

# Künstliche Intelligenz – Technische Autonomie

## Zur Vermenschlichung der Maschine und Maschinisierung des Menschen

■ Wolfgang Koch

*Alle Lebewesen verknüpfen Sinneseindrücke mit Erlerntem und Mitgeteiltem. So verschaffen sie sich ein Bild ihrer Umwelt und wirken zielgemäß. Künstliche Intelligenz automatisiert diese Vorgänge und steigert sie weit über natürliches Maß. Technische Autonomie automatisiert Wirkungsketten. Wie aber bleiben ‚kognitive Maschinen‘ nicht nur technisch, sondern auch ethisch und psychologisch beherrschbar? Was machen sie mit uns, wenn wir naiv sind? Letztlich wird Künstliche Intelligenz durch natürliche Dummheit zum Problem, durch die Weigerung, in vollem Sinne Mensch zu sein. Bleiben wir verantwortliche „Herren“ unserer Maschinen, bleiben sie nützlich – wenn aber nicht, drohen Gefahren. Vor allem Christen müssen das rechte Bild vom Menschen lebendig halten.*

Eine Bildersuche im Internet bietet aufschlußreiche Psychogramme der vernetzten Menschheit, ihrer Assoziationen, Erwartungen, Befürchtungen. Sucht man nach „Künstlicher Intelligenz“, KI, so erscheinen in dominierendem Schwarz, Dunkeltürkis, Blau oder Weiß intelligent anmutende Wesen, die sich kühl und überlegen aus Schaltungen, Hardware und Datensymbolen erheben – ganz anders und doch „nach dem Bilde des Menschen“. Michelangelos Deckenfresko der „Schöpfung“ lebt offenbar tief im kollektiven Unbewußten. Denn es prägt diese Bildsprache. Auch als Rodins „Denker“ posieren die Roboter im *artist's view*. Meist symbolisiert ein Lichtblitz spontan entstehende „Intelligenz“ technischer Systeme, die in ihren Komponenten nicht angelegt ist.

Neu ist diese Metapher nicht. Sprach nicht schon Konrad Lorenz von „Fulguration“? Abgeleitet von *fulgur* – Blitz, nennt er so den vage bleibenden Vorgang, wie menschliches Erkennen blitzhaft im Evoluti-

onsgeschehen entstand.<sup>1</sup> Für die alten Philosophen und Mystiker bezeichnet *fulguratio* dagegen das Wirken Gottes. Und wie die Transhumanisten der Gegenwart verspricht schon Lorenz kollektive Unsterblichkeit: „Potentiell ist unsere Kultur durch die von ihrer Naturwissenschaft geleistete Reflexion in die Lage versetzt, dem Untergang zu entgehen, dem bisher alle Hochkulturen zum Opfer gefallen sind“, schließt sein „Versuch einer Naturgeschichte menschlichen Erkennens“. *Zum ersten Mal* in der Weltgeschichte sei das so.<sup>2</sup>

Aber im Google-Psychogramm schwingt auch Bedrohliches mit, ein Lebensgefühl, das bedeutende Wissenschaftler und Unternehmer zum Ausdruck bringen: „Die Entwicklung der Künstlichen Intelligenz könnte das Ende der menschlichen Rasse bedeuten“, prophezeit Stephen Hawking. „Aus menschlicher Kontrolle gelöst, würde sie sich immer rascher entwickeln. Menschen, an die langsame biologische Evolution gebunden, könnten nicht konkurrieren und würden

verdrängt.“<sup>3</sup> *Mark my words — A.I. is far more dangerous than nukes*, warnt Elon Musk,<sup>4</sup> der an vielen Fronten Technologien vorantreibt: autonome Elektroautos, Fernverkehr als „Rohrpost“, Low-Cost-Solar- und Raketentechnik. Selbst ein KI-Pionier, fordert der liberale Unternehmer staatliche Intervention und läßt *brain machine interfaces* entwickeln, Kopplungen zwischen Menschenhirn und Computerprothese, um die menschliche Intelligenz der künstlichen ebenbürtig zu halten. Wer KI-Algorithmen erforscht, weiß, daß es um Leben und Tod gehen kann: Wie entscheiden autonome Fahrzeuge bei einer unvermeidlichen Kollision? Für den Insassen oder Passanten? Verschärft gilt dies für autonome Waffen und den Schutz vor Angreifern, die solche Waffen wohl bald einsetzen.

Viele Länder treiben Künstliche Intelligenz und Technische Autonomie voran. „Insbesondere wollen wir Deutschland zu einem weltweit führenden Standort bei der Erforschung von künstlicher Intelligenz machen“,

»Wer als Christ inmitten der Digitalisierung urteils- und handlungsfähig bleiben will, **muß sich zunächst seiner Fundamente vergewissern.**«

steht im Koalitionsvertrag der Bundesregierung. „Wir wollen ein Mobilitätsforschungsprogramm inklusive der Erforschung der autonomen Mobilität auflegen“, heißt es weiter. Man wolle man die rechtlichen Voraussetzungen für vollautonome Fahrzeuge schaffen. Autonomen Waffen möchte die regierende Koalition jedoch Grenzen setzen: „Autonome Waffensysteme, die der Verfügung des Menschen entzogen sind, lehnen wir ab. Wir wollen sie weltweit ächten.“<sup>45</sup>

Wer als Christ inmitten der Digitalisierung urteils- und handlungsfähig bleiben will, muß sich zunächst seiner Fundamente vergewissern. Sie geben Orientierung, um die neuen Technologien einzuordnen, zu verstehen, was sie *nicht* sind, und warum sie einen „Herrn“ brauchen. Ein Extrembeispiel ist die militärische Nutzung der KI, an der sich die Probleme bei ziviler Nutzung wie im Scheinwerferlicht zeigen. Wer kognitive Maschinen zum Guten nutzen will, darf sie nicht in die Hand von Ideologen fallen lassen, die das Menschenbild verdunkeln, Profit und politische Macht erstreben. „Starke KI“, sich ihrer selbst bewußte, intelligent und autonom handelnde Maschinen, wird es nicht geben. Real ist die transhumanistische Versuchung, durch Technologie den Menschen selbst anzutasten.

### Natürliches und Technisches

Wenn die Begriffe „Künstliche Intelligenz“ und „Technische Autonomie“ fallen, muß klar bleiben, daß von Informations- und Automatisierungstechnik die Rede ist. Was ist also Technik? Der Verein Deutscher Ingenieure definiert Technik als „Menge der nutzenorien-

tierten, künstlichen, gegenständlichen Sachsysteme“.<sup>46</sup> Da Nutzen im Vordergrund steht, impliziert Technik den sie nutzenden Menschen. Technikdisziplinen erstreben demnach instrumentelle Kenntnisse zur Weltgestaltung, sind also *artes serviles*, dienende Künste, gemäß der klassischen Lebenslehre. Definitionsgemäß sind technische Systeme „gegenständliche Sachsysteme“, Objekte, keine Subjekte.

Entscheidend ist das Attribut „künstlich“. Wer wissen will, was unter „Künstlichem“ zu verstehen ist, muß begreifen, was „Natürliches“ ist. Denn das Begriffspaar natürlich – künstlich steht als fundamentale Dichotomie am Anfang abendländischen Denkens. Mit ihr beginnt Aristoteles das zweite Buch seiner *Physik*: „Von dem, was ist, ist einiges von Natur, anderes durch andere Ursachen. [...] Was nicht naturbeschaffen ist, ist vom Menschen durch Kunst erschaffen und erhalten und hat keinen Anfang in sich selbst.“<sup>47</sup>

Wenn Aristoteles von „Natur“ spricht, gebraucht er das griechische Wort *physis*, das die Natur einer bestimmten einzelnen Sache benennt, so wie etwa Juristen von der Natur eines Sachverhalts sprechen. Für „Natur“ im Sprachgebrauch etwa des Naturschutzes kennt er die Worte *kosmos*, *taxis*, *uranos*. Der Physiker und Philosoph Carl Friedrich von Weizsäcker bekennt über die *Physik* des Aristoteles: „Als Schüler Bohrs und Heisenbergs kann ich nicht umhin, meine Begeisterung über eine physikalisch so gesunde Denkweise auszusprechen.“<sup>48</sup>

Künstliches ist also negativ definiert. Fundamentalere ist Natürliches: Was nicht natürlich ist, ist „durch Kunst erschaffen“. Das griechische



»Was nicht naturbeschaffen ist, ist vom Menschen durch Kunst erschaffen und erhalten und hat keinen Anfang in sich selbst ...« (Aristoteles) Bild: Freskoausschnitt aus „Die Schule von Athen“ von Raffael (1483 – 1520) in den Stanzen des Vatikan.

Wort dafür ist *technē*. Künstliches, also durch Kunst Erschaffenes, ist demnach ‚Technisches‘. Offenbar ist der moderne Technikbegriff enger gefaßt als dasjenige, was Aristoteles durch *technē* benennt. Aber insbesondere ist ‚Technisches‘ im heutigen Sinne ‚durch Kunst Erschaffenes‘. Was aber kennzeichnet ‚Natürliches‘, dessen Gegenteil ‚Technisches‘ ist?

Natürliches habe „einen Anfang in sich selbst“. Die Bedeutung des griechischen Wortes *archē*, Anfang, spiegelt sich im deutschen „Ursprung“, lateinischen *principium* wider. Aristoteles definiert präziser: „was in sich selbst *einen* Anfang seines Wirkens und Bestands hat“. *Einen* Anfang oder Ursprung hat Natürliches „in sich selbst“ – es mag weitere Ursachen ►

seines Wirkens und Bestands geben. Neben äußeren Faktoren ist es jedoch der innere „Anfang“, der natürliches Wirken und Entfalten, natürliche „Bewegung“ bestimmt. Wie definiert Aristoteles natürliches Wirken? Es ist „Zur-Wirklichkeit-Kommen eines bloß der Möglichkeit nach Vorhandenen“.<sup>9</sup> Genau dieses macht für ihn das Wesen eines Naturbeschaffenen aus.

Bei Technischem gibt es keinen inneren „Anfang“ seines Wirkens in dem, was es von sich aus ist. Daher kann Robert Spaemann formulieren: „Kunst Dinge, so Aristoteles, sind in der Tat dadurch charakterisiert, dass sie selbst aus einem ‚Was‘ und einem ‚Woraus‘ bestehen.“ Ihr ‚Wie‘ und ‚Wozu‘ liege nicht in ihnen, sondern im Menschen. „Ein Naturding ist hingegen dadurch charakterisiert, dass ‚Was‘ und ‚Wozu‘ *in ihm selbst in eins fallen*. Der Zweck ist die Form der Sache selbst, darum also auch das Wort *entelecheia*: ich trage das Ziel in mir.“<sup>10</sup> Naturdinge, *physei onta*, sind also Entelechien, tragen in sich selbst eine immanente Prozeßhaftigkeit, die sie zu dem macht, was sie ihrer Anlage nach sind. Daher kann Thomas von Aquin kommentieren: „Natur‘ ist also nichts anderes als ein Prinzip des Werdens und ruht in dem, in dem es primär und *per se* und nicht akzidentiell ist.“<sup>11</sup>

Für Josef Pieper ist dies „ein Urgedanke der abendländischen Metaphysik überhaupt: der Gedanke von der Entelechie als dem inneren *telos*. Der Sinn, d.i. der innere Zweck, das immanente Ziel jedes Dinges ist sein eigenes Wesen, seine innere Form. Vollenendet sein heißt: ganz sein Wesen sein. Werden heißt: Formverwirklichung, Ausverwirklichung des Sinnes, des inneren Zweckes. Darum ist Abkehr vom Sinn das Gegenteil von Werden, ‚Verderb‘; darum ist Rückkehr zum Sinn: ‚Gesundung.‘“<sup>12</sup> Technische Dinge sind gerade *keine* Entelechien. Auf dieser Unterscheidung beruhe „die ganze überlieferte Lebenslehre, die Überzeugung, daß im ganzen Bereich menschlichen Wirkens das Erste immer das ist, was der Mensch und die Dinge ‚von Natur‘ sind – wobei ‚von Natur‘ soviel besagt wie: kraft der Erschaffung, von Schöpfungen wegen“.<sup>13</sup>

## Was ist KI, was ist sie nicht?

Bereits vor jeder wissenschaftlichen

Reflexion oder technischen Realisierung sind ‚Intelligenz‘ und ‚Autonomie‘ allgegenwärtig: Jedes Lebewesen verknüpft unterschiedliche Sinnesindrücke mit erlerntem Wissen und Mitteilungen anderer Lebewesen. So entsteht ein Bild seiner Umwelt, die Voraussetzung für situationsgerechtes Wirken in der Biosphäre. Ein Beispiel wäre die natürlich-intelligente Jagd eines Wolfsrudels auf einen Bison, der sich ebenso natürlich-intelligent zu wehren weiß. Lebewesen, vor allem der Mensch, erschöpfen sich natürlich nicht in solchen Funktionen.

KI-basierte Systeme versuchen, das Sammeln von Daten durch Sensoren und aus Datenbanken technisch nachzubilden und vor allem ihre Verknüpfung, die Sensordaten- und Informationsfusion, wissenschaftlich zu verstehen, soweit wie möglich zu auto-



»Michelangelos Deckenfresko der „Schöpfung“ lebt offenbar tief im kollektiven Unbewußten.«

Bild: creative commons Lizenz

matisieren und in bestimmten Aspekten weit über das natürliche Wahrnehmungsvermögen hinaus zu steigern.<sup>14</sup> Technische Autonomie befaßt sich mit der Automatisierung von Wirkungsketten, der Steuerung von Sensoren, Plattformen, Wirkmitteln.

Es geht also um eine Art „Servolenkung fürs Hirn“,<sup>15</sup> kognitive Tools, die natürliche Fähigkeiten zur Wahrnehmung ebenso steigern wie mechanische Werkzeuge die Körperkraft. Sie sind erforderlich, sobald Daten in solcher Fülle einströmen, daß sie anders schon rein quantitativ nicht mehr weiterzuverarbeiten, zu verknüpfen, zu bewerten sind. Nur so bleibt verantwortetes Wirken in der zunehmend komplexen Technosphäre möglich. Diese Unterstützung der Wahrnehmung erfolgt auf tieferer Ebene als durch Brillen oder Hörgeräte, Fernrohre oder Mikroskope. Der Maschinenbau für kognitive Tools hat gewaltiges wirtschaftliches Potential, profitiert von Trends wie

Vernetzung, Mobilität, Sensor- und Datenbanktechnologie und schafft die Basis für *Manned-unmanned Teaming*.

Insofern künstlich intelligente und technisch autonome Systeme intelligentes Wahrnehmen und autonomes Wirken von Personen unterstützen, die für ihr Handeln verantwortlich sind, bleiben sie ein ‚etwas‘ für ‚jemanden‘, sind Objekte, keine Subjekte. Daher sind Künstliche Intelligenz und Technische Autonomie von natürlicher Intelligenz und Autonomie zu unterscheiden. Bei allem Nutzen können von derartiger Automatisierungstechnik Gefahren ausgehen, wie von anderer Automatisierung auch.

KI-assistiertes Wissen um Bedrohungen, Unbeteiligte und Handlungsoptionen, sei es im Straßenverkehr oder im Gefecht, helfen, komplexe Missionen situationsadäquater zu

meistern, menschliche Subjektivismen auszugleichen, den Schutz eigener Kräfte und Unbeteiligter zu verbessern. Drei Bereiche sind zu unterscheiden: Steigerung der Wahrnehmung durch *sensorielle Assistenz*, des Lagebewußtseins durch *kognitive Assistenz*, der Wirkung durch *physische Assistenz*. Die physische Präsenz von Menschen wird in gefährlichen Situationen entbehrlich.

Im Zentrum steht kognitive Assistenz, um 1) massenhaft einströmende Sensordaten auszuwerten, 2) komplexes Kontextwissen mit aktuellen Daten zu verknüpfen, 3) sich ergänzende und verschiedenartige Sensoren zu verknüpfen, 4) die Plausibilität der gewonnenen Information abzuschätzen, 5) zielentsprechende Wirkungsoptionen, ggf. basierend auf mobilen und interagierenden Plattformen, anzubieten und dabei 6) den ethischen Wirkungsrahmen zu respektieren. Schließlich soll sich KI-basierte Technische

»Insofern künstlich intelligente und technisch autonome Systeme intelligentes Wahrnehmen und autonomes Wirken von Personen unterstützen, die für ihr Handeln verantwortlich sind, **bleiben sie ein ‚etwas‘ für ‚jemanden‘, sind Objekte, keine Subjekte.**«

Autonomie individuellen Aufgaben und Fähigkeiten ihrer Nutzer anpassen. Sie soll also Menschen entlasten, damit diese tun können, was *nur* Menschen können, nämlich intelligent und in verantworteter Autonomie wirken. Eine wichtige Forderung ist Nachvollziehbarkeit der Angebote.

Steht die Menschheit angesichts Künstlicher Intelligenz und Technischer Autonomie also vor fundamental neuen Herausforderungen? Nein. Neue Technologien haben in der Geschichte immer wieder zu gesteigerten Möglichkeiten der Wahrnehmung und einem vergrößerten Wirkungskreis geführt. Der Unterschied zu früheren technologischen Umbrüchen liegt also eher auf quantitativer als auf qualitativer Ebene. Daher stellt sich auch heute eine zeitlose Frage: Wie wirken wir gut? Jede Antwort darauf setzt informationswissenschaftliche Antworten auf zwei verschiedene Fragen voraus:

- 1) Wie bleiben kognitive Maschinen von Menschen beherrschbar?
- 2) Wie ist ihre verantwortbare Nutzung technisch zu ermöglichen?

### Romano Guardinis Analyse

Wissenschaftliches Nachdenken über „Elektronengehirne“ beginnt in den 1940er und 1950er Jahren mit der Erfindung des Computers und der Entdeckung, daß hochvernetzte Nervenzellen untereinander elektrische Impulse austauschen. Gedankliche Voraussetzungen schaffen Norbert Wiensers Kybernetik<sup>16</sup> zur Steuerung elektrischer Netze, Claude Shannons Informationstheorie und Alan Turings Berechenbarkeitstheorie.

Das *Dartmouth Summer Research Project on Artificial Intelligence* bringt 1956 die neue informationswissen-

schaftliche Disziplin hervor, die von John McCarthy<sup>17</sup> und Marvin Minsky<sup>18</sup> organisiert und von den Informatikpionieren Nathaniel Rochester<sup>19</sup> und Claude Shannon unterstützt wird. Schon der Antrag formuliert die These, „daß jeder Aspekt des Lernens oder jedes andere Merkmal der Intelligenz im Prinzip so genau beschrieben werden kann, daß eine Maschine sie simulieren kann. Untersucht werden soll, wie man Maschinen dazu bringt, Sprache zu nutzen, Abstraktionen und Begriffe zu bilden, Probleme zu lösen, die heute dem Menschen vorbehalten sind, und sich selbst zu verbessern.“<sup>20</sup>

Den Umbruch, den wir heute „Digitalisierung“ nennen, erspürt Romano Guardini bereits 1950 mit seismographischer Sensibilität als ein sich radikal wandelndes Menschen- und Naturbild. Die Entfesselung natürlicher Möglichkeiten durch Technik gebe dem Verhältnis des Menschen zur Natur „den Charakter äußerster Entscheidung: entweder gelingt es dem Menschen, das Herrschaftswerk richtig zu machen, und dann wird es gewaltig – oder aber alles geht zu Ende“.<sup>21</sup>

„Es gibt kein Seiendes, das herrenlos ist“, stellt Guardini fest. „Die Natur als nicht-personale Schöpfung gehört Gott. Was aber im Freiheitsbereich des Menschen erscheint, muß einem Menschen gehören und von ihm verantwortet werden“, folgert er gemäß den einleitenden Unterscheidungen von Natürlichem und Technischem. „Geschicht das nicht, dann wird es nicht wieder zu ‚Natur‘ [...]; es bleibt nicht einfach disponibel, sondern etwas Anonymes ergreift von ihm Besitz.“<sup>22</sup>

Unmißverständlich benennt Guardini, was aus seiner Sicht geschehen kann: „Von der Macht des Menschen



»Den Umbruch, den wir heute „Digitalisierung“ nennen, erspürt Romano Guardini bereits 1950 mit seismographischer Sensibilität als ein sich radikal wandelndes Menschen- und Naturbild.« Bild: Romano Guardinis (1885 – 1968) Grab in St. Ludwig in München.

ergreifen die Dämonen Besitz [...], geistige Wesen, die von Gott gut geschaffen sind, aber von ihm abgefallen; die sich für das Böse entschieden haben und nun entschlossen sind, Seine Schöpfung zu verderben. Diese Dämonen sind es, welche dann die Macht des Menschen regieren: durch seine scheinbar natürlichen, in Wahrheit so widersprüchlichen Instinkte; durch seine scheinbar folgerichtige, in Wahrheit so leicht beeinflussbare Logik; ►►

durch seine unter aller Gewalttätigkeit so hilflose Selbstsucht.“

Offenbar ist auch Papst Franziskus von dieser Problematik bewegt. Er spricht sie nicht nur bei zahlreichen Gelegenheiten an, sondern läßt Konferenzen organisieren, um sie wissenschaftlich zu untersuchen, zuletzt im Mai 2019.<sup>23</sup> Im Programmheft zitiert Franziskus aus dem *Ende der Neuzeit*: Es sei eine Tatsache, „daß der moderne Mensch nicht zum richtigen Gebrauch der Macht erzogen wird; ja daß weithin sogar das Bewußtsein des Problems fehlt“.<sup>24</sup> Guardini steht Franziskus nahe. 1986 beginnt er in Frankfurt über ihn zu arbeiten und ermutigt die Ein-

eine verwundete, zum Bösen geneigte Natur habe, führe zu schlimmen Irrtümern im Bereich der Erziehung, der Politik, des gesellschaftlichen Handelns und der Sittlichkeit.<sup>27</sup> Wer dieses uralte Wissen um den Menschen abtut, erzeugt einen „blinden Fleck“ seiner Wahrnehmung. Ein Kenner der menschlichen Natur wie Goethe, den man kaum einen Christen nennen kann, denkt vergleichbar: „Wenn gewisse Erscheinungen der menschlichen Natur, betrachtet von seiten der Sittlichkeit, uns nötigen, ihr eine Art von radikalem Bösen, eine Erbsünde, zuzuschreiben, so fordern andere Manifestationen derselben, ihr gleichfalls eine Erbtugend [...] zuzugestehen.“<sup>28</sup>

### Was sind Neuronale Netze?

Algorithmen der Künstlichen Intelligenz und Technischen Autonomie treiben den Informationskreislauf an, der handelnde Personen befähigt, das Geschehen in komplexen Situationen wahrzunehmen und darin zu wirken. Ausgehend von physikalischen Signalen, vorverarbeiteten Sensordaten, Datenbankeinträgen oder sprachlichem Input, liefert KI Lagebilder. Technische Autonomie wandelt Absichten in komplexe Steuerbefehle für Sensoren, Plattformen und Wirkmittel um. Generell sind *datengetriebene* von *modellbasierten* Algorithmen zu unterscheiden. Die erste Familie entspricht unreflektierter Sinneswahrnehmung – Was sehe ich? Die zweite ermöglicht Kausalitätsverstehen – Warum geschieht etwas?

In bestimmten Anwendungen überaus erfolgreiche Algorithmen zu Beantwortung von Was-Fragen sind Neuronale Netze, die sich extrem vergrößert an der Signalverarbeitung vernetzter Nervenzellen orientieren. Auch diese Ansätze reichen bis in die 1950er Jahre zurück. Wer ihre Chancen und Risiken einschätzen will, muß ihren Grundgedanken verstehen.

Neuronale Netze sind mathematisch formulierte Vorschriften, Funktionen, die einem Input einen Output zuordnen. Dieser Output beschreibt, was der Input für den Nutzer bedeuten soll. Die Besonderheit besteht darin, daß derartige Funktionen extrem viele Parameter, wählbare Zahlenwerte, besitzen. In einer „Trainingsphase“ werden sie eingestellt, indem man etwa

für „verständene“ Fotos dem Neuronalen Netz zugleich mitteilt, was darauf zu sehen ist. Dieses *labeling* nutzt menschliches Verstehen. Hat man ‚lange genug‘ trainiert, bietet man dem Netz beliebige Fotos an und betrachtet den Output als das erkannte Was auf dem Foto. Neuronale Netze sind also Funktionsapproximatoren. Wer das massenhafte Anbieten von Stützstellen „Lernen“ nennt, weckt bei Nicht-Informatikern irriige Assoziationen.

Der Durchbruch bei Neuronalen Netzen mit ihren zum Teil spektakulären Erfolgen hat drei Ursachen: Nicht zuletzt durch Soziale Netze stehen erstens überaus große Datenmengen mit zugehörigen „Metadaten“ zu Verfügung, um auch „tiefe Netze“ mit unvorstellbar vielen Parametern trainieren zu können (*deep learning*). Zweitens schaffen die boomenden Computerspiele einen Massenmarkt für schnelle Graphikkarten und von ihnen abgeleitete Rechnerarchitekturen, durch die sich die Trainingsphase massiv beschleunigen läßt. Und schließlich bietet etwa die individualisierte Werbung aufgrund erkannter Käuferprofile einen auch finanziell höchst lukrativen *business case*.

Eine Schockwelle durchlief allerdings 2015 die *AI Community*, als Ian Goodfellow auf einer der führenden Konferenzen über *poisonous noise*, „vergiftetes Rauschen“ sprach. Er präsentierte ein Neuronales Netz, das zuverlässig Tiere auf Fotos „erkennt“, und beschrieb, wie er mit einem anderen Netz dessen Schwächen untersuchte. Wie er zeigen konnte, muß man nur wenige Bildpunkte auf einem Foto verändern, es gezielt „verrauschen“, um ein gut trainiertes Netz völlig in die Irre führen: Einen für Menschen unverändert erscheinenden Pandabären „erkennt“ das Netz als Gibbon-Affen und ist sich dabei völlig sicher.<sup>29</sup>

Die Hoffnung, darin zeige sich nur eine schlecht gewählte Netzarchitektur, zerschlägt sich 2017. Ein Netz zur Erkennung von Verkehrszeichen für autonomes Fahren liefert „selbstbewußt“ kraß falsche Ergebnisse, sobald ihm, für Menschen nicht wahrnehmbar, speziell „verschmutzte“ Verkehrschilder angeboten werden.<sup>30</sup>

Sogleich entsteht ein Zweig der Informatik – *the Byzantine science of deceiving artificial intelligence*: „Es ist



»Wer das massenhafte Anbieten von Stützstellen „Lernen“ nennt, weckt bei Nicht-Informatikern irriige Assoziationen.«

leitung des Seligsprechungsprozesses für ihn.<sup>25</sup> Dies gilt wohl insbesondere für den Gedanken einer dämonischen Besitzergreifung, da für Franziskus das Wirken solcher Kräfte eine Gegebenheit ist.

Wer die menschliche Verantwortung für Technik betrachtet, muß um das Wesen des Menschen wissen. Für Guardini „bleibt das Wissen um die Person mit dem christlichen Glauben verbunden. Ihre Bejahung und ihre Pflege überdauern wohl eine Weile das Erlöschen des Glaubens, gehen aber dann allmählich verloren.“<sup>26</sup> Dazu gehört das Wissen um die Realität der Erbsünde, durch die „der Teufel eine gewisse Herrschaft über den Menschen erlangt, obwohl der Mensch frei bleibt“, wie der Katechismus formuliert. Zu übersehen, daß der Mensch

»Von der Macht des Menschen ergreifen die Dämonen Besitz [...], Diese Dämonen sind es, welche dann die Macht des Menschen regieren: durch seine scheinbar natürlichen, in Wahrheit so widersprüchlichen Instinkte; **durch seine scheinbar folgerichtige, in Wahrheit so leicht beeinflussbare Logik**; durch seine unter aller Gewalttätigkeit so hilflose Selbstsucht.« (Romano Guardini)



»Wer KI-Algorithmen erforscht, weiß, daß es um Leben und Tod gehen kann: Wie entscheiden autonome Fahrzeuge bei einer unvermeidlichen Kollision? Für den Insassen oder Passanten?«

on. Zudem erzeugen KI-Algorithmen stets auch Artefakte, Dinge, die es in Wirklichkeit nicht gibt, oder weisen „blinde Flecken“ der Wahrnehmung auf, sehen nicht, was es dennoch gibt. Zu nennen ist auch die Cyber-Bedrohung, die Übernahme von Sensoren oder Subsystemen durch Gegner, die dann „bösaartig“ werden, täuschende Daten liefern oder sich gegen den Nut-

wahrscheinlich für einen Gegner recht einfach, uns hereinzulegen, uns zu täuschen. Manche diese Versuche können harmlos sein, manche sind es nicht“,<sup>31</sup> faßt das *US Army Research Laboratory* zusammen. Die militärische Relevanz dieser Erkenntnis ist offensichtlich. Längst werden Angriffssysteme gegen feindliche KI entwickelt, eigene KI-Systeme gegen solche Angriffe gehärtet, Angriffsstrategien gegen solchermaßen gehärtete gegnerische KI-Systeme erforscht usw.

Neben der Täuschbarkeit datengetriebener Algorithmen gibt es in vielen Anwendungen nicht genügend viele und ausreichend repräsentative Trainingsdaten (*sparse data*). Für militärische Nutzung ist dies typisch. Zudem sind Neuronale Netze *black boxes*. Niemand weiß, wie sie zu ihren Ergebnissen gelangen. Ferner läßt sich Kontextwissen nur aus den Daten selbst erlernen; Datenbanken sind nicht „anschließbar“. Kurz gesagt: sie sind *greedy, brittle, and opaque*, datenhungrig, störbar und undurchsichtig.

### Technische Beherrschbarkeit

Jede verantwortbare Nutzung KI-basierter Systeme setzt durchgängige technische Beherrschbarkeit voraus – *meaningful human control*. Bei individualisierter Werbung sind gelegentliche Fehler folgenlos, vielleicht sogar komisch. Sicherheitskritische Anwendungen bei Fahrerassistenzsystemen im Verkehr, erst recht bei autonomem Fahren, oder in der Luftfahrt erfordern dagegen eine Zulassung mit allen rechtlichen Folgen, etwa im Sinne der Haftung oder Versicherbarkeit. Die tödlichen Folgen unausgereifter Software mit gravierenden ökonomischen Konsequenzen für ein Weltunternehmen zeigt das Flugzeugmodell *Boeing 373 MAX*.

Eine Schwachstelle ist die Datenintegrität: Sind die Sensor- und Kontextdaten korrekt, und entsprechen die stets unvermeidlichen Fehler den Annahmen über sie? Bei verletzter Datenintegrität wird in naiven Systemen aus Datenfusion leicht Konfusi-



»Die tödlichen Folgen unausgereifter Software mit gravierenden ökonomischen Konsequenzen für ein Weltunternehmen zeigt das Flugzeugmodell *Boeing 373 MAX*.« Bild: Nach zwei Abstürzen von *Boeing 737 MAX 8* – zuletzt mit 346 Todesopfern - aufgrund identischer Fehlfunktionen sind seit März 2019 alle 371 Maschinen des Typs weltweit mit Flugverbot belegt. ▶▶

zer wenden. Zu fordern ist daher von jeder KI sozusagen eine Fähigkeit zu „Selbstkritik“.

Aber auch die Nutzer müssen Künstlicher Intelligenz gegenüber kritikfähig bleiben. Sonst entsteht die Gefahr freiwilliger Unterordnung und kritikloser Akzeptanz der Angebote, die mentale Weigerung, Verantwortung tatsächlich zu tragen, blindes Vertrauen. KI-basierte Systeme müssen die Wachsamkeit ihrer Nutzer trainieren, ihnen ein *Tell me why* für Lösungsangebote vermitteln und dürfen sie nicht verdummen. Nur menschliche Intelligenz vermag Plausibilitäten einzuschätzen, tatsächliches Verständnis zu entwickeln.

Eine Schlüsselfunktion besitzen modellbasierte Algorithmen, die den Namen des Pioniers der mathematischen Statistik und anglikanischen Geistlichen Thomas Bayes tragen. Sie

Da mathematische Modellierung oft zu komplex oder gar nicht möglich ist, bleiben *datengetriebene* Algorithmen unverzichtbar. Also liegt es nahe, beide Ansätze zu verknüpfen und nach *Explainable AI* zu suchen. Seit kurzem befaßt sich eine rasch wachsende *Community* damit, auch für Neuronale Netze „Unschärfe“ quantitativ darzustellen, stochastische Methoden anzuwenden und probabilistische Qualitätsmaße für den Output Neuronaler Netze anzugeben. Fesselnd ist ein Sachbuch Judea Pearls, des Erfinders der Bayes-Netze, das philosophische Fragen berührt und etwa bei der Suche nach algorithmischen Antworten auf Warum-Fragen die aristotelische *causa finalis* zu Ehren bringt.<sup>33</sup>

Was-Fragen führen zum philosophischen Problem der *qualia*, des subjektiven Erlebnisgehalts mentaler

All diese Forschungsfragen sind noch unbeantwortet. Dennoch sind Zertifizierung und Zulassung bereits jetzt gefordert. Robuste und praktisch nutzbare KI-Systeme werden sowohl datengetriebene als auch modellbasierte Algorithmen umfassen. Um die Leistungsfähigkeit des *deep learning* zu nutzen und ihre Probleme einzudämmen (*greedy, brittle, opaque*), nutzt man das Prinzip: *AI in the Box*. Systemisch werden datengetriebene Algorithmen durch modellbasiertes *Reasoning* „eingehegt“. Nachzuweisen sind vorhersagbare Systemeigenschaften, Insensitivität gegenüber unbekanntem Einflüssen,<sup>35</sup> Adaptivität gegenüber variablem Einsatzkontext und *graceful degradation*, graduelle und beherrschbare Leistungseinbußen beim Ausfall von Komponenten.

### Beispiel: Militärische KI

Ein Extrembeispiel, das zugleich Grundprobleme ziviler KI-Systeme wie im Scheinwerferlicht sichtbar macht, ist ihre militärische Nutzung. Denn Künstliche Intelligenz und Technische Autonomie verändern nicht nur die Wirtschaft und die Weise, wie wir leben, sondern auch Verteidigung und Sicherheit. „Wer in diesem Bereich die Führung übernimmt, wird Herrscher der Welt“, erklärt Wladimir Putin.<sup>36</sup> Besorgt um seine Souveränität in der sich auflösenden Nachkriegsordnung, beginnt Europa die Notwendigkeit zu erkennen, sich selbst verteidigen zu können.

Konflikte werden künftig automatisierter ausgetragen als je zuvor. Entwickelt werden nicht nur Schwärme bewaffneter Mikrodrohnen oder hypersonische Waffensysteme, die weltweit ihr Ziel in weniger als zehn Minuten treffen. KI-basierte Autonomie ist auch das Grundprinzip unbenannt interagierender Panzer im Verbund mit Drohnen, die als *Main Ground Combat System* MGCS traditionelle Panzer ablösen und die europäischen Außengrenzen verteidigen sollen. Weiter vorangeschritten ist das *Future Combat Air System* FCAS, um den europäischen Luftraum zu schützen. Der Startschuß für die Entwicklung erfolgte am 17. Juni 2019.

Bei FCAS sollen unbemannte, künstlich intelligente und technisch autonome Luftfahrzeuge bemannte Kampffjets neuester Generation als lo-



»Ein Extrembeispiel, das zugleich Grundprobleme ziviler KI-Systeme wie im Scheinwerferlicht sichtbar macht, ist ihre militärische Nutzung.«

ermöglichen logisches Schließen auch bei unscharfen Ausgangsdaten.<sup>32</sup> Dadurch werden wahrscheinliche Kausalstrukturen aufgedeckt und mögliche Hypothesen quantitativ bewertbar. Zugleich führt *Bayesian Learning* nicht zu *black boxes*, sondern ermöglicht ein systematisches Algorithmendesign, in das Kontextinformation oder Expertenwissen explizit integrierbar ist. Für Massendatenanwendungen, *big data*, ist jedoch die Verfügbarkeit hochleistungsfähiger Rechnerarchitekturen noch ein *bottleneck*.

Zustände, zum Kernproblem jeder Philosophie des Geistes also. *What is it like?*, fragt Thomas Nagel, für den darauf nur Wesen mit ‚Geist‘ antworten können.<sup>34</sup> Den Informatikern ergeht es heute offenbar wie den Physikern vor hundert Jahren, als sie von physikalischen Problemen zur Philosophie fanden. Für den Verlauf der technologischen, aber auch geistesgeschichtlichen Entwicklung wird viel davon abhängen, ob sie unter Philosophen und Theologen die rechten Ansprechpartner finden.

»Nachdem wir die Menschheit über die animalische Ebene des Überlebenskampfes hinausgehoben haben, werden wir nun danach streben, **Menschen in Götter zu verwandeln** und aus dem **Homo sapiens** den **Homo deus** zu machen.« (Der Transhumanist Yuval Harari)

yale *wing men* begleiten Bei Angriffen gegen den Jet schützen sie den Piloten und lenken von ihm ab. In Kampfmissionen fliegen sie als *remote carrier* weit voraus, klären koordiniert auf und bekämpfen gegnerische Ziele. Auch andere Komponenten der Luftverteidigung werden durch FCAS als *System-of-Systems* vernetzt: der Eurofighter, die Militärtransporter, Lenkflugkörper oder AWACS. Als ambitioniertestes Rüstungsprojekt Europas soll FCAS zugleich europäische Spitzentechnologie vorantreiben.

Bereits jetzt weisen Forschungsprojekte in Feldtests nach, wie koordiniert interagierende Bodenroboter und Drohnen als Multisensorträger die Soldaten durch umfassende Lagebilder unterstützen.<sup>37</sup> So werden etwa getarnte Scharfschützen durch Fusion räumlich verteilter akustischer und multispektraler Sensorik lokalisiert, denen die Soldaten sonst ausgeliefert wären, das Kaliber verschossener Munition klassifiziert und Funkverbindungen aufgeklärt, über die gegnerische Kräfte kommunizieren. Fähigkeiten zu vernetzter Operation werden so zum Aufklärungsziel. Kampffähige Drohnen wären vor allem in unübersichtlicher, Hinterhalte provozierender Umgebung erwünscht.

Ein solches Beispiel wird im Bundestag geschildert: Der Abzug der letzten 441 Soldaten aus dem Feldlager Kunduz sei eine der größten Operationen dieser Art in der Geschichte der Bundeswehr gewesen. Zwei Tage lang habe sich der kilometerlange Konvoi aus 119 Militärfahrzeugen durch unübersehbares Gelände bewegt, Aufklärungsdrohnen ständig Patrouillenflüge durchgeführt. Ein Angriff wäre rechtzeitig erkannt worden. Seine Bekämpfung hätte jedoch nicht unmittelbar erfolgen können, weil dazu Luftunterstützung anzufordern gewesen wäre.

Entscheidende, das Leben der Soldaten akut gefährdende Zeit wäre verflissen. Die geschilderte Operation sei glücklich verlaufen. Davon sei aber nicht immer auszugehen. Der Bundeswehr gehe es in der Forderung nach bewaffneten Drohnen, wie das Beispiel veranschauliche, um das Schließen von Schutzlücken.<sup>38</sup>

Ein Forschungsprojekt untersucht dieses Problem und betrachtet in städtischer Umgebung einen Konvoi, der durch eine Sprengfalle gestoppt und aus dem Hinterhalt angegriffen wird.<sup>39</sup> Herausfordernd für die Verteidigung sind fehlende Sichtverbindungen zu den Angreifern, ein Nachteil, der durch einen Verbund aus flugzeug- und helikopterartigen *Unmanned Aerial Vehicles*, UAV, ausgeglichen wird. Drohnen der ersten Kategorie verfügen über längere Einsatzzeiten und können größere Nutzlasten tragen als die der zweiten Kategorie.

„Drehflügler“ sind flexibler einsetzbar als „Starrflügler“, da sie vom Dach eines Militärfahrzeugs aus starten und dort wieder landen, sowie für genauere Lageeinschätzungen in der Luft über einem bestimmten Ort verweilen können. Der Drohneneinsatz erfordert robuste und sichere Kommunikation sowie enge Verzahnung der Datenfusion mit adaptivem Management verfügbarer Sensor-, Plattform- und Kommunikationsressourcen. Koordinierter Drohneneinsatz ermöglicht, zu erwartende Kollateralschäden abzuschätzen, und verschafft dem *Forward Air Controller* ein umfassendes Lagebild. Er allein trifft und verantwortet möglicherweise letale Entscheidungen. Wie werden sie verantwortbar?

### Ethik als Designprinzip?

Lagebilder stellen komplexe Abläufe für Menschen faßlich dar und beantworten Einzelfragen wie: Existieren

im Sichtbereich der Sensoren Objekte von Interesse? Sind sie zu lokalisieren, zu verfolgen, zu klassifizieren oder gar zu identifizieren? Welche Interrelationen bestehen zwischen ihnen? Lassen sich Anomalien erkennen? Wesentlich ist die Darstellung verbleibender Ungenauigkeiten, Mehrdeutigkeiten, nicht oder noch nicht beantworteter Fragen. Es gehört also zu einem Lagebild, auch über unaufgeklärte Aspekte aufzuklären.

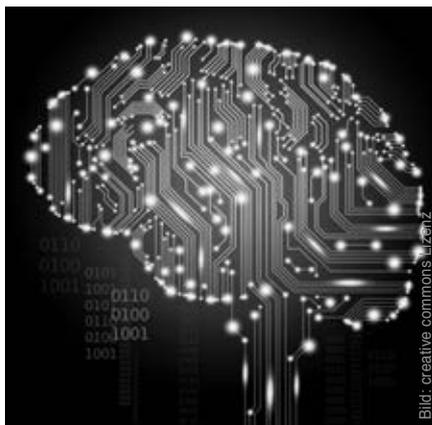
Herausfordernd ist die Tatsache, daß zu fusionierende Daten unvermeidlich unvollkommen sind, ungenau, unvollständig, nicht aufgelöst, falsch, unabsichtlich oder durch Gegner absichtlich verfälscht, schwer zu formalisieren, teilweise widersprüchlich. Durch mathematische Statistik gelingt es jedoch, auch für unpräzise Daten präzise arbeitende Algorithmen abzuleiten, die aus ihnen genauere, vollständigere, schneller verfügbare Information gewinnen, als es Menschen je vermöchten. Insbesondere werden Kollateralschäden auf bisher unerreichte Weise abschätzbar: *pre-engagement collateral damage prediction*. Im besprochenen Beispiel kommandiert der *Forward Air Controller* auf dieser Grundlage den Einsatz eines *Unmanned Combat Aerial Vehicle*UCAV. Durch Künstliche Intelligenz und Technische Autonomie wurde eine verlässliche oder im Vergleich zu anderen Systemen deutlich verlässlichere Zielaufklärung und Waffenkontrolle erreicht. Dies ist die technische Voraussetzung für verantworteten Waffeneinsatz.

Alle Planungsdokumente der Bundeswehr schreiben vor, daß jeder Waffeneinsatz von Menschen zu verantworten ist: „Für unbemannte Luftfahrzeuge ist das Prinzip des *Human-In-The-Loop* und damit die verzuglose Möglichkeit zum Bediene- ▶▶

reingriff jederzeit sicherzustellen. Es ist und bleibt die Linie der Bundesregierung, daß ein Waffeneinsatz von unbemannten Luftfahrzeugen ausschließlich unter Kontrolle des Menschen und nur in dafür durch den Bundestag mandatierten Einsätzen erfolgt.<sup>40</sup> Es besteht kein Grund zur Annahme, eine vom Deutschen Bundestag kontrollierte Parlamentsarmee verfolge bei dieser sensiblen Frage eine *hidden agenda*. Dies gilt jedoch nicht für alle Staaten.

Gemäß der Eigengesetzlichkeit wehrtechnischer Entwicklungen ist mit der Bedrohung eigener Kräfte durch *Lethal Autonomous Weapon Systems*, LAWS, zu rechnen, die Ziele automatisch auswählen und bekämpfen. Daher muß die Abwehr dieser Bedrohung untersucht werden. Ein Beispiel ist Drohnenabwehr, gerade auch im Kontext der öffentlichen Sicherheit.<sup>41</sup> Hochreaktiver Schutz ist als rein reaktive Abwehr kein automatisches Angriffssystem. Und auch hier ist der Mensch eingebunden, nämlich mit der Entscheidung, das System einzuschalten oder den Parameterrahmen zur Freigabe selbstständiger Abwehr zu wählen.

Entscheidungen zum Waffeneinsatz erfolgen gemäß den *Rules-of-En-*



»Es entstehen Gefahren aus der psychologischen „Maschinisierung des Menschen“, seiner Behandlung als „Biomachine im Sinne der Reiz-Antwort-Maschine eines kybernetischen Menschenbilds.«

einsatzspezifischen Handlungsrahmen fest. Gemäß rechtlichen, politischen, militärstrategischen sowie operativen Vorgaben konkretisieren sie die völkerrechtlichen Prinzipien des *ius in bello* und des internationalen *soft law*.

Beispiele sind *Diskrimination* (Waffeneinsatz nur bei lückenloser Zielerfassung) oder *Proportionalität* (Wahl bedrohungsadäquater Waffen). Darin eingeschlossen ist mögliche Vermeidung von Kollateralschäden.



Die Katholikin Yvonne Hofstetter (1966) ist eine deutsche Juristin und Sachbuchautorin. 2018 erhielt sie den Theodor Heuss-Preis für ihr empfehlenswertes Buch „Das Ende der Demokratie“.

*agement*, RoE, den Einsatzregeln, die bis in das informationstechnische Design zu berücksichtigen sind. Dabei machen sie keine taktischen Vorgaben, sondern legen rechtlich bindend den

Zu nennen sind auch *Vorsorge* (Vorrang der ungünstigen Prognose gegenüber der günstigen) und *Zurechenbarkeit*, wodurch *meaningful human control* impliziert wird. Schließlich

wäre *Rechtssicherheit* durch transparente Handlungsoptionen und Dokumentation zu nennen.

Offenbar stellen sich bei autonomem Fahren analoge Fragen. Ein rechtlich bindender Rahmen ist die Straßenverkehrsordnung. Auch hier stellt die urbane Umgebung besondere Anforderungen an die lückenlose Erfassung des Verkehrsgeschehens, wobei die Wahl bedrohungsadäquater Wirkmittel den Lenk- und Bremsentscheidungen entspricht. Dem Prinzip der *Vorsorge* ist zu genügen, wenn sich bei einem Unfall die Alternative zwischen der Gefährdung der Passanten oder Fahrzeuginsassen stellt. Im Verkehr wird man wohl noch lange nicht auf *meaningful human control* verzichten. Vergleichbare Probleme entstehen bei ziviler Drohnenutzung, Industrie 4.0, *consumer robots*, Finanzmarkt-Bots oder *smart cities*.

Mut macht die Tatsache, daß FCAS als erstes wehrtechnisches Großprojekt von Anfang an durch einen kompetent und hochrangig besetzten Ethikbeirat begleitet wird. So wird konkretisiert, wie ethische und rechtliche Prinzipien sowie der gesellschaftliche *sensus communis* zu technischen Designprinzipien werden, die bis zuletzt verantwortetes Wirken im Sinne von *human command & machine control* ermöglichen. Möge dies auf andere Rüstungsprojekte und zivile Anwendungen ausstrahlen.<sup>42</sup>

## KI und natürliche Dummheit

Wir haben überlegt, wie Künstliche Intelligenz technisch beherrschbar und verantwortlich nutzbar bleiben kann. Was aber geschieht, wenn wir zulassen, daß KI-basierte Systeme Einfluß auf uns selbst nehmen, weil wir unsere Verantwortung abgeben, aufhören, „Herren der KI“ zu sein? Auch wenn wir bestreiten, daß selbstbewußte Intelligenz und Autonomie aus Künstlicher Intelligenz und Technischer Autonomie „fulgurieren“, entstehen Gefahren aus der psychologischen „Maschinisierung des Menschen“, seiner Behandlung als „Biomachine im Sinne der Reiz-Antwort-Maschine eines kybernetischen Menschenbilds“.<sup>43</sup>

Aufgrund der Komplexität dieser Problematik beschränken wir uns darauf, Gefährdungen der persönlichen und politischen Freiheit durch ideologisch instrumentalisierte KI zu

»Was geschieht, wenn wir zulassen, daß KI-basierte Systeme Einfluß auf uns selbst nehmen, **weil wir unsere Verantwortung abgeben**, aufhören, „Herren der KI“ zu sein?«

betrachten. Zur Verleihung des Theodor-Heuss-Preises an sie formuliert die IT-Unternehmerin, Juristin und politische Publizistin Yvonne Hofstetter dazu drei prägnante Thesen.

1. *Algorithmische Personalisierung lähmt die politische Freiheit.*<sup>44</sup> Wer die Angebote des Internets nutzt, um sich zu informieren oder zu unterhalten, wer seine Einkäufe über Online-Plattformen tätigt oder durch *social media* mit seinem Familien-, Freundes- oder Bekanntenkreis Nachrichten oder Fotos austauscht, wer *location-based services* nutzt, um selbst zu wissen, wo er ist, dies anderen mitzuteilen, zu erfahren, wo andere sind oder wie er einen bestimmten Ort erreicht, wundert sich, wie gut ihn das Netz „kennt“. Denn es paßt sich seinen Bedürfnissen, seinem Profil an. Und meist sind es ja wirklich gute Ideen, das vorgeschlagene Buch, das Restaurant in unserer Nähe oder das Kleidungsstück, das zu uns paßt.

Was steht dahinter? Warum werden uns diese Services kostenfrei angeboten? „Die für Nutzer unsichtbare funktionale Schicht sozialer Medien sind Börsenhandelsplätze, die mit der Aufmerksamkeit ihrer Nutzer handeln“, erläutert die Theodor-Heuss-Preisträgerin. „Sie leiten die von uns freiwillig eingespeisten Daten unseres Lebens in Sekunden als Offerte weiter: Hier unser Angebot für die Aufmerksamkeit von Nutzer N. N., der Preis beträgt so und so viel: Deal.“ Darum geht es also: „Das Narrativ von der Personalisierung ist nichts als Ideologie und die Basis für einen Hyperkapitalismus, symbolisiert durch Totalkommunikation, die immer mehr mit der Totalüberwachung und Totalausbeutung zusammenfällt.“

Aber die auf diese Weise entstehende Überwachungs- und Manipulationsinfrastruktur ist nur der erste Schritt in eine un gute Richtung, die nicht nur

uns, sondern unser Rechtssystem verändern kann:

2. *Aus der Objektivierung der Person zum Datenhaufen entsteht der Neue Mensch.*<sup>45</sup> „Freiheit, wie wir Europäer sie meinen, beruht auf einem bestimmten Menschenbild, eben jener Vorstellung von der vernunftbegabten Person, die ‚selbst denken‘ und auch selbstverantwortlich entscheiden kann“, erläutert die Juristin und zeigt, welche Konsequenzen sich daraus ergeben können: „Wird der Mensch nach dem Modell eines steuerbaren Algorithmus erklärt und dauerhaft so behandelt, verändert das neue Verständnis auch den Menschen selbst. Selbstwahrnehmung, Wertesystem, Zusammenleben bleiben nicht ohne Folgen für die künftige Gestaltung unserer Gesellschaft durch das Rechtssystem.“

„Nichts ist so entscheidend für den Stil eines Rechtszeitalters wie die Auffassung vom Menschen, an der es sich orientiert“, zitiert der Jurist Friedrich von Westphalen den großen Rechtsphilosophen Gustav Radbruch. Seine Schlußfolgerung macht nachdenklich: „Eine Gesellschaft, die soziologisch normiert, dass der Mensch auf das naturwissenschaftliche Messbare – eben auf Big Data –, auf das empirisch Beobachtbare reduziert wird, muss und wird dieses Menschenbild früher oder später auch im hoheitlich gesetzten Recht manifestieren. Denn Recht ist nichts anderes als ‚geronnene Politik.‘“<sup>46</sup>

Die „Objektivierung der Person zum Datenhaufen“ steht jedenfalls mit der geltenden Rechtsprechung des Bundesverfassungsgerichts nicht in Einklang: „Mit der Menschenwürde wäre es nicht zu vereinbaren, den Menschen in seiner ganzen Persönlichkeit zu registrieren und zu katalogisieren, sei es auch in der Anonymität einer statistischen Erhebung, und ihn damit wie eine Sache zu behandeln,



»Es gibt für den Menschen Normen, die aus dem Wesen und Sein Gottes selber fließen und daher unverbrüchlich sind und nicht angetastet werden dürfen. Wenn man das Bestehen solcher Normen nicht anerkennt, dann gleitet ein Volk abwärts in Diktatur und Gewalt.« (Konrad Adenauer) : Bild: Fronleichnam 1962 in Bonn.

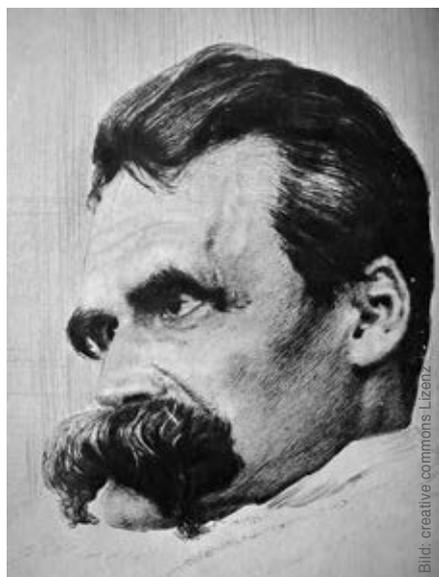
die einer Bestandsaufnahme in jeder Beziehung zugänglich ist.“<sup>47</sup> Die massenhafte „Erhebung“ und „Bestandsaufnahme“ persönlichster Daten durch transnationale Technologiegiganten, die sich gerichtlicher oder politischer Kontrolle weitgehend entziehen, ist demnach grundgesetzwidrig.

### Totalitäre Gefahren

Nach einer entsetzlichen Lektion über den Totalitarismus und zu Beginn einer ungeheuren Wiederaufbauleistung, gerade auch auf moralischer und geistiger Ebene, formuliert Konrad Adenauer die Grundlagen einer freiheitlichen Demokratie: „Wir nennen uns christliche Demokraten, weil wir der tiefen Überzeugung sind, dass nur eine Demokratie, die in der christlich-abendländischen Weltanschauung, in dem christlichen Naturrecht, in den Grundsätzen der christlichen Ethik wurzelt, die große erzieherische Aufgabe am deutschen Volk erfüllen und seinen ►►

Wiederaufstieg herbeiführen kann.“ Auch nach seiner Kanzlerschaft mahnt Adenauer: „Unsere Partei – ich wiederhole es – steht fest und unverbrüchlich auf dem Boden dieser, der christlichen Weltanschauung, dass es für den Menschen Normen gibt, die aus dem Wesen und Sein Gottes selber fließen und daher unverbrüchlich sind und nicht angetastet werden dürfen. Wenn man das Bestehen solcher Normen nicht anerkennt, dann gleitet ein Volk abwärts in Diktatur und Gewalt.“<sup>48</sup>

Wer rechtsphilosophischem Denken nachgeht, das Adenauers Mahnung ignoriert oder explizit ablehnt, nimmt einigermaßen verstört zur Kenntnis, welche scheinbar selbstverständlichen Grundbegriffe in Frage



»Offenbar teilen Nietzsche und die Transhumanisten viele Aspekte in ihren Anthropologien und Werten. [...] Ihr Konzept des Transhumanen weist viele signifikante Ähnlichkeiten mit Nietzsches ‚Übermensch‘ auf.« (Radierung nach der Fotoserie Der kranke Nietzsche, 1899)

stehen: „Offensichtlich ist das Konzept des ‚menschlichen Handelns‘ ein Paradebeispiel für ein im Wesentlichen umstrittenes Konzept, was bedeutet, dass es keinen Konsens über die Bedeutung dieses Begriffs zwischen seinen verschiedenen Nutzern gibt“, faßt ein rechtsphilosophisches Werk die Diskussion über sich den auflösenden Begriff des verantwortlichen Rechtssubjekts zusammen.<sup>49</sup>

Insbesondere zielt poststrukturalistisches Denken in diese Richtung. So

scheint Michel Foucault der Vorstellung des menschlichen Handelns keine Bedeutung mehr beimessen zu können und spricht vom „Tod des Subjekts“, wie Nietzsche vom „Tod Gottes“: „Man braucht sich nicht sonderlich über das Ende des Menschen aufzuregen; das ist nur ein Sonderfall oder, wenn Sie so wollen, eine der sichtbaren Formen eines weitaus allgemeineren Sterbens. Damit meine ich nicht den Tod Gottes, sondern den Tod des Subjekts, des Subjekts als Ursprung und Grundlage des Wissens, der Freiheit, der Sprache und der Geschichte.“<sup>50</sup> Wenn der Mensch nach Gottes Ebenbild geschaffen ist, Gott aber für tot erklärt wird, sind das Menschenbild einer biokybernetischen Reiz-Antwort-Maschine und der Tod des Menschen als Subjekt nur folgerichtig. Was folgt daraus auf politischer Ebene?

3. *Die Fragmentierung und Schwächung des Souveräns ist Nährboden für Tyrannei und für eine neue – nicht notwendigerweise demokratische Gesellschaft*,<sup>51</sup> formuliert Yvonne Hofstetter als dritte These. „Empirisch beobachtbar ist, dass sich um Narrative und Kampagnen, die in sozialen Netzwerken lanciert werden, eine Aggregation von Anhängern, Freunden, Followers bildet mit dem Potenzial, Schneeballeffekte auszulösen.“ Dadurch gehe die Wahrheit, das Verständnis einer gemeinsamen Wirklichkeit verloren, ohne die in einer Gesellschaft politische Meinungsbildung unmöglich werde.

Hannah Arendt habe erkannt, wie sehr politische Freiheit Wahrheit voraussetzt: „Solange die Presse frei und nicht korrupt ist, hat sie eine ungeheuer wichtige Aufgabe zu erfüllen und kann mit Recht die vierte öffentliche Gewalt genannt werden“. Das Recht auf nicht manipulierte Tatsacheninformation, hält sie für die „wesentlichste politische Freiheit“, ohne welche „die ganze Meinungsfreiheit zu einem entsetzlichen Schwindel wird“.<sup>52</sup>

Für Katholiken, die inmitten dieser programmatischen Angriffe auf die Freiheit leben, ist es schmerzlich, wenn ihre Hirten programmatisch die Irrelevanz der Wahrheit zu proklamieren scheinen: „Im Konzil ist geradezu eine Art Kopernikanischer Wende erfolgt: Die Wahrheit wird nicht geleugnet, tritt aber in der Hintergrund“,

erklärt Karl Kardinal Lehmann an der Heinrich-Heine-Universität. „Ich kenne keine Beispiele, wo eine Lehrentwicklung solche Differenzen in sich birgt“.<sup>53</sup>

## Human Enhancement Engineering

Es bleibt ein Mythos, daß selbstbewußte, intelligent und autonom handelnde Wesen aus Künstlicher Intelligenz und Technischer Autonomie hervorgehen und den Menschen verdrängen. Dennoch bergen diese Technologien nicht nur psychologische Gefahren, sondern gefährden den Menschen als solchen, wenn sie zu Instrumenten transhumanistischer Ideologen werden, die *human nature* als *work-in-progress* betrachten.<sup>54</sup>

Wie die Bildersuche im Internet spiegeln auch *Comics* und Filme das kollektive Unbewußte einer Zeit, die sich dieser Versuchung zu öffnen scheint. Zugleich prägen sie dieses Unbewußte, so daß es aufschlußreich ist, über die Wellen der *Super Heroes* nachzudenken, Wesen mit übermenschlichen Fähigkeiten und High-Tech-Ausrüstung, die immer wieder die Kinos und Bildschirme überspülen. Sie bekämpfen Böses, meist Superschurken. In der Regel werden sie edel gezeichnet, es gibt jedoch auch ambivalente Figuren. Am Anfang der gegenwärtigen Welle steht der Zeichentrickfilm *The Incredibles* (2004), der an *Comics* anknüpft, die ab 1956 erfolgreich werden, ausgerechnet im Geburtsjahr der KI also.

Die merkwürdige Koinzidenz der Konjunkturen, die KI und *Super Heroes* zum Faszinosum machen, läßt nach gemeinsamen gedanklichen Wurzeln fragen. Eine Studie deckt Friedrich Nietzsches Vorstellung vom „Übermensch“ auf.<sup>55</sup> Aber auch KI ist von Beginn an nicht allein eine Spezialdisziplin der Informatik, sondern will innerweltliche Erlösungssehnsüchte stillen. Marvin Minsky, einer der Initiatoren der *Dartmouth Conference 1956*, formuliert sie klar: „Das Ziel der KI ist die Überwindung des Todes“<sup>56</sup>, und liegt mit Konrad Lorenz auf einer gedanklichen Linie.

2006 erscheint in der Fachzeitschrift *Applied Ergonomics* ein vielbeachteter Aufsatz über Revolutionen und sich wandelnde Paradigmen

»Im Konzil ist geradezu **eine Art Kopernikanischer Wende** erfolgt: Die Wahrheit wird nicht geleugnet, tritt aber in der Hintergrund.«

(Karl Kardinal Lehmann)



»Wie in Aldous Huxleys (1894 – 1963) „Schöner Neuer Welt“, die sich durch Drogen stabilisiert, die „alle Vorteile des Christentums und des Alkohols besitzen, aber keines ihrer Defekte“, kennt das moderne Berufsleben längst mentales Doping.«

der Mensch-Maschine-Interaktion.<sup>57</sup> Der Autor ist Kenneth R. Boff, *Chief Scientist* des *Human Effectiveness Directorate* am *US Air Force Research Laboratory*. Die Revolution der Informationstechnologie habe „eine ethische Kluft geöffnet zwischen denen, welche die kognitive und physische Steigerung des Menschen als großen Nutzen für die Gesellschaft sehen, und denen, die sie als Manipulation der Grundlagen der menschlichen Natur wahrnehmen“. Worum geht es?

Wie beschrieben, ermöglichen Künstliche Intelligenz und Technische Autonomie „kognitive Maschinen“, die unsere Fähigkeiten des Wahrnehmens und Wirkens weit über natürliches Maß hinaus steigern. Boff betrachtet die *Revolution in Information Technology* allerdings unter der Perspektive des *Human Enhancement Engineering*, also nicht mehr nur als Mittel zur Weltgestaltung, sondern als Kräfte, die uns nicht nur psychologisch beeinflussen, sondern uns selbst verändern können. Er identifiziert vier Generationen:

1. *Physical Fit*: Anpassung von Geräten, Arbeitsplätzen und Aufgaben an den Menschen und seine Grenzen. 2. *Cognitive Fit*: Harmonische Integrati-

on der Technik, des Menschen und seiner Aufgaben zu effektiven Systemen. 3. *Neural Fit*: Steigerung physischer und kognitiver Fähigkeiten des Menschen durch symbiotische Kopplung mit Technologie. 4. *Biological Fit*: Biologische Veränderung der physischen und kognitiven Fähigkeiten zur Maximierung der menschlichen Effektivität. Es ist aufschlußreich, vor allem der dritten und vierten Generation nachzugehen und Boffs Erwartungen, über die wir 2005 mit ihm selbst beim abendlichen Rotwein diskutiert haben, nach 13 Jahren mit aktuellen Trends zu vergleichen. Wir kommentieren einige Zitate aus Boffs Aufsatz.

1. „Die kognitiven und physischen Fähigkeiten des Menschen können weit über die reguläre biologische Variation hinaus gesteigert werden, wodurch sich traditionelle Grenzen für die Anpassungsfähigkeit des Menschen in komplexen technischen Systemen verändern“. Boffs Prognose ist eingetroffen.

2. Die US-amerikanische Forschung zielt darauf ab, „menschliche kognitive Fähigkeiten durch die Vereinigung von Menschen und Computersystemen zu maximieren“ – unser Thema hier. Darüber hinaus können schon

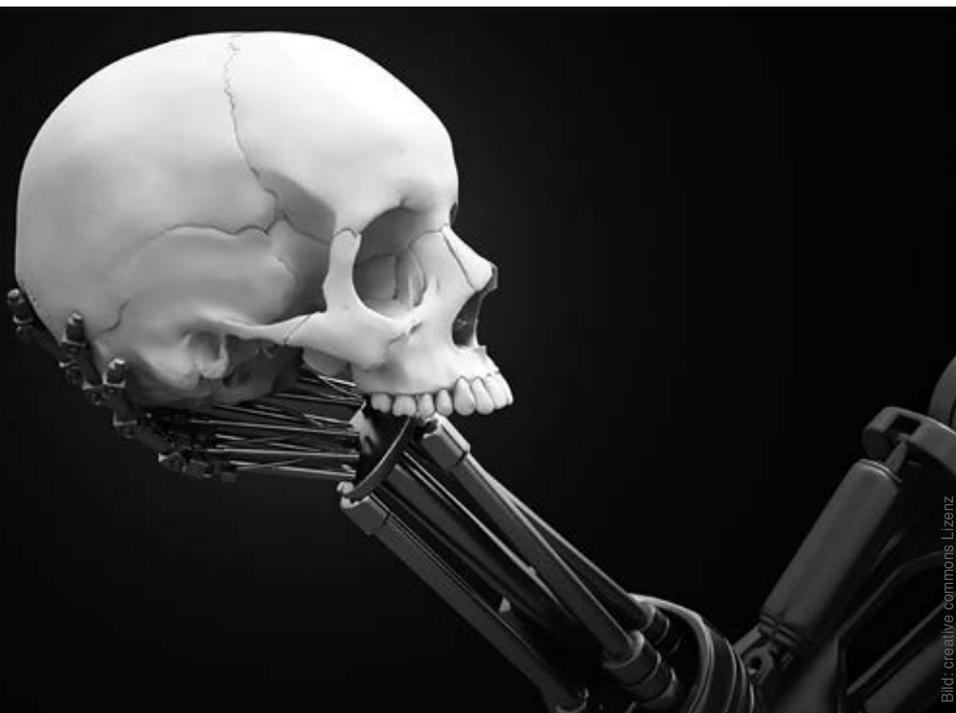
heute Menschen durch Gehirnaktivität, über EEG-Elektroden aufgenommen, technische Systeme rein mental nach entsprechendem Training steuern. Eine unmittelbare Kopplung des menschlichen Gehirns mit Computersystemen scheint in Reichweite.<sup>58</sup>

3. Neben der prothetischen Wiederherstellung beeinträchtigter kognitiver Funktionen sei „das Rennen nach dem nächsten ‚Viagra‘ eröffnet. Allerdings wird dieses Potenzmittel auf das Gehirn wirken und normal begabte Menschen in einen *hyper-range* kognitiver Fähigkeiten katapultieren“. Wie in Aldous Huxleys „Schöner Neuer Welt“, die sich durch Drogen stabilisiert, die „alle Vorteile des Christentums und des Alkohols besitzen, aber keines ihrer Defekte“, kennt das moderne Berufsleben längst mentales Doping. Die philosophische Debatte über „Neuroenhancement“ ist entbrannt.<sup>59</sup>

4. „Der Markt für kognitives Enhancement ist extrem groß und wird zu einem neuen Zweig der Medizin führen, ‚Kosmetische Neurologie‘“. Angesichts des ökonomischen Gewichts der Schönheitschirurgie ist Boffs Vision für „Alpha-plus-Intellektuelle“ gemäß Huxleys Taxonomie nicht unrealistisch.

### Transhumane Versuchungen

Als Kunstwerke sprechen manche Science-Fiction-Filme philosophische und theologische Grundfragen an, die jede Menschenseele bewegen. Daher schärfen ihre Visionen den Blick ►►



»Das Ziel der KI ist die Überwindung des Todes.« (Der Transhumanist Marvin Minsky)

auf sich anbahnende Entwicklungen und stählen den Willen, Weichen in die rechte Richtung zu stellen oder wenigstens dazu beizutragen. Denn technischer Fortschritt ist kein unausweichlicher Prozeß, sondern wird durch freie Entscheidungen von Menschen gestaltet, die für ihr Handeln verantwortlich sind.

Ein solcher Film thematisiert bereits 1982 *Human Enhancement Engineering*. Die Hauptfigur in Ridley Scotts *Blade Runner* ist Roy Batty, ein biogenetisch über Menschenmaß gesteigertes Wesen im Sinne von Boffis *4th Generation of Human Factors and Ergonomics*.<sup>60</sup> Gemäß seiner Funktionsdefinition *combat, colonization defense program* wird dieser *replicant* wie seine Mitwesen als Sklave in den Weltraumkolonien einer künftigen Menschheit ausgebeutet. Die Rückkehr zur Erde ist ihnen verboten. Im November des Jahres 2019 der Filmhandlung kommt er auf der Suche nach seinem „Schöpfer“, dem genialen Biogenetik-Unternehmer Eldon Tyrell, dennoch zur Erde und wird gejagt.

Wie bei Lorenz und Minsky geht es eigentlich um den Griff nach dem „Baum des Lebens“, nachdem die Frucht vom „Baum der Erkenntnis“ konsumiert ist. Denn aus Sicherheitsgründen besitzen Replikanten eine nur 4-jährige Lebenserwartung. Als Roy Batty seinen „Schöpfer“ findet und als

*deus malignus* erkennt, tötet er Edlon Tyrell, der an High-Tech-Unternehmer der Gegenwart denken läßt.<sup>61</sup> Ridley Scott läßt keinen Zweifel daran, daß er diese Szene als Gottesmord versteht. Die Schlußszene *Tears in the rain* ist eine der ergreifendsten Sequenzen der Filmgeschichte.<sup>62</sup> Auf seine Art variiert sie das Lebensgefühl des „Tollen Menschen“ bei Nietzsche:

„Wohin ist Gott? rief er, ich will es euch sagen! Wir haben ihn getötet, – ihr und ich! Wir Alle sind seine Mörder! Aber wie haben wir dies gemacht? Wie vermochten wir das Meer auszutrinken? Wer gab uns den Schwamm, um den ganzen Horizont wegzuwischen? Was taten wir, als wir diese Erde von ihrer Sonne losketteten? Wohin bewegt sie sich nun? Wohin bewegen wir uns? Fort von allen Sonnen? [...] Gott ist tot! Gott bleibt tot! Und wir haben ihn getötet! Wie trösten wir uns, die Mörder aller Mörder?“<sup>63</sup>

Als *deus malignus* sehen sich anscheinend die Transhumanisten, die offen bekennen: „Transhumanisten betrachten die menschliche Natur als *work-in-progress*“.<sup>64</sup> Orientierung bietet ihnen Nietzsche: „Offenbar teilen Nietzsche und die Transhumanisten viele Aspekte in ihren Anthropologien und Werten. [...] Ihr Konzept des Transhumanen weist viele signifikante Ähnlichkeiten mit Nietzsches ‚Übermensch‘ auf“.<sup>65</sup> Den Atem hält an,

wer sich klar macht, aus welcher Position der Historiker Yuval Harari, in Millionenaufgabe und in 24 Sprachen übersetzt, glaubt, sprechen zu können: „Und nachdem wir die Menschheit über die animalische Ebene des Überlebenskampfes hinausgehoben haben, werden wir nun danach streben, Menschen in Götter zu verwandeln und aus dem *Homo sapiens* den *Homo deus* zu machen.“<sup>66</sup>

Wer Boffis nüchterne Prognosen mit dem transhumanen Programm zusammendenkt, stößt auf eine Haltung, wie sie Nietzsche beschreibt: „Hybris ist heute unsre ganze Stellung zur Natur, unsre Natur-Vergewaltigung mit Hilfe der Maschinen und der so unbedenklichen Techniker- und Ingenieur-Erfindsamkeit; Hybris ist unsre Stellung zu Gott. [...] Hybris ist unsre Stellung zu uns, – denn wir experimentieren mit uns, wie wir es uns mit keinem Tiere erlauben würden, und schlitzen uns vergnügt und neugierig die Seele bei lebendigem Leibe auf: was liegt uns noch am »Heil« der Seele!“<sup>67</sup>

## Die erste und letzte Sünde

„Welche Idee wird die gefährlichste Idee der Welt?“, fragt sich der Politikwissenschaftler Francis Fukuyama und nennt den Transhumanismus „eine seltsame Befreiungsbewegung, deren Kreuzzügler weitaus höhere Ziel anstreben als Bürgerrechtler, Feministinnen oder Gender-Aktivist\*innen. Sie möchte nicht weniger, als die menschliche Rasse aus ihren biologischen Grenzen befreien.“<sup>68</sup> „Francis Fukuyama glaubt das“, antwortet ihm der Transhumanist Nick Bostrom, aber seine Antwort beunruhigt: „Wirklich gefährlich wird der Transhumanismus nur für einen reaktionären Biokonservativismus.“<sup>69</sup>

Fällt eine leistungsfähige Technologie in die Hände von Techno-Ideologen mit seltsamen Vorstellungen über die Zukunft der menschlichen Rasse, wird es in der Tat gefährlich. Wenigstens Deutsche wissen, was dann auch mit Kulturnationen geschehen kann. Vielleicht gewinnen Gardinis Dämonen durch unsere „scheinbar natürlichen, in Wahrheit so widersprüchlichen Instinkte“, durch unsere „scheinbar folgerichtige, in Wahrheit so leicht beeinflussbare Logik“, durch unsere „unter aller Gewalttätigkeit so hilflose Selbstsucht“ die Macht, unse-

»Was heute vielen Menschen als borniertes Festhalten der Kirche an traditionellen Verhaltensmustern erscheint, **muss in einem neuen Lichte gesehen werden.**« (Robert Spaemann)



»Was heute vielen Menschen als borniertes Festhalten der Kirche an traditionellen Verhaltensmustern erscheint, muss in einem neuen Lichte gesehen werden: als Widerstand gegen das, was C. S. Lewis ‚die Abschaffung des Menschen‘, die abolition of man, genannt hat.« (Robert Spaemann)

re ganze Gesellschaft zu einem Experiment mit uns selbst zu verführen, „wie wir es uns mit keinem Tier erlauben würden“.

Radikal widersprechen transhumanistische Auffassungen einer Anthropologie, die den Menschen in seiner

leib-seelischen Polarität und Begrenztheit auch in seiner Verwundung als „gut geschaffen“ und von Gott geliebt annimmt. Die Hysterie, mit der über selbstbewußte und „übermenschliche“ KI spekuliert wird, lenkt von einer realen Gefahr ab, der Instrumentalisierung von KI für *human enhancement*.

Um besser zu verstehen, was Transhumanismus letzten Endes und im Grunde ist, sollte man „den tiefen Zusammenhang des Menschen mit Gott beachten“, wie es im Katechismus heißt. „In einem Zustand der Heiligkeit erschaffen, war der Mensch dazu bestimmt, von Gott in der Herrlichkeit völlig ‚vergöttlicht‘ zu werden.“ Die erste Sünde wendet sich gegen diese Bestimmung: „Vom Teufel versucht, wollte er ‚wie Gott sein‘, aber ohne Gott und vor Gott und nicht Gott gemäß.“<sup>70</sup> So betrachtet, entspräche Transhumanismus der ersten Sünde, die sich anschickt, auch nach dem „Baum des Lebens“ zu greifen, und zur letzten Sünde wird.

„Was heute vielen Menschen als borniertes Festhalten der Kirche an traditionellen Verhaltensmustern erscheint, muss in einem neuen Lichte gesehen werden: als Widerstand gegen das, was C. S. Lewis ‚die Abschaffung des Menschen‘, die *abolition of man*, genannt hat“, kommentiert Ro-

bert Spaemann weitsichtig eine Entwicklung, die schon lange absehbar ist und sich nun realisieren könnte: „Die szientistische Zivilisation mit ihrer Tendenz zum Spiritualismus und Materialismus, zur Dekomposition der menschlichen Natur ist die Tendenz zu dieser Abschaffung. Wenn Europa nicht die kostbare Perle wiederfindet, die seine Mitte war, wird es zum Ort, von dem die Abschaffung des Menschen auf diesem Planeten ausgeht.“<sup>71</sup>

In diesen Auseinandersetzungen, die sich weiter verschärfen werden, muß vor allem Christen bewußt bleiben, was Künstliche Intelligenz und Technische Autonomie NICHT sind: Künstliche Dinge bestehen aus einem Woraus und einem Was. Ihr Wie und Wozu liegt NICHT in ihnen, sondern im Menschen. Daher wäre insbesondere ‚Sprachverschmutzung‘ durch unangemessene Anthropomorphismen wie „Intelligenz“, „Autonomie“ oder „Lernen“ zu vermeiden. Durch *overpromising* schadet dies auf lange Sicht der Vertrauenswürdigkeit eines in sich überaus nützlichen Zweigs der Informatik und vernebelt unser Denken. Denn angemessene Sprache ist nicht nur ein Kommunikationsmittel, sondern das Medium unseres Denkens schlechthin. ■

## Anmerkungen

1 K. Lorenz (1977). *Die Rückseite des Spiegels. Versuch einer Naturgeschichte menschlichen Erkennens*. München, Zürich 1975, S. 48ff.

2 *Die Rückseite des Spiegels*, S. 321.

3 BBC, 2.12.2014. <https://www.bbc.com/news/technology-30290540>.

4 SXSW 2018, 11.3.2018. [https://www.youtube.com/watch?time\\_continue=7&v=kz-IUyrcbos](https://www.youtube.com/watch?time_continue=7&v=kz-IUyrcbos).

5 <https://www.cdu.de/system/>

[tdf/media/dokumente/koalitionsvertrag\\_2018.pdf?file=1](tdf/media/dokumente/koalitionsvertrag_2018.pdf?file=1).

6 VDI 3780: Technikbewertung. Begriffe und Grundlagen. September 2000, S. 2

7 *Physik*. Übers. von Christian Hermann Weiße. II.1.

8 C. F. von Weizsäcker (1967). *Möglichkeit und Bewegung. Eine Notiz zur aristotelischen Physik*. In: ders. *Die Einheit der Natur*. München <sup>2</sup>1981, S. 432.

9 *Physik*. III.3.

10 R. Spaemann, R. Löw (2005). *Natürliche Ziele. Geschichte und Wiederentdeckung*

*des teleologischen Denkens*. Stuttgart, S. 51f.

11 *Ergo natura nihil aliud est quam principium motus et quietis in eo in quo est primo et per se et non secundum accidens*. In *phys. lib. 2 l. 1 n. 5*. [www.corpusthomicum.org](http://www.corpusthomicum.org).

12 J. Pieper (1929). *Einleitung zu Rerum novarum*. In: *Werke*. EB 1, S. 7.

13 J. Pieper (1977). *Über den Begriff der Sünde*. In: *Werke*. Bd. 5, S. 233.

14 W. Koch (2014). *Tracking and Sensor Data Fusion. Me-*

*thodological Framework and Selected Applications*. Springer.

15 F. F. Weyh (2014). *Intelligenter Staub. Sensornetzwerke durchdringen unsere Umwelt*. Deutschlandfunk Kultur. 8.5.2014, [https://www.deutschlandfunkkultur.de/technik-intelligenter-staub.1088.de.html?dram:article\\_id=284823](https://www.deutschlandfunkkultur.de/technik-intelligenter-staub.1088.de.html?dram:article_id=284823).

16 N. Wiener (1948). *Kybernetik. Regelung und Nachrichtenübertragung in Lebewesen und Maschine*. München 1968.

17 Programmiersprache LISP, Turing Award 1971.

- 18 Neuronale Netze: SNARC (Stochastik NeuralAnalog Reinforcement Computer).
- 19 Konstrukteur der IBM 701, erste Assembler-Sprache.
- 20 <http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf>.
- 21 R. Guardini (1950). *Das Ende der Neuzeit*. Ostfildern, Paderborn 1986, S. 51.
- 22 *Ende der Neuzeit*, S. 71.
- 23 *Robotics, AI, and Humanity: Science, Ethics, and Policy*. 16-17 May 2019 | Casina Pio IV | Vatican City. [http://www.pas.va/content/dam/accademia/booklet/booklet\\_robotics.pdf](http://www.pas.va/content/dam/accademia/booklet/booklet_robotics.pdf).
- 24 *Ende der Neuzeit*, S. 70.
- 25 <https://www.katholisch.de/aktuelles/aktuelle-artikel/selig-sprechungsprozesse-fur-gerlich-und-guardini-eroffnet>.
- Ende der Neuzeit*, S. 86.
- 27 *Katechismus der Katholischen Kirche* (1997), Nr. 407.
- 28 FA.I.22, S. 93.
- 29 I. Goodfellow et al., "Explaining and Harnessing Adversarial Examples", Int. Conf. on Learning Representations, 2015.
- 30 N. Papernot et al., "Practical Black-Box Attacks against Machine Learning", 2017 ACM Conf. on Computer and Communications Security.
- 31 D. Gershgorin (2016). *Fooling the machine*. 30.3.2016. <https://www.popsci.com/byzantine-science-deceiving-artificial-intelligence>.
- 32 S. B. McGrayne (2012). *The Theory That Would Not Die*. Yale University Press.
- 33 J. Pearl (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books.
- 34 Th. Nagel (1974). *What Is it Like to Be a Bat? Wie ist es, eine Fledermaus zu sein?* Übers. von Ulrich Diehl. Reclam 2016.
- 35 *There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – there are things we do not know we don't know.* Donald Rumsfeld, 12.2.2002.
- 36 Putin ergänzt, daß es nicht wünschenswert sei, daß jemand in diesem Bereich eine Monopolstellung erreiche. Sollte Rußland ein solcher Durchbruch gelingen, werde sein Land diese Technik aber mit dem Rest der Welt teilen, „so wie es heute bereits mit der Nukleartechnik der Fall ist“. TASS, 1.9.2017. <https://tass.com/society/963209>.
- 37 *Automatische Augmented Gefechtsfeldaufklärung (Auto-Auge)*. [https://www.fkie.fraunhofer.de/de/Veranstaltungen/abschlussdemo\\_autoauge.html](https://www.fkie.fraunhofer.de/de/Veranstaltungen/abschlussdemo_autoauge.html).
- 38 Aktuelle Stunde am 2. Juli 2014. [http://www.bundestag.de/dokumente/textarchiv/2014/kw27\\_de\\_drohnen/286034](http://www.bundestag.de/dokumente/textarchiv/2014/kw27_de_drohnen/286034).
- 39 *Aufklärungsunterstützung aus der Luft – Agile UAV in vernetzter Umgebung*. <https://www.fkie.fraunhofer.de/de/forschungsabteilungen/sdf/AgileUAV.html#tabpanel-1>.
- 40 <https://www.bmvg.de/de/aktuelles/verteidigungsministerium-veroeffentlicht-militaerische-luftfahrtstrategie-11506>.
- 41 W. Koch (2016). *Multisensorielle Drohnenabwehr. Forschungsaspekte einer Herausforderung*. Behördenspiegel, Mai 2016. [https://issuu.com/behoerden\\_spiegel/docs/2016\\_mai/56](https://issuu.com/behoerden_spiegel/docs/2016_mai/56).
- 42 Y. Hofstetter, W. Koch, F. v. Westphalen (2019). *Autonome Waffen. Das 5. Gebot im KI-Krieg*. Spektrum der Wissenschaft, .6.7.2019. <https://www.spektrum.de/kolumne/der-krieg-der-zukunft-wird-dank-robotern-und-kuenstlicher-intelligenz-ein-problem/1655406>.
- 43 Y. Hofstetter (2018). *Programmierte Freiheit – Spielräume für Verantwortung: Dank*. 16.06.2018, S. 2. <https://www.theodor-heuss-stiftung.de/wp-content/uploads/pv-2018-yvonne-hofstetter-ths-dank-redeformat-1.pdf>.
- 44 *Programmierte Freiheit*, S. 6.
- 45 *Programmierte Freiheit*, S. 9.
- 46 F. v. Westphalen, Y. Hofstetter (2017). *Der drohende Verlust der Privatautonomie des Verbrauchers. Wie Big Data die Welt verändert – und was das für das Menschenbild im Recht bedeutet*. In: *Anwaltsblatt* 12/2017, S. 1174–1185.
- 47 BVerfG, Beschluß vom 16.07.1969 - 1 BvL 19/63.
- 48 In: D. u. W. Koch (2013). *Konrad Adenauer. Der Katholik und sein Europa*. Kiflegg 2018, S. 25, 119.
- 49 M. Hildebrandt, A. Rouvroy (2011). *Law, Human Agency and Autonomic Computing: The Philosophy of Law Meets the Philosophy of Technology*. Abington, S. 5.
- 50 Die vielzitierte Äußerung entstammt einem Gespräch mit Jean-Michel Palmier vom 3. Mai 1969. In: *Michel Foucault. Schriften in vier Bänden. Dits et Ecrits*. Hrsg. von Daniel Defert und François Ewald. Frankfurt a. M. 2003. Bd. I, S. 1002.
- 51 *Programmierte Freiheit*, S. 13.
- 52 H. Arendt (1967). *Wahrheit und Lüge in der Politik. Zwei Essays*. München, Berlin 2016, S. 42.
- 53 K. Lehmann (2012). *Vom Durchbruch auf dem Zweiten Vatikanischen Konzil bis zur heutigen Zeit*. 2. Vorlesung der Heinrich-Heine-Gastprofessur Toleranz in der europäischen Geschichte und Gegenwart am 11.12.2012, Universität Düsseldorf. <https://mediathek.hhu.de/watch/aca75d7d-282c-40dc-a83a-ba95765fd747>.
- 54 <https://nickbostrom.com/ethics/values.html>.
- 55 A. Anton (2007). *The Nietzschean Influence in The Incredibles (2004)*. In: *The Amazing Transforming Superhero!* Hrsg. von Terrence R. Wandtke. McFarland 2007.
- 56 [https://de.wikipedia.org/wiki/Geschichte\\_der\\_k%C3%BCnstlichen\\_Intelligenz](https://de.wikipedia.org/wiki/Geschichte_der_k%C3%BCnstlichen_Intelligenz).
- 57 K. Boff (2006). *Revolutions and shifting paradigms in human factors & ergonomics*. In: *Applied Ergonomics* 37(4):391-9. [https://www.researchgate.net/publication/7025963\\_Revolutions\\_and\\_shifting\\_paradigms\\_in\\_human\\_factors\\_ergonomics](https://www.researchgate.net/publication/7025963_Revolutions_and_shifting_paradigms_in_human_factors_ergonomics).
- 58 *Neurotechnologie-Startup "Neuralink". Elon Musk will Gehirn und Computer verbinden*. Focus, 17.7.2019. [https://www.focus.de/finanzen/boerse/neurotechnologie-startup-neuralink-elon-musk-will-gehirn-und-computer-verbinden-affenversuche-bereits-erfolgreich\\_id\\_10933217.html](https://www.focus.de/finanzen/boerse/neurotechnologie-startup-neuralink-elon-musk-will-gehirn-und-computer-verbinden-affenversuche-bereits-erfolgreich_id_10933217.html).
- 59 *Neuroenhancement: Die philosophische Debatte*. Hrsg. von Klaus Viertbauer und Reinhart Kögerler. Frankfurt a. M. 2019.
- 60 [https://de.wikipedia.org/wiki/Blade\\_Runner](https://de.wikipedia.org/wiki/Blade_Runner).
- 61 Tyrell: *I'm surprised you didn't come here sooner*. Roy: *It's not an easy thing to meet your Maker*. Tyrell: *What can He do for you?* Roy: *Can the Maker repair what He makes?* Tyrell: *Would you like to be modified?* Roy: *I had in mind something a little more radical*. Tyrell: *What seems to be the problem?* Roy: *Death*. Tyrell: *Death. Well, I'm afraid that's a little out of my jurisdiction, you ...* Roy: *I want more life, father*.
- 62 *I've seen things you people wouldn't believe. Attack ships on fire off the shoulder of Orion. – I watched c-beams – glitter in the dark near Tannhäuser Gate... All those moments will be – lost in time – like tears in rain. – Time – to die*. <https://www.youtube.com/watch?v=NoAzpa1x7jU>
- 63 *Die fröhliche Wissenschaft*, III, Nr. 125.
- 64 <https://nickbostrom.com/ethics/values.html>.
- 65 S. L. Sorgner (2009). *Nietzsche, the Overhuman, and Transhumanism*. *Journal of Evolution and Technology*, Vol. 20, Issue 1, March 2009, pgs 29-42.
- 66 Y. Harari (2015). *Homo deus. Eine Geschichte von Morgen*. Übers. von Andreas Wirthensohn. München 2018, S. 34.
- 67 *Zur Genealogie der Moral*. § III.9.
- 68 F. Fukuyama (2004). *The world's most dangerous idea: transhumanism*. *Foreign Policy*, Vol. 144 (2004), 42–43.
- 69 N. Bostrom (2004). *Transhumanism: The World's Most Dangerous Idea?* <https://nickbostrom.com/papers/dangerous.html>.
- 70 Nr. 386, Nr. 398.
- 71 R. Spaemann (1991). *Die europäische Kultur und der banale Nihilismus*. *Umkehr* 1 (1993), S. 9. <http://www.kath-info.de/kultur.html>.