginnt und endet das menschliche Leben mit schwächeren kognitiven Leistungen, etwa weil Säuglinge noch nicht sprechen können und weil höchstaltrige Menschen über 100 Jahre zu einem sehr hohen Anteil demenziell verändert sind. Zu kognitiven Höchstleistungen sind sie dann nicht mehr fähig. Zweitens wirft die Diskussion über die Programmierung von selbstfahrenden Fahrzeugen in ihrer utilitaristischen Ausprägung sehr wohl die Frage nach einer letztlich kognitivistischen Bewertung menschlichen Lebens im Fall eines ethischen Dilemmas auf (vgl. C. Misselhorn 2018).

Soll etwa bei einem unvermeidbaren Unfall die deontologische These gelten „ein Leben ist ein Leben", oder ist die folgende utilitaristische Abwägung ethisch, sozial und politisch legitim: „Besser ist es, bei einem unvermeidbaren Unfall einen sehr leistungsfähigen Menschen als zwei oder drei alte und schwache Personen zu retten".

Wer sich an der Logik des kategorischen Imperativs ausrichtet, wird Menschen grundsätzlich als „Selbstzweck" betrachten und dazu neigen, weitergehende Nützlichkeitserwägungen abzuwehren. Wer ein kognitives Weltmodell menschlicher Selbstbeschreibung wählt, bei dem der Wert eines Menschen von seiner intellektuellen Leistungsfähigkeit abhängt, der dürfte eher dem utilitaristischen Nutzenkalkül wie beschrieben folgen (zur Diskussion: P. Mastronardi, D. Taubert 2004, C. Misselhorn 2018, U. Hemel 2020a, P. Kirchschläger 2021).

In der sozialen und politischen Welt, die durch Gesetze und Richtlinien gestaltet wird, sind utilitaristische Abwägungen übrigens bereits heute weitgehend bekannt und akzeptiert. Beispiele sind die Bergrettung und die medizinische Triage. So ist eine aufwändige Bergrettung unter Einsatz von teuren Hubschraubern und vielen Personen ökonomisch aufwändig. Es gibt daher Regeln für den Abbruch einer Rettungsaktion, die im Einzelfall sehr schmerzlich sind, die aber überwiegend auf soziale Akzeptanz stoßen.

Ähnliche Verhaltensregeln gibt es bei Engpässen der medizinischen Versorgung, die zu einer „Sortierung" oder „Zuteilung" von Behandlungsmöglichkeiten führt. Dabei gilt beispielsweise der Grundsatz „Beati possidentes", also „selig die Besitzenden" deshalb, weil niemand aus einem Behandlungsbett mehr herausgeworfen werden darf, wenn ein Patient oder eine Patientin kommt, die die Behandlung eigentlich noch nötiger hätte.

Nun ist es nicht neu, dass philosophische Fragen und Antworten sich auf das Feld ethischen und sozialen Handelns auswirken. Neu ist allerdings die Dimension der Folgewirkungen in einer Welt der sehr schnellen digitalen Transformation in einem Umfeld der Klimakrise.

Erstaunlich ist dabei der Umstand, dass die normativen Voraussetzungen und Implikationen von Weltmodellen menschlicher Selbstbeschreibung in der akademischen Diskussion im Vergleich mit anderen Themenwelten derzeit ein Schattendasein fristen. Der zugrundeliegende Konflikt zwischen dem Anspruch wertfreier, objektivierbarer Wissenschaft und zielgerichteten, aber auch normativ verantworteten Eingreifens in die Welt wird dabei überwiegend verdrängt. Dies ist fatal, weil ein solches Verdrängungs- und Vermeidungsverhalten nicht nur gegen das allgemein übliche Transparenzgebot zu den eigenen Voraussetzungen des Denkens, sondern auch zu einer „Wiederkehr des Verdrängten" führen kann.

Dies gilt speziell auch für den Bereich der Künstlichen Intelligenz. Maschinelles Lernen anhand von Trainingsdaten aus der realen Welt führt schließlich über kurz oder lang zu wirksamen Lerneffekten in digitalen Programmen (C. Misselhorn 2018).

Im Extremfall kann dies zu außerordentlich sexistischen und rassistischen Social Bots führen (G. Neff, P. Nagy 2016, A. Schlesinger u.a. 2018). Entsprechende Beispiele sind auch bereits bekannt geworden, so aus Südkorea. Dann stellt sich aber zwangsläufig die Frage nach Abschaltung solcher Bots oder nach einer verantwortlichen Programmkorrektur.

Mit der „Wiederkehr des Verdrängten" kann dann gemeint sein, dass sich ganz unmittelbar normative Fragen bei der Praxis der Programmierung stellen. Wer darf in das Programm eingreifen? Wie tief und wie stark darf der Eingriff sein? Woraus leitet sich die Legitimation zu ethisch motivierten Eingriffen überhaupt ab?

Speziell in der Ethik der digitalen Welt tauchen also normative Fragen in einer Weise auf, die nicht allein Programmierern und Informatikerinnen zuzumuten ist, sondern die zum Gegenstand philosophischer Diskussion werden sollten.

2. Die Angst vor Normativität in der Wissenschaftspraxis und der ethische Entscheidungszwang in der digitalen Praxis

Angst ist zunächst ein subjektiver Gemütszustand (vgl. F. Riemann 2011, B. Bandelow 2004). Ob das Vermeiden normativer Aussagen in der Wissenschaftspraxis diesem Gefühl der Angst zugeschrieben werden sollte, wäre eine eigene Diskussion wert. Eigene Beobachtungen mit Vertretern so unterschiedlicher Disziplinen wie Informatik, Psychologie, Biologie, Sozial- und Wirtschaftswissenschaften, ja teilweise sogar Philosophie, Erziehungswissenschaften und Theologie zeigen zumindest ein weit verbreitetes Unbehagen gegenüber normativen Aussagen, weil mit diesen eine vermeintliche Minderung wissenschaftlicher Validität und Reputation einhergeht (vgl. S. Blackburn 1992; A. Gibbard 1986). Im Einzelfall kann es dann sogar um das Gefühl gehen, die eigene wissenschaftliche Laufbahn durch „normative" Aussagen zu beschädigen.

Im Gegensatz dazu wissen die axiomatisch begründeten Disziplinen wie die Rechtswissenschaft, die Mathematik und die Theologie um die Unausweichlichkeit normativer Haltungen und normativer Aussagen (vgl. K. Engisch 1971). Deren Nachteil ist die Festlegung auf

eine Erkenntnisperspektive, die eingestandenermaßen auch anders ausfallen könnte und die eine Abwägung von Vor- und Nachteilen voraussetzt. Ein solcher Nachteil kann ideologisch genau dann werden, wenn die eigene Perspektive zu einer Blickverengung in der Art und Weise führt, dass wertvolle Gegenargumente übergangen, bagatellisiert oder ausgeschlossen werden.

Die Vorzüge normativer Aussagen in der Wissenschaft werden selten beleuchtet. Ein Hintergrund dafür dürfte es sein, dass die Ideale der Wissenschaftsfreiheit, der Reproduzierbarkeit und der Unvoreingenommenheit im Widerspruch zur definierten Perspektivität normativer Herangehensweisen zu stehen scheinen.

Ideologisch wird das Vermeiden normativer Aussagen dann, wenn sie in die Richtung einer akademischen Lebenslüge abgleitet. Es ist schließlich eine biologische Tatsache, dass beispielsweise das menschliche Sehvermögen eine bestimmte Perspektive voraussetzt, die mit der Position unserer Augen und mit der Richtung unseres Blicks verbunden ist. Menschen sind keine von ihren biologischen Voraussetzungen freien Beobachter und Gestalter des Universums. Wir sind hingegen zwangsläufig an unsere eigene Perspektive gebunden, als Spezies Mensch und als Individuum.

Eine solche **Rückbindung an Perspektivität** gilt auch in der Wissenschaft. Jede Forschungsfrage, jedes Forschungsvorhaben ist mit der Eigenschaft verbunden, auf bestimmten Voraussetzungen in der Fragehaltung und in der Fragerichtung verbunden zu sein. Dies gilt auch und sogar für experimentelle Naturwissenschaften, wie es in den Arbeiten von Imre Lakatos und Paul Feyerabend schon vor über 50 Jahren kritisch beleuchtet wurde (vgl. I. Lakatos 1976, I. Lakatos 1979, P. Feyerabend 1976). Für die Physik geht es hier um die Theoriegeladenheit von Beobachtungen und die Beobachtungsgeladenheit von Theorien. In eine ähnliche Richtung geht die noch heute lesenswerte Studie von Georges Devereux über „Angst und Methode in den Verhaltenswissenschaften" (G. Devereux 1976). Aus den verschiedensten Blickwinkeln heraus besteht Konsens über die Schlussfolgerung: Zur wissenschaftlichen Methode gehört mit Blick auf diese Erkenntnisse **nicht das Vermeiden einer eigenen Perspektive, sondern deren transparenter Ausweis.**

Widersprüchlich an der heutigen Wissenschaftspraxis wirkt dabei nicht zuletzt die massive Wiederkehr normativer Forderungen in den Bereichen des Sprachgebrauchs und der ethischen Bewertung historischer und aktueller Verhaltensweisen. Mit dem Stichwort der „Responsible Science" (vgl. M. Solomon 2012, D. B. Resnik, K. C. Elliott 2016) wird weit darüber hinaus nach der Verantwortung von Wissenschaft angesichts der weltweiten Klimakrise und anderen, politischen Herausforderungen gefragt. Dabei gilt die Verwendung gendergerechter Sprache ebenso wie die bedingungslose Verurteilung von Rassismus und Kolonialismus als gesetzte, sanktionsbewerte Norm (vgl. auch E. W. Said 2009).

Damit kommen aber Normen und Normativität durch die Hintertür zurück in die Wissenschaftspraxis. Positionen zugunsten gendergerechter Sprache und für den Kampf gegen Rassismus und Kolonialismus sind zwar gut begründbar. Sie sollten aber als Positionen bezeichnet, transparent begründet und ausgewiesen werden. Andernfalls besteht trotz bester Absichten auch bei diesen normativen Zielen die Gefahr einer gewissen Ideologisierung, die Gegenpositionen, nachdenkliche Anfragen und kritische Formen der Auseinandersetzung mit ihnen mit dem Bannstrahl der kategorischen Ablehnung versieht.

Perspektivität und Transparenzgebot sind gerade auch im Umfeld von möglicher Diskriminierung und mit Blick auf Gendergerechtigkeit, aber auch latenten Rassismus durchaus Themen der Forschung und Praxis im Bereich Künstliche Intelligenz. Die verbreitete Forderung nach „Ethics by Design" (P. Verbeek 2008, K. Shilton 2013) oder „Value Sensitive Design" (B. Friedman, D. Hendry 2019) bedeutet ja ganz praktisch, dass Programmiererinnen und Programmierer sich von Anfang an über die ethischen und sozialen Folgen ihrer Entwicklungen Gedanken machen und darauf achten sollten, dass ihre Programmierung mit den geltenden sozialen Normen ihres Umfelds kompatibel wird.

Der Angst vor offener Normativität in der gängigen Wissenschaftspraxis steht folglich der faktische Zwang zu ethischen Entscheidungen in der Praxis digitaler Programmierung gegenüber. Ob diese Spannung als „Double Bind", als persönlicher Gewissenskonflikt oder als unausweichliche Begleiterscheinung der heutigen Welt gesehen wird, spielt dabei im ersten Schritt keine Rolle. In vielen Fällen zeigt sich aber, dass KI-Forscherinnen und KI-Forscher die angemessene Berücksichtigung ethischer und sozialer Fragen als eher unangenehme gesellschaftliche Anforderung in einem Kontext empfinden, bei dem sie im Grunde nur bestimmte technische Probleme lösen wollen. Am liebsten wäre ihnen dann eine ethische Musterlösung.

Genau das bieten Philosophinnen und Ethiker nicht. Die praktischen Auswirkungen eines „**philosophischen Schweigens**" zu den ethischen Implikationen der Entwicklung von KI-Programmen ist allerdings verheerend, weil mangelnde ethische Sprach- und Handlungsfähigkeit sich in eher unreflektierten Positionen spiegelt. Dringend erforderlich ist also eine öffentliche Diskussion über verfügbare Alternativen und ihre Konsequenzen.

Neu an der Situation ist die direkte Handlungsrelevanz philosophischer Entscheidungen. Denn die philosophische Unvermeidbarkeit einer eigenen Erkenntnis- und Handlungsperspektive spiegelt sich unmittelbar in der Praxis der KI-Entwicklung, bis hin zur Frage nach dem Selbstbild von Menschen in einer hochtechnologischen Umgebung.

Die Unausweichlichkeit normativer ethischer Positionen in der KI-Praxis geht in die gleiche Richtung wie die angesprochene Unvermeidlichkeit eines Weltmodells menschlicher Selbstbeschreibung in KI-Kontexten. Es lohnt sich daher, dem Transparenzgebot normativer Annahmen zu folgen und zunächst einmal auf das vorherrschende anthropologische Modell einzugehen: Die Sicht auf den Menschen als kognitive Defizitgestalt gut funktionierender KI.

3. Das Weltmodell menschlicher Selbstbeschreibung als defizitärer KI

Technischer Fortschritt führt und verführt zu Analogien im menschlichen Selbstbild. Ob Gott in der mechanisch geprägten Zeit nach der Aufklärung als oberster Uhrmacher oder als oberster Computer gedeutet wird, ist zeittypisch zu erklären. Nun spiegeln Gottesbilder immer auch Menschenbilder. Dies gilt auch in Zeiten geringerer öffentlicher Relevanz von Religion in philosophischen Debatten. Denn es lassen sich durchaus Strukturähnlichkeiten zwischen Menschen- und Gottesbildern finden. So wie Gott in der scholastischen Theologie als Inbegriff der Vernunft, sozusagen als absoluter Logos, gedeutet wurde, so spricht man heute von Superintelligenzen, die den Menschen im Grunde überflüssig machen.

Die Gottebenbildlichkeit aus Genesis 1, 26 wurde lange als menschliche Vernunftfähigkeit wahrgenommen (vgl. H. Schilling 1961). Nicht zuletzt Entwicklungen der KI zeigen nun auf, dass ein solches Menschenbild an Grenzen stößt. Wäre Vernunft wirklich die hervorstechende Eigenschaft des Menschen, dann verliert der Mensch im Kosmos genau dann seine privilegierte Bedeutung, wenn höhere Formen der Vernunft auftreten.

Die digitale Abbildung und Gestaltung der Welt als binäre Abfolge von Nullen und Einsen in jedem Rechenprogramm führt über verschiedene evolutionäre Stufen der Technikentwicklung nunmehr im 21. Jahrhundert zu einer neuen Situation: Menschen erfahren, dass Computer kognitive Aufgaben besser, ermüdungsfreier und weniger fehleranfällig erledigen können als sie selbst: sie erfahren sich insoweit als „unterlegen" (A. Grunwald 2019).

Dies lässt sich als Kränkung beschreiben, als Herausforderung an das eigene Selbstbild (vgl. U. Hemel 2020a, 123-166). Umgekehrt sind Kränkungen auch Lernchancen für umfassendes Identitätslernen nach dem Motto: „Wenn ich also gar nicht so bin, wie ich dachte, wer bin ich dann?"

Wenn man Lernprozesse in Lernen 1, Lernen 2 und Lernen 3 unterscheidet, sind Formen des Identitätslernens hoch spannend (vgl. U. Hemel 2017). Sie unterscheiden sich vom Lernen 1 (dem akkumulativen Lernen von Fakten) und Lernen 2 (als dem einsichtigen Verstehen von Zusammenhängen), weil sie als Lernen 3 und als „Identitätslernen" eine wirksame Selbstbeschreibung der eigenen Person nach sich ziehen, die im Sinn einer Rückkopplungsschleife auf Alltagshandlungen Einfluss nimmt.

Solche Rückkopplungsprozesse sind aus der Psychologie und Verhaltenstherapie wohl bekannt (vgl. J. Bauer 2015). Filme und Romane zeigen immer wieder, wie aus der Kindheit erlernte negative Selbstbilder ein Leben prägen und negativ beeinflussen können. Im Hintergrund steht dann aber grundsätzlich das Lernen 3 als einem Identitätslernen, welches auch problematische Züge annehmen kann.

Die Konstruktion der sozialen Wirklichkeit auf einer kollektiven Ebene folgt ähnlichen Zusammenhängen. Wer den Menschen überwiegend von seiner Vernunftfähigkeit und seiner kognitiven Leistungsfähigkeit her definiert, der kommt angesichts der Leistungsfähigkeit von KI in das Fahrwasser schwerer psychologischer Herausforderungen. Gerade aus solchen Kontexten kommen ja Überlegungen zur Abschaffung der Menschheit, zum Überflüssig-Werden menschlicher Arbeitskraft und zur Kritik an der echten oder vermeintlichen Überlegenheit des Menschen über andere Spezies. Wer den Menschen überwiegend von seiner Vernunft her definiert, der oder die gerät leicht in die Versuchung, im Menschen allenfalls eine defizitäre KI mit hoher Fehleranfälligkeit zu sehen.

Zu philosophischen und – auf anderer Ebene – theologischen Reflexionen gehört aber immer wieder die Überprüfung eigener Denkprämissen. Menschen geraten ja schon biologisch in vielen Disziplinen ins Hintertreffen mit anderen Lebewesen: Sie laufen weniger schnell als die besten Raubkatzen, sie haben weniger Kraft als Elefanten, eine geringe Ausdauer als Kamele, sie sehen und hören schlechter als bestimmte Lebewesen wie etwa Adler, Turmfalken und Fledermäuse.

Aufgrund ihrer besonderen Fähigkeit zur Zusammenarbeit und kulturellen Evolution haben Menschen aber in ihrer Entwicklung immer wieder neue Technologien hervorgebracht und

sich an diese angepasst. Menschen leben daher, ob es ihnen bewusst ist oder nicht, von jeher in ko-evolutiven Welten, die sie gestalten und von denen sie selbst geprägt sind.

So ermöglicht Sprache Kommunikation über konkrete Situationen hinaus. So entlastet die Schrift das Gedächtnis. Und so bewegen sich Autos und Flugzeuge deutlich schneller, als es Menschen physiologisch möglich ist. Auf der Linie dieser Argumentation könnte man sagen: „Wenn wir wissen, dass Autos schneller fahren als Menschen laufen können, dann können wir uns auch daran gewöhnen, dass digitale Applikationen besser rechnen und bessere kognitive Leistungen erbringen können als wir Menschen." Noch weiter: Wenn wir Ko-Evolution als Möglichkeit menschlicher Entwicklung betrachten, liegt auch der Gedanke einer Ko-Evolution für das Verhältnis Mensch und Maschine nahe.

Über das Argument der gewöhnungsbedürftigen „Alltagswelt der Ko-Evolution" hinaus greift das Zehn-Kampf-Argument. So wie der Gewinner der olympischen Goldmedaille im Zehnkampf in keiner einzigen der zehn Disziplinen Weltmeister ist, so sehr beeindruckt er durch die Kombination eines hohen Leistungsniveaus in allen zehn Disziplinen. Angewendet auf den Menschen könnte daraus folgen, dass das mangelnde Spitzenniveau bei Sinnesleistungen wie dem Sehen und Hören ebenso wie unsere Mängel in der kognitiven Verarbeitung von Informationen aus unserer Umwelt kompensiert werden durch das Gesamtbild des „Zehnkampfs", also durch die komplexe Fähigkeit, verschiedene Bereiche günstig, zielführend und überlebenswirksam zu organisieren. Die kognitive Überlegenheit einer guten KI könnte dann als punktuelle Leistung wahrgenommen werden, ohne dass die ko-evolutionäre Komplexität menschlicher Lebensformen in Frage gestellt würde.

Schon aus diesen beiden Argumenten heraus zeigt sich, dass das Bild von Menschen als einer defizitären KI allenfalls einen kleinen Ausschnitt der Wirklichkeit beleuchtet (vgl. M. Mitchell 2019).

Im Rückgriff auf den wirkmächtigen Gedanken der menschlichen Gottebenbildlichkeit (Schilling, H., 1961). lässt sich die Argumentation allerdings noch weiterführen. Schließlich legt das Christentum hohen Wert darauf, dass Gott die Liebe sei. Wenn dies so ist, dann kann Gott nicht nur auf den Gedanken der höchsten Vernunft reduziert werden. Gottebenbildlichkeit müsste sich in der wechselseitigen Durchdringung und Verwiesenheit von Vernunft- und Liebesfähigkeit zeigen.

Philosophisch und nicht theologisch argumentierend geraten wir hier in das Feld des Zusammenhangs rationaler und emotionaler Fähigkeiten des Menschen. Wir könnten also aufatmen und behaupten: „Computer können gut rechnen, aber nicht lieben".

Eine solche, vereinfachende Behauptung reicht aber schon technisch nicht aus. Digitale Anwendungen können sehr wohl auf emotionale Interaktionen hin trainiert werden. Genauer gesagt geht es um maschinelle Interaktionen, die von Menschen als Emotionen wahrgenommen oder decodiert werden. Der legendäre Turing-Test (vgl. J.H. Moor 1976) legt als Kriterium fest, ob Menschen die Interaktionen digitaler Interaktionspartner von menschlichen Interaktionen unterscheiden können (vgl. A. Saygin u.a. 2000). Es gibt nun bereits heute zahlreiche Mensch-Maschine-Interaktionen, die tatsächlich als Mensch-Mensch-Interaktionen wahrgenommen werden. Maschinen gewinnen folglich den Status von Quasi-Personen (U. Hemel 2020, 356-365, T. Störzinger u.a. 2020). Dabei spielt die

technische Herstellung der Anmutung von Emotionalität im Erleben nicht mehr die entscheidende Rolle.

Der Verweis auf die menschliche Alltagswelt der Ko-Evolution und die menschliche „Zehnkampf-Fähigkeit" mögen folglich als Indikatoren dafür gelten, dass Menschen sich nicht selbst als Rumpfform einer kognitiv überlegenen KI deuten sollten.

Sicherlich ist die besondere emotional-affektive Verfassung menschlicher Personen mit ihrer speziellen „Liebesfähigkeit" weit entfernt von der Konstitution performanter digitaler Systeme. Es könnte aber durchaus strittig sein, ob sich aus diesen heute noch faktischen Unterschieden tatsächlich eine schlüssige Argumentationskette für einen prinzipiellen Unterschied zwischen Menschen und digitalen Maschinen herleiten lassen. Anders gesagt: Wir müssen uns der Frage nach den Unterschieden zwischen Menschen und Computern neu stellen (vgl. dazu auch EKD 2021).

Ich werde mich der Frage in dreifacher Richtung zuwenden: von der Frage der Leiblichkeit und Widersprüchlichkeit her, von der Besonderheit menschlichen Selbstbewusstseins und von der speziellen Symbol-, Kooperations- und Kulturfähigkeit von Menschen.

4. Leiblichkeit, Widersprüchlichkeit und Ambivalenzkompetenz

Die einfachste Unterscheidung zwischen Menschen und KI ist die **Ebene der Stofflichkeit und der Leiblichkeit** (vgl. U. H. Körtner 2010, T. Fuchs 2020). Dieser Aspekt ist so offensichtlich, dass er entweder übersehen wird oder als trivial gilt. So wird der Mensch in KI-Kontexten gelegentlich als „Thinking Body" beschrieben. Schon im Begriff „Thinking Bodies" wirkt aber nach, was im Modell menschlicher Selbstbeschreibung als defizitärer KI implizit unterstellt wird, nämlich die weitgehende Reduktion von Menschen auf ihre kognitive Leistung. Der Körper ist dann sozusagen der stoffliche Träger einer geistigen Performance.

Menschliche Körper sind zur Erbringung kognitiver Leistungen auf bestimmte Umweltbedingungen oder „Systemparameter" ausgelegt wie beispielsweise Raumtemperatur, Flüssigkeits- und Nahrungszufuhr oder die Abwesenheit emotionaler Extremzustände. Digitale Interaktionspartner in der Gestalt von digitalen Programmen oder auch verkörpert in Computern oder gar Androiden benötigen aber auch ihrerseits einen initialen Input wie Mikrochips, Silizium, bestimmte Rohstoffe, Verfahren und technische Prozesse, um ins Stadium der Performanz zu treten. Dieser wiederum benötigt eine bestimmte Form der Energiezufuhr wie Strom aus dem öffentlichen Netz oder aus Akkus oder Batterien. Auch bei digitalen Interaktionspartnern und bei der KI ist die Abwesenheit extremer Störparameter wesentlich. Dazu können je nach Design mechanische Einwirkungen wie Stürze oder Hammerschläge, Flüssigkeiten wie Kaffee auf einer Tastatur oder auch leere Akkus und Stromunterbrechungen zählen. Die „Materialität" oder „Verkörperung" und Stofflichkeit ist folglich eine Eigenschaft von KI, die stets gilt, häufig aber übersehen wird (vgl. R. Pfeifer, F. Iida 2004).

Der gemeinsame Horizont kognitiver Aufgabenerfüllung ist dann KI und Menschen gemeinsam. Die stoffliche Trägerschaft durch Hardware und Software im Unterschied zu Fleisch und Blut unterscheidet Menschen und KI.

Eine solche menschliche Selbstbeschreibung löst aber mit Recht Störgefühle aus. Die Artikulation von Unbehagen erklärt aber nicht, wie sich Menschen und KI tatsächlich unterscheiden. Nun ist der Umgang mit kognitiv sehr leistungsfähiger Technik wie der KI insgesamt als **großer gesellschaftlicher Lernprozess** zu begreifen. In dessen Folge werden Hypothesen gebildet, verworfen oder verfeinert. In die gleiche Richtung gehen die folgenden Überlegungen, die sich auf Fragen der nicht-funktionalen Parametrisierung, der Widersprüchlichkeit, des Selbstbewusstseins und der sozialen Iteration beziehen.

Mit dem Begriff der Parametrisierung ist hier die Voreinstellung relevanter Parameter gemeint. Diese ist teilweise fest codiert, teilweise wird sie in einer Anwendungssituation festgelegt. Wenn jemand mithilfe von KI ein selbstfahrendes Auto entwickelt, dann braucht er bekanntermaßen ein Modell der relevanten Umweltparameter wie etwa Straßenbeschaffenheit, Wetterbedingungen, mechanische Hindernisse und potenzielle Störungen freier Fahrt (vgl. R. Gutsche 1994, E. D. Dickmanns 2005, C. Engemann, F. Sprenger 2015, D. Burkard u.a. 2019). Die Funktionalität technischer Prozesse führt jedenfalls bei allen digitalen Anwendungen zu einem funktionalen Design. Erkennbar unwichtige Dinge oder Parameter werden schon deshalb nicht einprogrammiert, weil dadurch der technische und finanzielle Aufwand ohne erkennbaren Zweck steigen würde.

Das ist bei Menschen grundsätzlich anders. Menschen zeichnen sich durch eine „Parametrisierung" in reichhaltigen, funktionalen und nicht-funktionalen Kontexten aus. Menschen sind von vornherein gerade nicht auf einen bestimmten Zweck ausgelegt. Sie sind von daher auch nicht vollkommen planbar und berechenbar. Sie nehmen Umweltparameter auf, die nichts mit dem sachlichen Zweck einer Interaktionskette zu tun haben, etwa einen bestimmten Geruch im Gebäude oder eine bestimmte Intensität der Beleuchtung oder einen Fleck auf der Hose des Gesprächspartners. Anders gesagt: Menschen nehmen Geräusche und Gerüche wahr, die sie mit ihrer biographischen Erinnerung verbinden und in ihre persönliche Lebensgeschichte integrieren können, so etwa den Fliedergeruch im Frühling beim Kennenlernen einer geliebten Person.

Eine solche komplexe Speicherung als identitätsbildende Erinnerung ist digitalen Anwendungen fremd. Sie hätte nur dann einen Sinn, wenn sie einem funktionalen Zweck zugeordnet wäre. Selbst wenn es gelänge, solche Formen von „Erinnerungsspeichern" nachzubauen oder zu emulieren, gäbe es einen Unterschied. Denn Menschen nutzen ihre „reichhaltige Parametrisierung" eben polyvalent, assoziativ und nicht ausschließlich funktional und zweckrational.

Hinter diesem Gedanken erscheint erneut das Zehn-Kampf-Argument im Vergleich von Menschen und KI. Doch selbst der Zehnkampf ist ja in sich eine zweckrationale Anordnung sportlicher oder sonstiger Leistungen. Menschen scheinen aber ihre eigene Rationalität in gewisser Weise dosieren zu können. Weder folgen sie stets festgelegten Zwecken noch sind sie immer rational. Das aber ist eine Stärke, nicht eine Schwäche von Menschen. Auf die daraus und aus anderen Faktoren folgende, mögliche potenzielle Minderleistung des Menschen im Vergleich zu den kognitiven Möglichkeiten einer guten, lernenden KI zu verweisen, ist daher nur ein Blick wie durch ein Schlüsselloch. Dieser Blick ist zwar fokussiert, er wird aber wesentliche Aspekte des Gesamtbildes nicht erfassen können.

Dabei kann der Reichtum menschlicher Ausdrucksformen durchaus auch reduktionistisch gedeutet werden. Wenn Menschen in technischer Sprache nicht-funktional parametrisiert

sind, dann sind alle nicht-funktionalen Lebensäußerungen von Menschen im Grunde ein Hintergrundgeräusch, eine Störung oder eine Leistungsminderung.

Eine solche Aussage ist aber nur dann richtig und sinnvoll, wenn ästhetische, emotionale und auch im engeren Sinn ethische Aspekte des Menschseins ausgeblendet werden. Dies gilt auch mit Blick auf die im Alltag erfahrbare Widersprüchlichkeit von Menschen, die heute dieses und morgen jenes sagen, die heute Position für und morgen gegen eine bestimmte Maßnahme wie etwa eine Geschwindigkeitsbegrenzung auf Autobahnen zum Ausdruck bringen. Widersprüchlichkeit kann nämlich mit und ohne temporalen Bezug gedacht werden.

Daher ist die logische Widerspruchslosigkeit und Widersprüchlichkeit von der typisch menschlichen Widersprüchlichkeit auf der Zeitachse zu unterscheiden. Denn im menschlichen Leben hängen Entscheidungen eben auch von der praktischen Interpretation konkreter Lebenssituationen ab. Und da kommt es vor, dass ein bestimmter, üblicherweise auch passender Handlungsmaßstab für konkrete Situationen ungeeignet ist. Ein klassisches Beispiel ist das Überfahren einer roten Ampel, wenn jemand seine hochschwangere Frau mit schon geplatzter Fruchtblase ins Krankenhaus bringt.

In der philosophischen Ethik spricht man hier von „Epikie" (vgl. K. Demmer 2014), also einem ethisch gerechtfertigten Übertreten von Normen und Handlungsrichtlinien zur Verfolgung eines höheren Gutes. Das Sensorium für Epikie ist allerdings digital schwer abzubilden. Der Versuch dazu würde jedenfalls die nötigen Programmierzeilen oder lines of code multiplizieren und die Fehleranfälligkeit einer digitalen Anwendung enorm steigern.

Menschliche Widersprüchlichkeit kann als **situative Flexibilität** und damit als evolutionäre Errungenschaft im Sinn einer individuellen und kollektiven Lernfähigkeit gedeutet werden, die sogar im „laufenden Betrieb" eine Art Umprogrammierung ermöglicht. Diese menschliche Eigenschaft ist folglich ihrerseits Teil der evolutionären Anpassung des Menschen an seine Umgebung. Sie kann im Extremfall auch als Katastrophenkompetenz gedeutet werden. Menschen sind tatsächlich gut darin, in Katastrophen sich irgendwie durchzuwursteln und nach einem Ausweg zu suchen, den sie dann manchmal finden, manchmal eben auch nicht. Zur damit verbundenen Ambivalenzkompetenz von Menschen gehört allerdings auch die Einsicht, dass Menschen häufig auch gut darin sind, Katastrophen anzurichten. Im Managementleben kommt es dann manchmal zur humorvollen Aussage: „Manager oder Managerinnen sind Menschen, die Probleme lösen, welche sie zuvor selbst verursacht haben."

Widersprüchlichkeit, Ambivalenzkompetenz und höchste situative Flexibilität sind jedenfalls bis heute eine hohe Hürde für jedwede KI. Dies gilt ganz pragmatisch, aber auch funktional, weil die entsprechende Programmierung jede Anwendung teuer macht. Die Aussage gilt letztlich aber auch grundsätzlich, eben weil Widersprüchlichkeit kein logisches Programmierziel und auch kein technisch abbildbares Ziel maschinellen Lernens darstellt.

5. Selbstbewusstsein und soziale Iteration im symbolischen Universum

Aus dem Gesagten folgt nicht nur ein weiterhin bedeutender Unterschied bei der Input-Output-Gestaltung der Relationen zur äußeren Welt, wenn es um Menschen im Vergleich

mit digitalen Maschinen geht. Die Interpretation möglicher Widersprüchlichkeit als situativer Flexibilität ist ja Teil der speziellen Anpassungsfähigkeit von Menschen an ihre Umwelt. Sie ist aber auch Ausdruck einer bestimmten Form reflexiver Distanzierung von unmittelbar gegebenen Weltzusammenhängen. Die Reflexionsfähigkeit des Menschen geht daher in einer kognitiven Rechenleistung nicht auf. Sie hat insbesondere die Eigenschaft einer impliziten und expliziten Mehrstufigkeit, weil wir Menschen uns im Wachzustand in einem komplexen Kontinuum von Denken und Erleben wahrnehmen, welches hoch persönliche, situative, aber auch kulturelle Komponenten gleichzeitig umfasst.

Menschliches Selbstbewusstsein ist daher nicht einfach die Wahrnehmung des eigenen Selbst als einer Handlungsmitte oder als Ausgangspunkt selbstgesetzter Handlungen. Es umfasst vielmehr auch eine relationale Situiertheit in persönlichen, gruppenbezogenen und kulturellen Kontexten (vgl. G.H. Mead 1978). Weiterhin ist menschliches Selbstbewusstsein gelegentlich ein vorherrschender, meistens aber ein begleitender, „konkomitanter" Zustand der Selbstwahrnehmung. Dieser Zustand hat seinerseits ganz unterschiedliche Facetten.

Ein wesentlicher Teil menschlicher Selbstwahrnehmung besteht in der situativen Vergewisserung der eigenen Identität. Damit einher geht eine Art von Konsistenz- und Plausibilitätsprüfung äußerer Situationen und des eigenen, inneren und äußeren Verhaltens in praktischen Alltagssituationen. Wer sich als Mann in einem Restaurant zufällig auf die Damentoilette verirrt oder wer sich umgekehrt als Frau vor einer Reihe von Urinalen wiederfindet, der stutzt und wird seinen Handlungsstrang korrigieren.

Identität und Rolle fließen im sozialen Leben allerdings eng ineinander. Ein Bankmanager vor 30 Jahren war in seiner Rolle ohne Krawatte nicht vorstellbar, heute hingegen schon. Fluide Identität in Rolle und Selbstwahrnehmung ist folglich Teil jeder menschlichen Lebensgeschichte.

Interessant an diesen Alltagsbeobachtungen ist nun gerade der Vergleich mit digitalen Erzeugnissen. Diese wissen nicht um sich selbst. Sie haben auch dann kein Selbstbewusstsein menschlicher Art, wenn sie einen Arbeitsspeicher haben, der Erinnerungen emuliert oder wenn sie auf die Planung von funktionalen Handlungssträngen ausgelegt sind. Hintergrund dabei ist insbesondere ihre Funktionalität. Denn die Ausführung von Aktionen (die wir hier vereinfachend „Handlungen" nennen) folgt grundsätzlich einem auslösenden Reiz, der eine absehbare Reaktion auslöst. „Spontane" Handlungen von Computern, Robotern oder Androiden aber sind schon deshalb nicht vorgesehen, weil diese dann den Modus der Berechenbarkeit und Planbarkeit verlieren würden.

Die „Emanzipation" digitaler Interaktionspartner wird zwar in der Gestalt von Filmen, Romanen oder Diskussionsbeiträgen immer wieder thematisiert. Sie reicht aber allenfalls hinein in eine Art „Bedrohungsszenario", bei der die Menschheit geplant oder zufällig in eine Katastrophe gerät. Im Hintergrund steht, wie im klassischen James-Bond-Film, häufig eine machtgierige Person oder Organisation. Dies kann als Hinweis dafür gelten, dass die Besonderheit der menschlichen Willensbildung selbst in dystopischen und düsteren Phantasien nicht durch die „Handlungsautonomie" digitaler Apparaturen ersetzt werden kann.

Selbstbewusstsein und Selbstbestimmung gehen daher Hand in Hand, weil sie die Fähigkeit zur selbstreflexiven Distanz mit der Fähigkeit zum Setzen handlungsleitender, autonomer

Willensakte kombinieren (vgl. dazu J. Bauer 2015, C. List 2021). Die häufig anzutreffende Reduktion der Diskussion auf den Aspekt des Selbstbewusstseins greift daher zu kurz.

Noch weiter: Wer den Unterschied zwischen KI und Menschen begreifen will, tut gut daran, den einzelnen Menschen nicht als isolierte Größe, sondern als Mensch unter Menschen zu verstehen. Unter diesem Blickwinkel kommt das soziale und kulturelle Universum zur Geltung, dem wir Menschen ausgesetzt sind und das wir unsererseits mitgestalten.

Denn die Kombination von Selbstbewusstsein und willentlicher Selbstbestimmung wird ja grundsätzlich nur in einem zeitgeschichtlich bestimmten, örtlich und kulturell geprägten Kontext wirksam. Menschen handeln nicht ohne Rücksicht auf ihre soziale Umgebung. Sie antizipieren die Reaktionen anderer und modulieren ihr eigenes Handeln entsprechend. Dies kann in der Form größerer sozialer Abhängigkeit ebenso wie in der Gestalt maximaler Rücksichtslosigkeit geschehen. Es ist aber, so oder so, eine Randbedingung menschlichen Handelns.

Dies wird insbesondere im symbolischen Interaktionismus, aber auch in allen gängigen Theorien der Persönlichkeitsbildung und Erziehung mit großer Selbstverständlichkeit artikuliert (vgl. G. H. Mead 1978, G.-B. v. Carlsburg, A. M. Stross 2021). Menschen wissen um andere Personen und nehmen deren Reaktionen in ihren Handlungen vorweg oder richten ihre Handlungen von vornherein an den erwarteten Reaktionen aus.

Diese Kette sozialer Interaktionen ließe sich grundsätzlich auch in einer KI-geprägten Landschaft abbilden. Sie wird aber teuer, denn auch selbstlernende digitale Systeme brauchen klug vorbedachte Trainingsdaten. Zumindest in der aktuellen Landschaft sind komplexe Vernetzungen mit unklar festgelegten Zielen und Routinen über verschiedene Systeme hinweg allenfalls Gegenstand von Arbeiten im Rahmen der sogenannten „Industrie 4.0". Hier gilt es beispielsweise, Wartungsdaten von Maschinen aus unterschiedlichen Sensor-Systemen miteinander zu vernetzen (C. Engemann F. Sprenger 2015, D. Burkart u.a. 2019). In jedem Fall aber geht es um eine konkrete Vernetzung mit strikt festgelegten funktionalen Zwecken. Von autonomen Willensakten im Kontext sozialer Iteration wie oben beschrieben sind solche Systeme noch meilenweit entfernt.

6. Mehrstufige Symbol- und Kooperationsfähigkeit und die Kulturfähigkeit von Menschen

Fassen wir die verschiedenen Ebenen gradueller und eher grundsätzlicher Unterschiede zwischen Menschen und Künstlicher Intelligenz zusammen, ergibt sich ein komplexes Bild. Nicht-funktionale Willensakte und Handlungen, Widersprüchlichkeit oder situative Flexibilität, Selbstbewusstsein und soziale Iteration in Handlungsketten sind Menschen von Kindheit an vertraut. Menschen lernen, ein Bewusstsein ihrer selbst in der Gestalt einer sich dynamisch entwickelnden Identität zu haben (vgl. U. Hemel 2017). In jedem Augenblick ihres Lebens liegen Kontinuität und Diskontinuität nahe beieinander. Menschen sind dabei sowohl völlig individuell wie auch sozial beides: einzigartig und Teile eines größeren Ganzen.

Die grundlegende und mehrstufige Symbol- und Kooperationsfähigkeit von Menschen spiegelt sich in ihrer Fähigkeit zur Bildung von Kulturen, so wie sie in der Anthropologie immer

wieder beschrieben und analysiert werden. Die Vielgestalt von Kulturen ist ihrerseits jeweils ein Spiegel bestimmter technologischer Entwicklungen wie z.B. nach dem Übergang vom Stadium des Jägers und Sammlers zum Stadium der bäuerlichen Landwirtschaft.

Gegenwärtig steht die Menschheit vor der gewaltigen Transformation hin zu einer digital-analogen, hybriden Form der Existenz. Daher stellen sich auch Abgrenzungsfragen zwischen Menschen und Künstlicher Intelligenz.

Menschliche Kulturen sind nicht nur als Sprachgemeinschaften zu begreifen. So hat sich heute über Sprach- und Landesgrenzen hinweg eine globale Zivilgesellschaft herausgebildet, die über digitale Information und Kommunikation, aber auch über den globalen Austausch von Gütern und Dienstleistungen miteinander verbunden ist und die angesichts der aktuellen Klimakrise auch vor gemeinsamen Herausforderungen steht.

Im Unterschied zu allen bisher bekannten Formen von KI führt die menschliche Symbol- und Kooperationsfähigkeit zu komplexen und nicht letztlich aufklärbaren Gebilden wie beispielsweise einem internationalen Finanzsystem, einer Europäischen Zentralbank oder gar zu Cyberwährungen wie Bitcoin und anderen. Letztere sind bereits analog-digitale Mischformen, deren Hintergrund aber eine so schwer greifbare Fähigkeit wie Systemvertrauen zu sein scheint.

Unabhängig von solchen speziellen Ausprägungen zeigt sich die Kulturfähigkeit von Menschen auch in einem philharmonischen Orchester, in der Verleihung von Nobelpreisen für Literatur oder in Sportereignissen wie den Olympischen Spielen oder den viel kritisierten Formel-1-Rennen. Diese Beispiele zeigen aber auch die Wandlungsfähigkeit und das Anpassungstalent von Menschen: Gerade kulturelle Ausdrucksformen ändern sich im Sinne einer gesellschaftlichen Evolution teilweise sehr schnell, so dass sich auch die Lebenswelt von Menschen innerhalb weniger Jahrzehnte teilweise drastisch verändert. Damit ändern sich zwangsläufig auch ethische Fragen, ethische Anforderungen und ethische Handlungsrichtlinien (vgl. u.a. M. Düwell u.a. 2006).

Die zukünftigen Möglichkeiten der KI gehen zwar viel weiter, als es die meisten Menschen heute überhaupt für möglich halten. Die besonderen Möglichkeiten der digitalen Welt führen nicht zuletzt zu einer Expansion dessen, was wir „symbolisches Universum" des Menschen nennen können. Anders gesagt: Unser Alltagsleben spielt sich immer stärker in digitalen Räumen ab. Unser Leben wird durch unsere digitale Identität und eine hybride, analog-digitale Lebensform geprägt.

Eine der kollektiven Aufgaben in der globalen Zivilgesellschaft ist es daher, eine geeignete Form digitaler Humanität zu entwickeln, die den Menschen weder unter- noch überschätzt (vgl. U. Hemel 2020a, 366-372). Digitale Humanität kann dabei zur regulativen Idee für eine solche Form digital-analoger Existenz von Individuen, Gesellschaften und der Menschheit insgesamt werden, die in der Lage ist, Risiken aus KI-Anwendungen und aus der digitalen Welt zu erkennen und zu beherrschen, sie zugleich aber im Sinn der Entfaltung der besten menschlichen Fähigkeiten zu nutzen.

Der Begriff der „**digitalen Humanität**" bezeichnet, so gesehen, die besten Möglichkeiten von Menschen in ihrer individuellen und sozialen Existenz, bezogen auf ein friedliches Zusammenleben in der globalen Zivilgesellschaft, auf eine nachhaltige Existenzform mit Blick

auf den Planeten Erde und mit Blick auf die Gestaltung einer lebensfreundlichen digitalen Welt (vgl. U. Hemel 2021).

7. Verantwortung, Freiheit und Schuld im individuellen und sozialen Kontext menschlichen Handelns

Denn die besondere Struktur menschlicher Handlungen im schwer zu entwirrenden Ineinander eines reflexiven Selbstbewusstseins, eines autonomen Wollens und einer komplexen Abhängigkeit von sozialen und kulturellen Randbedingungen führt ja nicht nur zu kulturprägenden regulativen Ideen wie denen von Freiheit und Verantwortung. Sie hat vielmehr auch individuelle und kollektive Zuschreibungsakte zur Folge, die bei der ethischen Bewertung von Handlungen ebenso wie im juristischen Kontext bei der Zuschreibung der Verantwortung für Handlungsfolgen eine große Rolle spielt (vgl. U. Hemel, 2020a).

Vereinfacht gesagt: Die Rückseite der Freiheit zeigt sich in Verantwortung. Verantwortung aber zieht die Möglichkeit nach sich, schuldig zu werden oder zumindest haftbar zu sein. Die Grenzen von KI liegen auch dort, wo diese zwar lern- und handlungsfähig sein mag, aber nicht selbst Träger von Verantwortung werden kann. Andernfalls müssten wir uns Geld- und Haftstrafen für KI vorstellen können, etwa als Handlungsfolge fahrlässiger oder gar im ethischen Sinn schuldhafter Verfehlungen.

Digitale Programme und Vorrichtungen können als Pseudopersonen und als Quasipersonen wirken und in bestimmten Bereichen dem Handeln von Menschen sehr nahe kommen oder von ihm gar nicht mehr offensichtlich unterschieden werden. Sie sind aber aufgrund ihres Mangels an Selbstbewusstsein, selbstbestimmter Programmautonomie und kultureller Koevolution eben gerade keine zu Freiheit und Schuld fähige Personen im menschlichen Sinn.

Digitale Programme mithilfe von KI erweitern folglich menschliche Handlungs- und Ausdrucksmöglichkeiten enorm. Die Beherrschung der mit ihnen einhergehenden Risiken ist aber unsere ureigene, menschliche Aufgabe, auf individueller und politischer Ebene. Tatsächlich ist es ja so, dass hinter militärischer, politischer und kommerzieller digitaler Macht konkrete Machtinteressen konkreter Personen, Organisationen und Staaten stehen, bis zum heutigen Tag.

Es gibt daher, so können wir schlussfolgern, keinerlei Anhaltspunkt dafür, dass eine heutige oder künftige KI den inneren und äußeren Reichtum an Ausdrucksmöglichkeiten von Menschen in ihrer individuellen und kulturellen Existenz auch nur annähernd erreicht. KI sollte daher zwar in ihren Risiken beherrscht, aber auch in ihren Chancen zur Alltagserleichterung gewürdigt werden. Am Ende gilt: Rechner rechnen, Menschen leben – mit ihren Handlungsmöglichkeiten und Träumen, aber auch mit ihren Verfehlungen und Grenzen (vgl. U. Hemel 2020a, 368).

LITERATUR

Ach, J.S. (2018). Tierwohl und Ethik. In J.S. Ach & D. Bochers (Hrsg.), Handbuch Tierethik. Grundlagen – Kontexte – Perspektiven. Stuttgart: J.B. Metzler.

Bandelow, B. (2004). Das Angstbuch. Woher Ängste kommen und wie man sie bekämpfen kann. Reinbek: Rowohlt 2004.

Bauer, J. (2015). Selbststeuerung (7. Aufl.). München: Blessing.

Blackburn, S. (1992). Gibbard on Normative Logic, Philosophy and Phenomenological Research, 52(4), pp. 947-952.

Bohlken, E., Thies, C. (Hrsg.) (2009). Handbuch Anthropologie. Der Mensch zwischen Natur, Kultur und Technik. Stuttgart: Springer.

Bostrom, N. (2018). Superintelligenz. Szenarien einer kommenden Revolution (3. Aufl.). Berlin: Suhrkamp.

Burkard, D., Kohler H., Kreuzkamp N., Schmid J. (Hrsg.) (2019). Smart Factory und Digitalisierung. Baden-Baden: Nomos.

Caffo, L., Horta, O., & Rude, M. (2015). Speziesismus. In A. Ferrari & K. Petrus (Hrsg.), Lexikon der Mensch-Tier-Beziehungen. Bielefeld: transcript. S. 318-323.

Carlsburg, B. v., Stroß, A.M. (Hrsg.) (2021). (Un)pädagogische Visionen für das 21. Jahrhundert, (Non-) Educational Visions for the 21st Century, Reihe: Baltische Studien der Erziehungs- und Sozialwissenschaft. Frankfurt u.a.: Peter Lang.

Demmer, K. (2014). Epikie. In K. Demmer (Hrsg.), Selbstaufklärung theologischer Ethik. Paderborn: Ferdinand Schöningh, S. 109-132.

Devereux, G. (1967). Angst und Methode in den Verhaltenswissenschaften. München: Hanser.

Dickmanns, E.D. (2005). Vision: Von Assistenz zum Autonomen Fahren. I M. Maurer & C. Stiller (Hrsg.), Fahrerassistenzsysteme mit maschineller Wahrnehmung. Berlin/Heidelberg: Springer. S. 203-237.

Dierksmeier, C., Hemel, U., Manemann, J. (Hrsg.) (2015). Wirtschaftsanthropologie. Baden-Baden: Nomos.

Düwell, M., Hübenthal, C., & Werner, M. H. (Hrsg.) (2006). Handbuch Ethik. Stuttgart: Springer.

EKD Denkschrift Freiheit Digital (2021). Die zehn Gebote in Zeiten des digitalen Wandels. Leipzig: Evangelische Verlagsanstalt.

Engemann, C., Sprenger, F. (Hrsg.) (2015). Internet der Dinge. Bielefeld: transcript.

Engisch, K. (1971). Einführung in das juristische Denken. Stuttgart u.a.: Kohlhammer (Erstausgabe 1956).

European Commission High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI," April 8, 2019, https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. Brussels, abgerufen am 4.11.2021 um 14.25h.

Feyerabend, P. (1976). Wider den Methodenzwang. Frankfurt a.M: Suhrkamp.

Frey, C. B., Osborne, M. (2017). The Future of Employment. In F.Y. Phillips (Eds.), Technological Forecasting and Social Change. Amsterdam, 114. pp. 254-280.

Friedman, B., Hendry D. G. (2013). Value Sensitive Design, Shaping Technology with Moral Imagination. Cambridge/Mass.: MIT Press.

Fuchs, T. (2020). Verteidigung des Menschen. Grundfragen einer verkörperten Anthropologie. Berlin: Suhrkamp.

Galupo, M.P., Pulice-Farrow, L., & Ramirez, J.L. (2017). "Like a Constantly Flowing River": Gender Identity Flexibility Among Nonbinary Transgender Individuals. In J.D. Sinnott (Ed.), Identity Flexibility During Adulthood. Cham: Springer. pp. 163-177.

Gibbard, A. (1986). An Expressivistic Theory of Normative Discourse. Ethics, 96(3), pp.472-485. Illinois: UCP.

Grunwald, A. (2019). Der unterlegene Mensch. Die Zukunft der Menschheit angesichts von Algorithmen, künstlicher Intelligenz und Robotern. München: Riva.

Gutsche, R. (1994). Das Transportsystem MONAMOVE. In Fahrerlose Transportsysteme, Fortschritte der Robotik, Vol 22. Braunschweig/Wiesbaden: Vieweg & Sohn Verlagsgesellschaft mbH.

Hemel, U. (2015). Wirtschaftsanthropologie. Grundlegung für eine Wissenschaft vom Menschen, der wirtschaftlich handelt. In C. Dierksmeier, U. Hemel & J. Manemann (Hrsg.), Wirtschaftsanthropologie. Baden-Baden: Nomos. S. 9-26.

Hemel, U. (2017). Heimat und personale Selbstbildung. Eine pädagogische Reflexion. In U. Hemel & J. Manemann (Hrsg.), Heimat finden – Heimat erfinden. Politisch-philosophische Reflexionen. Paderborn: W. Fink. S. 157-173.

Hemel, U. (2019). Mentale Architektur und Wirtschaftsanthropologie – eine Zukunftsaufgabe. In S. Kiessig & M. Kühnlein (Hrsg.), Anthropologie und Spiritualität für das 21.Jahrhundert. FS für Erwin Möde. Regensburg: Pustet. S. 335-350.

Hemel, U. (2020a). Kritik der digitalen Vernunft. Warum Humanität der Maßstab sein muss. Freiburg/Br.: Herder.

Hemel, U. (2020b). Unterscheidet die Fähigkeit zur Ökonomie den Menschen vom Tier? Eine Auslegung im Horizont menschlicher Symbolfähigkeit. In U. Lüke & G. Souvignier (Hrsg.), Der Mensch – ein Tier, Und sonst? Interdisziplinäre Annäherungen (Quaestiones Disputatae 307), Freiburg/Br. S.152-162.

Hemel, U. (2021). Digitale Fairness und digitale Humanität – was heißt Verantwortung in der digitalen Welt? In K. Kahle & N. Weidtmann (Hrsg.), Verantwortung, Ein Begriff in seiner Aktualität. Paderborn (Brill/Mentis), im Druck.

Hirschfelder, G., Lahoda, K. (2012). Wenn Menschen Tiere essen. Bemerkungen zu Geschichte, Struktur und Kultur der Mensch-Tier-Beziehungen und des Fleischkonsums. In J. Bucher-Fuchs & L. Rose (Hrsg.), Tierische Sozialarbeit. Wiesbaden: Verlag für Sozialwissenschaften. S. 147-166.

Kirchschläger, P.G. (2021). Digital Transformation and Ethics. Ethical Considerations on the Robotization and Automation of Society and the Economy and the Use of Artificial Intelligence. Baden Baden: Nomos.

Körtner, U.H. (2010). Leib und Leben. Bioethische Erkundungen zur Leiblichkeit des Menschen. Göttingen: Vandenhoeck & Ruprecht.

Lakatos, I. (1976). Falsification and The Methodology of Scientific Research Programmes. In S.G. Harding (Ed.), Can Theories be Refuted? Boston: D. Reidel, pp. 205-259.

Lakatos, I. (1979). Beweise und Widerlegungen. Braunschweig/Wiesbaden: Vieweg.

List, C. (Hrsg.) (2021). Warum der freie Wille existiert. Darmstadt: wbg.

Lüke, U., Souvignier, G. (Hrsg.) (2020). Der Mensch – ein Tier, Und sonst? Interdisziplinäre Annäherungen (Quaestiones Disputatae 307). Freiburg: Herder.

Lütteke, F. (2014). Vielseitiges autonomes Transportsystem basierend auf Weltmodellerstellung mittels Datenfusion von Deckenkameras und Fahrzeugsensoren. Erlangen-Nürnberg: Meisenbach.

Mastronardi, P. & Taubert, D. (2004). Das Recht im Spannungsfeld utilitaristischer und deontologischer Ethik. Wiesbaden: Franz Steiner.

Mead, G.H. (1978). Geist, Identität und Gesellschaft aus der Sicht des Sozialbehaviorismus, Frankfurt a.M.: Suhrkamp.

Misselhorn C. (2018). Grundfragen der Maschinenethik. Stuttgart: Reclam.

Misselhorn, C. (2018). Grundfragen der Maschinenethik. Stuttgart: Reclam.

Mitchell, M. (2019). Artificial intelligence: A guide for thinking humans. Penguin UK.

Moor, J.H. (1976). An Analysis of the Turing test. Philosophical Studies, 30(4), Dordrecht-Holland, pp.249-257.

Neff, G., Nagy, P. (2016). Automation, Algorithms, and Politics, Talking to Bots: Symbiotic Agency and the case of Tay. International Journal of Communication, Vol. 10, pp. 124-178.

Nida-Rümelin, J., Weidenfeld, N. (Hrsg.) (2018). Digitaler Humanismus: eine Ethik für das Zeitalter der künstlichen Intelligenz. München: Piper.

Pfeifer, R., Iida, F. (2004). Embodied Artificial Intelligence: Trends and Challenges. In F. Iida, R. Pfeifer, L. Steels & Y. Kuniyoshi (Eds.), Embodied Artificial Intelligence. Berlin/ Heidelberg: Springer. pp. 1-26.

Resnik, D.B., Elliott, K.C. (2016). The Ethical Challenges of Socially Responsible Science. Accountability in Research, 23(1), pp.31-46.

Riemann, F. (2011). Grundformen der Angst. München/Basel: Reinhardt.

Said, E.W. (2009). Orientalismus. Frankfurt a.M.: S. Fischer.

Saygin, A.P., Cicekli, I., Akman, V. (2000). Turing Test: 50 Years Later. Minds and Machines, 10(4), pp. 463-518.

Schilling, H. (1961). Bildung als Gottesbildlichkeit: eine motivgeschichtliche Studie zum Bildungsbegriff. Grundfragen der Pädagogik. Freiburg: Lambertus.

Schlesinger, A., O'Hara, K.P., Taylor, A.S. (2018). Let's Talk About Race: Identity, Chatbots, and AI. Proceedings of the 2018 Conference on Human Factors in Computing systems, pp. 1-14.

Shilton, K. (2013). Values Levers: Building Ethics into Design. Science, Technology, & Human Values, 38(3), pp.374-397.

Solomon, M. (2012). Socially Responsible Science and The Unity of Values. Perspectives on Science, 20(3), pp.331-338.

Steinbock, B. (1978). Speciesism and the Idea of Equality. Philosophy, Vol. 53, No. 204, Cambridge Univ. Press, pp. 247-256.

Störzinger, T., Carros, F., Wierling, A., Misselhorn, C., Wieching, R. (2020). Categorizing Social Robots with Respect to Dimensions Relevant to Ethical, Social and Legal Implications. i-com: Vol. 19, No. 1. Berlin: De Gruyter, pp. 47-57. DOI: 10.1515/icom-2020-0005, abgerufen am 04.11.2021 um 14.32h.

Sunderland, J., Litoselliti, L. (2002). Gender Identity and Discourse Analysis: Theoretical and Empirical Considerations. Gender Identity and Discourse Analysis Vol.2, pp.1-39.

Verbeek, P.P. (2008). Morality in Design: Design Ethics and the Morality of Technological Artifacts. In P. Kroes, P.E. Vermaas, A. Light & S.A. Moore (Eds.), Philosophy and Design: from Engineering to Architecture, pp. 91-103. ttps://doi.org/10.1007/978-1-4020-6591-0_7, abgerufen am 04.11.2021 um 14.46h.

Vogt, M. (2021). Prinzip Nachhaltigkeit, Grundlagen und zentrale Herausforderungen. Freiburg/Br.: Herder.

Wawrzyniak, D. (2019). Tierwohl und Tierethik. Empirische und moralphilosophische Perspektiven. Bielefeld: transcript.

Alle Rechte vorbehalten.
Abdruck oder vergleichbare Verwendung von Arbeiten des Instituts für Sozialstrategie ist auch in Auszügen nur mit vorheriger schriftlicher Genehmigung gestattet.

Publikationen des IfS unterliegen einem Begutachtungsverfahren durch Fachkolleginnen- und kollegen und durch die Institutsleitung. Sie geben ausschließlich die persönliche Auffassung der Autorinnen und Autoren wieder.