Künstliche Intelligenz und menschliches Maß

Ulrich Steckmann und Bert Heinrichs

Abstract: Die fortschreitende Verwendung von Verfahren künstlicher Intelligenz provoziert die Frage, inwiefern bei deren ethischer Bewertung auf anthropologische Reflexionen und insbesondere auf ein Kriterium des menschlichen Maßes zurückgegriffen werden kann. Der Beitrag thematisiert das schwierige Verhältnis von Ethik und Natur und erkundet über die Auseinandersetzung mit maßgeblichen kritischen Herausforderungen die Spielräume für den ethischen Rückgriff auf anthropologische Argumente, wobei der Einsatz von künstlicher Intelligenz in der Medizin den Anwendungsbezug bildet.

Keywords: künstliche Intelligenz, Medizinethik, philosophische Anthropologie, *animal rationale*, Raum der Gründe, Diskursivität, Anthropologiekritik, naturalistischer Fehlschluss, Posthumanismus.

1. KI als potentes Werkzeug und als anthropologische Herausforderung

Verfahren der künstlichen Intelligenz (KI) haben binnen weniger Jahre enorm an Bedeutung gewonnen und kommen heute in vielen Lebensbereichen zur Anwendung.¹ In einem Beitrag für das Magazin *Forbes* listet Bernard Marr zehn Beispiele dafür, dass KI bereits integraler Bestandteil des Alltags vieler Menschen ist. Dazu zählen Handys, Social-Media-Plattformen, E-Mail-Dienste, Internetsuchmaschinen, Smart Home Devices, Bankdienstleistungen und natürlich Empfehlungsalgorithmen, wie beispielsweise Amazon und Netflix sie verwenden (siehe Marr 2019). Auch in der Medizin kommt KI mittlerweile vermehrt zum Einsatz. Vor allem solche Teilbereiche, die mit bildgebenden Verfahren zu diagnostischen Zwecken arbeiten, bieten sich für den Einsatz von KI an. Sogenannte Convolutional Neural Networks (CNN) haben in den vergangenen Jahren enorme Erfolge in der Verarbeitung von Bilddaten erzielen können (siehe grundlegend hierzu Krizhevsky et al. 2012). Daneben gibt es aber auch andere KI-Methoden und Einsatzbereiche im medizinischen Bereich. So wird etwa Reinforcement Learning (RL) genutzt, um Diagnoseempfehlungen anhand von großen Datenmengen zu optimieren.

Die Verwendung von KI wirft eine ganze Reihe schwerwiegender Probleme auf. Eines der zentralen Probleme betrifft die epistemische Intransparenz einiger KI-Methoden, zu denen auch CNN und RL zählen. Ansätze unter dem Stichwort ‚Explainable AI‘ versuchen, diesem

¹ Der Begriff ‚künstliche Intelligenz‘ ist ein Oberbegriff, der eine Vielzahl von Methoden und Verfahren unter sich vereint und an seinen Rändern zunehmend unscharf wird. Zur geschichtlichen Entwicklung sowie zu einigen Binnendifferenzierungen vgl. Heinrichs et al. 2022, Kapitel 1. Einen tiefergehenden, philosophisch geprägten Überblick geben Bringsjord und Govindarajulu 2020.

Einwand zu begegnen und KI in epistemischer Hinsicht transparent zu machen (siehe Samek et al. 2019). Dies ist vor allem wichtig, um seltene, aber schwer nachvollziehbare Fehler von KI besser zu verstehen und zu vermeiden (siehe Heaven 2019). Jenseits dieses Problemkomplexes stellt sich die Frage, ob der Einsatz von KI speziell in der Medizin nicht eine noch grundlegendere anthropologische Herausforderung darstellt. Dies soll einleitend an zwei Beispielen illustriert werden.

Esteva publizierte im Jahr 2017 mit Kolleginnen und Kollegen eine Studie, in der sie eine KI zur Klassifizierung von Hautveränderungen vorstellen (siehe Esteva et al. 2017). Die Gruppe von KI-Spezialisten und Dermatologinnen und Dermatologen der Stanford University berichtete, dass sie ein CNN mit einem Datensatz von 129.450 klinischen Bildern trainierte. Die Wissenschaftlerinnen und Wissenschaftler verglichen anschließend an biopsiegeprüften klinischen Bildern die Leistung ihres CNN mit den Diagnosen von 21 anerkannten Dermatologinnen und Dermatologen. Die Aufgabe bestand zum einen darin, bösartige Karzinome von gutartigen Geschwüren (seborrhische Keratosen) zu unterscheiden; zum anderen mussten maligne Melanome von gutartigen Hautfehlbildung (Nävi) unterschieden werden. Im ersten Fall ging es um die Identifizierung der häufigsten Krebsarten, im zweiten Fall um die Identifizierung des tödlichsten Hautkrebses. Laut dem Bericht erreichte das CNN ebenso gute Ergebnisse wie die Dermatologinnen und Dermatologen. Die Autorinnen und Autoren der Studie folgerten, dass durch die Ausrüstung von Smartphones mit einem derartig trainierten CNN möglicherweise die dermatologische Vorsorge außerhalb von Kliniken deutlich erhöht werden könne. Gehe man davon aus, dass bis zum Jahr 2021 voraussichtlich 6,3 Milliarden Smartphones weltweit im Umlauf seien, dann könne durch diese Technologie ein kostengünstiger Zugang zu lebenswichtiger diagnostischer Versorgung angeboten werden. Dieses Beispiel zeigt eindrücklich das Potenzial, das KI-Systeme in der medizinischen Diagnostik haben. Solche Systeme können eine qualitativ hochwertige und gleichzeitig leicht verfügbare medizinische Diagnostik ermöglichen und auf diese Weise die medizinische Versorgung für sehr viele Menschen erheblich verbessern.²

In einer Studie aus dem Jahr 2018 berichteten Komorowski und Kollegen über eine KI zur Unterstützung von Therapieentscheidungen bei Sepsis (siehe Komorowski et al. 2018). Anders als bei dem zuvor beschriebenen diagnostischen System kam hier die Methode des verstärkten Lernens (RL) zum Einsatz. Die Belohnung, mit der verstärktes Lernen stets operiert, war in diesem Fall an das Überleben der Patientinnen und Patienten 90 Tage nach Einweisung auf die Intensivstation geknüpft. In der Studie wurde berichtet, dass die Behandlung mithilfe des *AI Clinicians* durchschnittlich besser war als ohne. Dies führte das Autorenteam darauf zurück, dass das System implizites Wissen aus einer großen Menge von

² Das Beispiel ist aber auch geeignet, die Herausforderungen zu beleuchten, die mit dem Einsatz von KI in der medizinischen Diagnostik verbunden sind. Bei einer kritischen Überprüfung stellte sich heraus, dass das System Bilder von Hautveränderungen, auf denen Markierungen zur Bestimmung der Größe mit abgebildet waren, übermäßig häufiger als bösartig einstufte als Bilder ohne ein entsprechendes Skalierungshilfsmittel. Es zeigt sich, dass in den Trainingsdaten Bilder mit Markierungen signifikant häufiger bösartige Hautveränderungen zeigten, was zur Folge hatte, dass das System lernte, die Abbildung von Markierungen als maligne Hautveränderungen zu deuten; vgl. Narla et al. 2018.

Patientendaten extrahiert habe, die die Lebenserfahrung menschlicher Klinikerinnen und Kliniker um ein Vielfaches übersteige, und optimale Behandlungsstrategien durch die Analyse einer Vielzahl von meist suboptimalen Behandlungsentscheidungen lernen konnte. Stellt man in Rechnung, dass Sepsis eine der Haupttodesursachen in Krankenhäusern ist und nach wie vor kein gesichertes Wissen über optimale Behandlungsstrategien vorliegt, ist die Entwicklung eines solchen Systems sicher ein großer Gewinn.

Auf den ersten Blick handelt es sich bei den beiden beschriebenen Systemen um Werkzeuge, die die medizinische Praxis verbessern können, indem sie präzisere Diagnosen und effektivere Therapieempfehlungen bieten können als es bislang möglich war. Dies ist ohne Zweifel eine positive Entwicklung. Gleichzeitig steigern die beschriebenen Systeme aber auch den ohnehin schon hohen Grad der Technisierung in der Medizin, der von vielen bereits heute als problematisch angesehen wird. Während ein undifferenzierter Technikpessimismus sicher haltlos ist, könnte das Unbehagen angesichts einer schnell voranschreitenden Technisierung der Medizin dann begründet sein, wenn dieser Prozess die Medizin als menschliche Praxis zu unterminieren droht. Diesem Gedanken liegt natürlich ein spezielles Verständnis von menschlicher Praxis zugrunde. In erster Näherung kann man sagen, dass sich eine menschliche Praxis dadurch auszeichnet, dass sie durch bestimmte Interaktionsformen konstituiert wird, die ihrerseits auf anthropologische Grundeigenschaften verweisen. Fraglich ist dann, ob die Art der Technisierung der Medizin, die durch den Einsatz von KI eingesetzt hat, mit solchen anthropologischen Grundeigenschaften in Konflikt stehen könnte. Oder, um es etwas pointierter zu formulieren: Könnte es sein, dass KI in der Medizin ein menschliches Maß fehlt?

2. Das menschliche Maß als ethisches Kriterium

Das Ansinnen, im Zuge einer ethischen Evaluation auf ein menschliches Maß oder ein Maß des Menschlichen zu rekurrieren, ist keine Selbstverständlichkeit und entsprechend erläuterungs- und rechtfertigungsbedürftig. In außerphilosophischen Diskussionskontexten trifft man nicht selten – und bevorzugt bei der moralischen Bewertung technischer Innovationen – auf Warnungen vor einem ‚Verlust des menschlichen Maßes‘ oder auf Mahnungen, ‚das Maß des Menschlichen nicht zu überschreiten‘. Doch in den meisten Fällen erweisen sich derartige Bezugnahmen als unpräzise oder sogar als dubios. Moralphilosophische Rekonstruktionen solcher Diskurspositionen bestehen daher in der Regel darin, diejenigen zugrunde liegenden moralischen Intuitionen, die als rechtfertigungsfähig gelten können, in den geläufigen Terminologien konsequentialistischer oder deontologischer Ethiken einzufangen, ohne dabei auf ein Konzept des menschlichen Maßes zurückzugreifen. Daneben zielt die ethische Analyse in diesen Fällen auch darauf, offenzulegen, inwieweit sich in Bezug auf ein angeblich allgemein Menschliches lediglich partikuläre Wertüberzeugungen verklausuliert finden.

Soll nun ein menschliches Maß als ethisches Kriterium in Anspruch genommen werden, stellen sich zwei grundlegende Aufgaben: Zum einen wäre zu zeigen, inwiefern eine solche Inanspruchnahme für die ethische Evaluation einen Gewinn darstellen kann. Zum anderen muss plausibel gemacht werden, dass der ethische Rückgriff auf anthropologische Argumente

auf die skeptischen Einwände reagieren kann, die einem solchen Vorhaben traditionell entgegengebracht werden.

Bevor jedoch zu erläutern sein wird, wie die Einbeziehung anthropologischer Argumente zu einem Mehrwert für die ethische Evaluation bestimmter technologischer Innovationen beitragen kann, ist zu klären, in welcher Weise hier die notorisch schillernden Ausdrücke „Anthropologie“ und „anthropologisch“ verwendet werden. Bekanntlich ist die Anthropologie keine einheitliche wissenschaftliche Disziplin – eine Wissenschaft vom Menschen existiert vielmehr nur im Plural. In den empirischen Wissenschaften bezeichnet ‚Anthropologie‘ zum Teil ein spezifisches Forschungsfeld im Rahmen etablierter Disziplinen (bspw. die biologische Anthropologie als ein Teilgebiet der Zoologie), zum Teil widmen sich eigenständige Disziplinen bestimmten Aspekten des Menschseins (bspw. die Kulturanthropologie). Innerhalb der Philosophie stößt man zwar auf den Begriff der philosophischen Anthropologie, doch bezieht er sich nicht auf eine etablierte Disziplin der Philosophie, sondern wiederum auf ein Forschungsfeld mit unscharfer Umgrenzung. Die Frage, „inwiefern ‚Was ist der Mensch?‘ überhaupt noch eine philosophische Frage ist“ (Schnädelbach 1992, S. 117), wird mithin skeptisch beantwortet. Diese disziplinäre Konstellation ist die Folge einer Entwicklung der neuzeitlichen Philosophie, im Zuge derer eine „Abkehr vom Anthropozentrismus“ (Schnädelbach 1992, S. 125) zu beobachten ist. Das Wesen des Menschen in den Mittelpunkt philosophischer Reflexion zu rücken, war ein zentrales Projekt der Aufklärungsphilosophie³ und sollte nicht zuletzt auch einen praktischen Humanismus begründen. Ein Echo findet dieser Ansatz noch bei Kant, wenn dieser in seiner *Logik* davon spricht, dass sich die Grundfragen der Philosophie – „1) Was kann ich wissen? 2) Was soll ich thun? 3) Was darf ich hoffen?“ – auf eine weitere Frage, nämlich „Was ist der Mensch?“, bezögen (Kant, *Logik*, 25). Doch dieser Bezug ist bei Kant schon nicht mehr so zu begreifen, dass von der Anthropologie eine Grundlegung der gesamten Philosophie oder auch nur ihres praktischen Teils zu erwarten wäre.⁴ In der nachkantischen Philosophie wird dieses Abrücken von einem philosophischen Anthropozentrismus weitestgehend beibehalten. Die empirischen Wissenschaften hingegen betrachten den Menschen lediglich als ein spezifisches Anwendungsgebiet ihrer jeweiligen disziplinären Forschungsperspektive. Insofern sie dies tun, zählen sie für Kant zur „Anthropologie in physiologischer Hinsicht“⁵ – wobei neben die naturwissenschaftliche Naturalisierung des Menschen in der nachkantischen Zeit noch seine geisteswissenschaftliche Historisierung getreten ist. Infolge dieser Verschiebung der wissenschaftlichen Beschäftigung mit dem Menschen hin zu den Realwissenschaften verändert sich auch der Charakter der Thematisierung: Es wird nicht länger der Anspruch erhoben, Wesensbestimmungen des Menschen zu generieren, und schon gar nicht solche, die auch normativ gehaltvoll wären. Damit aber findet so etwas wie eine Entanthropologisierung der wissenschaftlichen

³ Siehe Schnädelbach 1992a, S. 119: „Das Wesen des wirklichen Menschen in den Mittelpunkt der Philosophie zu stellen und es zur Richtschnur des Erkennens und zum Maßstab des Handelns zu machen war einmal das Programm der *Aufklärung* gewesen“.

⁴ Siehe dazu unten, Abschnitt 3.

⁵ Vgl. Kant, 399: „Eine Lehre von der Kenntnis des Menschen, systematisch abgefaßt (Anthropologie), kann es entweder in *physiologischer* oder in *pragmatischer* Hinsicht sein. – Die physiologische Menschenkenntnis geht auf die Erforschung dessen, was die *Natur* aus dem Menschen macht [...].“

Beschäftigung mit dem Menschen statt, insofern nämlich die anthropologische Frage „Was ist der Mensch?“ – und das auch noch in der von Kant in der *Logik* gegebenen Bestimmung – auf eine umfassende Verständigung des Menschen über sich selbst abzielte.⁶ Die philosophisch-anthropologische Grundfrage „Wer sind wir (als Menschen)?“ spielt für die empirischen Wissenschaften, die Menschen zu ihren Forschungsgegenständen zählen, keine Rolle.

Dass die anthropologische Frage in der Philosophie trotz der gerade grob skizzierten Entwicklung virulent geblieben ist, zeigen unterschiedliche Rehabilitationsbemühungen, die bisweilen auch als „Wiederkehr anthropologischen Denkens“ (Barkhaus et al. 1996) annonciert werden. Zu denken ist zunächst an ein Spezifikum in der historischen Entwicklung der deutschen Philosophie: die Philosophische Anthropologie als Theorieströmung in der ersten Hälfte des 20. Jahrhunderts. Sie ist in erster Linie mit den Namen von Max Scheler, Helmuth Plessner und Arnold Gehlen verbunden, deren theoretische Programme allerdings heterogener ausfallen, als es die vereinheitlichende Rede von der Philosophischen Anthropologie eventuell vermuten lässt (siehe Krüger 2006, S. 23–29).⁷ So zielt etwa Schelers Ansatz auf eine „Metaphysik des Menschen“ (Hartung 2018, S. 61, vgl. Habermas 1973, S. 96) ab, der wieder die Rolle einer *prima philosophia* zufallen soll, während Gehlens Verständnis von Philosophischer Anthropologie eher auf eine Integration der Forschungserträge empirischer Wissenschaften hin angelegt ist.

Der zunächst naheliegende Gedanke, für die anthropologischen Argumente, die in die ethische Bewertung der zu Beginn des Beitrags angesprochenen Problemstellungen einbezogen werden sollen, auf die deutsche Philosophische Anthropologie zurückzugreifen, verliert bei näherem Hinsehen an Attraktivität. Abgesehen davon, dass die anthropologischen Überlegungen teilweise in einen Rahmen metaphysischer Spekulationen eingespannt sind (Scheler) oder auf Forschungsergebnisse empirischer Wissenschaften zurückgreifen, die überwiegend als überholt anzusehen sind (Gehlen), gilt für die unterschiedlichen Ansätze der Philosophischen Anthropologie, dass sie methodisch auf einem Niveau operieren, das den durch die analytische Philosophie etablierten allgemeinen Standards nicht genügt.⁸ Das schließt allerdings die Möglichkeit nicht aus, einzelne Bestimmungen oder Theorieelemente jener Anthropologien nach Maßgabe der gegenwärtigen methodischen Standards gewinnbringend zu rekonstruieren. Insbesondere Plessners phänomenologische Analysen der menschlichen Lebensform könnten sich in diesem Sinne als ähnlich anschlussfähig erweisen,

⁶ Herbert Schnädelbach charakterisiert die „anthropologische Aufgabe der Philosophie“ deshalb auch als eine „*Interpretationsaufgabe*“ (Schnädelbach 1992a, 135).

⁷ Scheler, Plessner und Gehlen gelten philosophiehistorisch als „Kerngruppe“ (Fischer 2006, S. 63) der Philosophischen Anthropologie. Zur Frage, inwieweit von einem „Identitätskern“ der unterschiedlichen Entwürfe dieser Vertreter der Philosophischen Anthropologie gesprochen werden kann, siehe Fischer 2006.

⁸ Es besteht kein Konsens darüber, was genau die analytische Philosophie als distinkte Strömung innerhalb der neueren und neusten Philosophie ausmacht und von anderen Strömungen abgrenzt (siehe Beaney 2013; Basile 2014, S. 184–186; Glock 2008, Kap. 8; Newen 2007, S. 11–19). Die Bemühung um begriffliche Präzision, insbesondere im Hinblick auf die Formulierung von Problemstellungen, sowie die konsequente Vermeidung eines tradierten Fachjargons sind aber wohl Merkmale, die für das Selbstverständnis der analytischen Philosophie zentral waren und nach wie vor sind.

wie es etwa mit den phänomenologischen Ansätzen Martin Heideggers und Maurice Merleau-Pontys in der Philosophie des Geistes bereits geschieht.

Ausdrücklich auf Plessners Werk bezieht sich beispielsweise Herbert Schnädelbach, wenn er für die Philosophie die Aufgabe einer „Anthropologie der Vernunft“ (Schnädelbach 1992b, S. 36) formuliert, die jedoch von vornherein als interdisziplinäres Projekt anzulegen sei. Schnädelbachs Vorschlag richtet sich darauf, das menschliche Selbstverständnis als *animal rationale* philosophisch zu explizieren, um radikal vernunftkritischen und eliminativistischen Positionen erfolgreich entgegenzutreten zu können. Von ähnlichen Intentionen ist auch Charles Taylors Bestimmung des Menschen als „self-interpreting animal“ motiviert, deren theoretische Entfaltung er ausdrücklich der „philosophical anthropology“ (Taylor 1985, S. 1) zuschlägt.⁹ Taylors Zugriff trägt neoaristotelische Züge, was ihn wiederum mit neueren Ansätzen wie etwa demjenigen von Michael Thompson verbindet, für den der Aristotelismus den Vorzug besitzt, „daß er dem Begriff des *Menschen* einen besonderen Status in der praktischen Philosophie gibt“ (Thompson 2011, S. 14–15).¹⁰

Die ethische Anknüpfung an die anthropologische Frage, was der Mensch sei bzw. was die menschliche Lebensform ausmache, ist maßgeblich motiviert von der Einsicht, dass moralische Subjekte bzw. Personen sich nicht allein über ein schmales Set rationaler Kompetenzen charakterisieren lassen. Auch wenn es für manche Anwendungsbereiche der Ethik angemessen sein kann, mit der Voraussetzung eines einfachen Selbstbestimmungskonzepts zu operieren, verfehlen derartige Reduktionen in anderen Fällen vom Ansatz her das Ziel, Normen für die moralische Rücksicht auf die Integrität personaler Existenz zu rechtfertigen. Konstitutiv für ein personales Leben ist das komplexe Zusammenspiel unterschiedlicher Eigenschaften und Fähigkeiten, die mit Gewissheit bislang nur den Angehörigen der menschlichen Lebensform zugeschrieben werden können. Höherstufige kognitive Kompetenzen bilden den Kernbereich jener Fähigkeiten, die allgemein als wesentlich für eine Zuweisung des vollen moralischen Status an Angehörige der betreffenden Lebensform angesehen werden (vgl. dazu den Überblick in Jaworska und Tannenbaum 2021). In dieser Form fungiert die traditionelle Bestimmung des Menschen als *animal rationale* in den Hauptströmungen der Ethik als Bezugspunkt der Zuschreibung sowohl eines Status als moralisches Subjekt als auch als eines Gegenstandes mit herausgehobenem Wert und besonderer Schutzwürdigkeit. Kant hat bekanntlich die Selbstzweckformel und das daraus folgende Instrumentalisierungsverbot auf die „vernünftige Natur“ bezogen (siehe Kant, GMS, S. 429).¹¹ Ebenso weist beispielsweise der gegenwärtige Präferenzutilitarismus eines Peter Singer den Besitzern höherstufiger kognitiver Fähigkeiten – in diesem Fall der Fähigkeit, eine „Präferenz hinsichtlich [der] eigenen zukünftigen Existenz“ (Singer 2013, S. 158) auszubilden – gewichtigere Interessen und damit einen höheren moralischen Status zu,

⁹ Allerdings gesteht Taylor umgehend zu, der Terminus sei geeignet „to make English-speaking philosophers uneasy«. Zur anthropologischen Bestimmung der *self-interpreting animals* siehe Taylor 1985, S. 45–76.

¹⁰ Siehe dazu ausführlicher Abschnitt 3.

¹¹ Jedoch ist darauf hinzuweisen, dass Kant mit der Bestimmung „vernünftige Natur“ keine Beschränkung auf die menschliche Gattung vornimmt.

auch wenn die daraus folgende Schutzwürdigkeit einen von der kantianischen Ethiktradition stark abweichenden Zuschnitt hat.¹²

Die kognitiven Kompetenzen, die die menschliche Selbstbestimmung konstituieren, sind keine Fähigkeiten, die abgetrennt von sämtlichen weiteren Eigenschaften und Fähigkeiten des Menschen bestünden und gleichsam von außen auf diese einwirkten. Bereits Johann Gottfried Herder hat in seiner Anthropologie betont, dass „die Vernunft keine abgeteilte, einzelwirkende Kraft“ sei, sondern eine der menschlichen Gattung eigene „Richtung aller Kräfte“ (Herder 1997, S. 29). Seine Überlegungen zielten nicht zuletzt auf eine nicht-dualistische Anthropologie. In der gegenwärtigen Theorielandschaft vor allem der Philosophie des Geistes wird an Herders Vernunftkonzeption insofern angeknüpft, als vielfach – in einer dezidiert anti-cartesianischen Wendung – von der Situiertheit des Geistes im Allgemeinen und der Rationalität im Speziellen ausgegangen wird. Die körperliche oder soziale Situierung kognitiver Fähigkeiten und Prozesse wird in dem zwischen Kognitionswissenschaften und Philosophie des Geistes aufgespannten interdisziplinären Raum als *embodied cognition*, als *situated cognition* bzw. als *embedded mind* angesprochen (siehe etwa Haugeland 1998; Robbins und Aydede 2010; Shapiro und Spaulding 2021). Diese Begriffe beziehen sich unter anderem darauf, wie kognitive Vorgänge in der menschlichen Lebensform situiert sind, womit traditionell anthropologische Fragestellungen berührt sind. Insbesondere den oben genannten phänomenologischen Ansätzen wird eine bedeutsame Rolle als theoretische Vorläufer zuerkannt (siehe bspw. Anderson 2003, S. 103–104; Fingerhut et al. 2013, S. 29–32 und 50; Gallagher 2010). Inwieweit Konzepte wie Leiblichkeit (Merleau-Ponty) und In-der-Welt-Sein (Heidegger) oder die Idee der Verschränkung von Außenwelt, Innenwelt und Mitwelt (Plessner) anschlussfähig sind für die empirische Forschung der Kognitionswissenschaften, ist Gegenstand interdisziplinärer Verständigungsprozesse. Die entsprechenden Forschungserträge von Kognitionswissenschaften und Philosophie des Geistes lassen jedenfalls erwarten, dass die philosophisch-anthropologische Auffassung vom *animal rationale* beträchtliche Weiterungen und Präzisierungen erfahren kann, die sich auch ethisch niederschlagen. Denn ein revidiertes Verständnis des *animal rationale* zieht mindestens der Möglichkeit nach auch ein anderes Bild der menschlichen Vulnerabilität nach sich, was in ethischer Hinsicht eine Neujustierung der Normen moralischer Rücksichtnahme erforderlich machen kann oder der ethischen Evaluation neue Anwendungsfelder eröffnet.

Neben der phänomenologischen Denktradition wird auch dem Pragmatismus eine Vorläuferrolle für die gegenwärtige Situated-Cognition-Forschung zugesprochen (siehe Fingerhut et al. 2013, S. 32–43). Aktuelle Vertreterinnen und Vertreter der pragmatistischen Tradition tragen aber noch in anderer Hinsicht zu einer ‚Anthropologie der Vernunft‘ bei. Der sich selbst als Neopragmatist begreifende Robert Brandom scheint zwar keine anthropologischen Ambitionen zu besitzen, doch sein Ansatz erlaubt nichtsdestoweniger ein erweitertes Verständnis des Menschen als *animal rationale*. Brandom rekonstruiert die

¹² Zudem weist Singer eine starre Verknüpfung solcher Kompetenzen mit der biologischen Spezies Mensch als Speziesismus zurück; vgl. Singer 2013, S. 98-107.

menschliche Vernunft als Fähigkeit, sich im sogenannten Raum der Gründe zu orientieren.¹³ Dazu gehört es auch, die Frage danach, wer „wir“ sind, stets aufs Neue zu stellen (siehe Brandom 1994, S. 4–5).

3. Einwände gegen die Verwendung von anthropologischen Argumenten in ethischer Absicht

Wer vorhat, anthropologische Argumente in ethischer Absicht einzusetzen, hat in Rechnung zu stellen, dass er eine ganze Reihe von philosophischen Einwänden abwehren muss. Zunächst einmal begegnet einem von metaethischer Seite der Vorwurf, der Versuch, anthropologische Bestimmungen und ethische Bewertungen zusammenzubringen, führe zu einem naturalistischen Fehlschluss (siehe etwa Bohlken 2009, 421-422). Aus einer noch so treffenden deskriptiven Erfassung der menschlichen Lebensform seien aus logischen Gründen grundsätzlich keine normativen Orientierungen ableitbar. Ein zweiter Typus von Einwänden bestreitet der philosophischen Anthropologie ihren zentralen Gegenstand. Bei der Beantwortung der Frage, was der Mensch sei bzw. die menschliche Lebensform ausmache, gelange man nicht über historisch kontingente Ausformungen menschlichen Lebens hinaus, die im schlechten Fall für Wesensbestimmungen ausgegeben werden. Der Einwand setzt also dem Vorhaben, anthropologische Bestimmungen zu gewinnen, die Überzeugung von der grundsätzlichen Unbestimmbarkeit des Menschlichen entgegen. Je nach theoretischer Voraussetzung wird dabei entweder argumentiert, die – in einem teleologischen Sinne verstandene – Bestimmung des Menschen sei noch nicht erreicht und auch vom gegenwärtigen Standpunkt aus epistemisch unzugänglich, oder aber die menschliche Lebensform sei grundsätzlich entwicklungsoffen, so dass man statt auf ein allgemein Menschliches stets nur auf historisch und kulturell variante Erscheinungsformen ohne kontextübergreifende Gemeinsamkeiten stoße. Drittens ist in jüngerer Zeit von Vertreterinnen und Vertretern trans- und posthumanistischer Positionen ein ethischer Einwand hinzugekommen: Zwar wird nicht grundsätzlich in Abrede gestellt, dass die menschliche Lebensform über bestimmbare Grenzen verfügt, doch aus trans- bzw. posthumanistischer Perspektive wird, wie die Selbstbezeichnungen „Trans-“ und „Posthumanismus“ bereits erkennen lassen, kein ethisches Problem darin gesehen, diese Grenzen zu überschreiten, wobei die Vertreter des Transhumanismus vor allem technische Erweiterungen und Ersetzungen im Auge haben.¹⁴ Bisweilen wird sogar eine moralische Pflicht zur Überschreitung des menschlichen Maßes – das in erster Linie als Freiheitsbeschränkung

¹³ Den Begriff des logischen Raums der Gründe übernimmt Brandom von Wilfrid Sellars, siehe Sellars 1997, § 36.

¹⁴ Siehe unten zu den mit diesen Theorieströmungen verbundenen Intentionen. Wenn hier verallgemeinernd von Trans- und Posthumanismus die Rede ist, geschieht das im Wissen um die Unterschiede zwischen beiden Strömungen sowie um deren jeweilige interne Heterogenität. Einen hilfreichen knappen Überblick über die betreffende Theorielandschaft bietet Loh 2020, S. 10–16. Im vorliegenden Beitrag geht es nicht um eine Zurückweisung von Trans- und Posthumanismus *in toto*, sondern lediglich um die kritische Diskussion von Positionen, die von einigen Vertreterinnen und Vertretern dieser Theorieströmungen verfochten werden.

aufgefasst wird – postuliert bzw. eine Pflicht zur Verfügbarmachung der notwendigen Mittel oder zur Unterstützung entsprechender Forschungen.

Der erste Typus von Einwänden bezieht sich auf das Verhältnis von Natur und Ethik im Allgemeinen. In Frage steht dabei, inwiefern die Ethik in rechtfertigungsfähiger Weise auf Naturtatsachen Bezug nehmen kann. Dabei wird insbesondere die Ableitung von normativen oder evaluativen Aussagen aus Tatsachenbehauptungen als Schlussfehler zurückgewiesen. Der betreffende Fehler firmiert üblicherweise als naturalistischer Fehlschluss. Die Lehre vom naturalistischen Fehlschluss geht bekanntlich auf G. E. Moore zurück und bildet eines der prominentesten Theoriestücke der Metaethik des 20. Jahrhunderts. Mittlerweile hat sich allerdings der Vorwurf, es liege ein naturalistischer Fehlschluss vor, vielfach vom ursprünglichen Einführungskontext abgelöst und erfährt oftmals eine undifferenzierte Verwendung. Moore bezog sich mit „naturalistischer Fehlschluss“ auf einen Definitionsfehler, der darin besteht, die evaluative, mithin moralische Eigenschaft „gut“ durch eine natürliche Eigenschaft zu definieren (siehe Moore 1993, S. 62, vgl. Heinrichs 2019, S. 98–116). Die Aufdeckung dieses Fehlers sollte bereits bei Moore dazu dienen, eine Naturalisierung der Ethik zu vereiteln, die eben auch vorläge, wenn man versuchte, aus Tatsachenbehauptungen über die menschliche Natur ethische Normen abzuleiten.

Auch wenn sich der Vorwurf des naturalistischen Fehlschlusses immer noch beträchtlicher Beliebtheit erfreut, gilt Moores Argumentation doch inzwischen als wenig überzeugend (siehe Heinrichs 2022, S. 174–175). Nichtsdestoweniger scheint Moore mit seiner Lehre einen bedeutsamen Punkt berührt zu haben: die Autonomie der Ethik. Die These von der Autonomie der Ethik hat mit Blick auf das Verhältnis von Ethik und Natur in erster Linie einen negativen Zuschnitt, indem sie nämlich die Unabhängigkeit der Ethik von Naturbestimmungen behauptet. Die Autonomie-These ist in der analytischen Ethik des 20. Jahrhunderts ausführlich diskutiert worden, doch sie führt philosophiegeschichtlich weit zurück. Eine maßgebliche Verteidigung der These stammt von Kant, der gleich in der Vorrede zu seiner *Grundlegung zur Metaphysik der Sitten* unmissverständlich klar macht, dass „der Grund der Verbindlichkeit“ moralischer Normen „nicht in der Natur des Menschen oder den Umständen der Welt, darin er gesetzt ist, gesucht werden müsse“ (Kant, *GMS*, 389). Die Begründung des moralischen Gesetzes muss Kant zufolge gänzlich frei gehalten werden von empirischen Bestimmungen. Sie sei stattdessen „*a priori* lediglich in Begriffen der reinen Vernunft“ (Kant, *GMS*, 389) zu suchen. Für das Verhältnis von (menschlicher) Natur und Ethik bedeutet das, dass beide in Kants Verständnis voneinander „sorgfältig gesäubert“ (Kant, *GMS*, 388) seien. Insofern scheinen sich in der Ethik positive Bezugnahmen auf die menschliche Natur grundsätzlich zu verbieten. Ebendas soll unter anderem auch die Lehre vom naturalistischen Fehlschluss nahelegen.

Doch daraus, dass Kants radikalisiertes Begründungsprogramm einen Graben zwischen Natur und Ethik aufreißt, entsteht ein Anschlussproblem, dessen sich Kant auch bewusst war. Denn die zunächst ohne Inanspruchnahme von Tatsachenbehauptungen über die Natur gerechtfertigten ethischen Prinzipien müssen auf die empirische Welt bezogen – und das heißt: in materiale Bestimmungen übersetzt – werden, wenn sie in irgendeiner Form

handlungsleitend werden sollen. Es stellt sich das „Anwendungsproblem der Ethik“ (Heinrichs 2022, S. 178). Dadurch gelangt die für die Begründungsfrage ausgeschlossene Natur der moralischen Subjekte zurück in die ethische Reflexion: „[E]ine Metaphysik der Sitten kann nicht auf Anthropologie gegründet, aber doch auf sie angewandt werden“ (Kant, MdS, 217). Allerdings hat Kant diesen Schritt zur Anwendung selbst nur um den Preis vollzogen, damit den Bereich des Wissenschaftlichen zu verlassen. Eine angewandte Ethik war für Kant aufgrund der Vielzahl und der Heterogenität der zu berücksichtigenden Gesichtspunkte nicht als integraler Bestandteil des von ihm konzipierten rationalistischen Systems der Ethik denkbar, sondern lediglich als dessen nichtwissenschaftlicher Anhang (siehe Kant, MdS, 468-469). Die restriktive kantische Auffassung von Wissenschaftlichkeit kann jedoch für die gegenwärtige Moralphilosophie wohl keine Verbindlichkeit mehr beanspruchen, so dass darin kein Hindernis mehr besteht für eine methodisch gesicherte Einbeziehung anthropologischer Argumente in die angewandte Ethik.

Mit Blick auf das Verhältnis von Natur und Ethik ist aber noch ein alternativer Argumentationsweg zugunsten der ethischen Bezugnahme auf die menschliche Natur zu erwägen. Verfechterinnen und Verfechter der oben bereits angesprochenen neoaristotelischen Ansätze halten Kants ‚sorgfältige Säuberung‘ der Begründung ethischer Normen von Naturbestimmungen und folglich auch das sich anschließende ethische Anwendungsproblem für einen Irrweg. Für einen Aristoteliker wie Michael Thompson sollen praktische Ziele und Normen vielmehr durch einen Begriff den Menschen bzw. der menschlichen Lebensform fundiert werden, wobei dieser aristotelisch geprägte Begriff im Unterschied zum kantischen Verständnis nicht als ein empirischer Begriff aufgefasst werde (siehe Thompson 2011, S. 16). In dieser Herangehensweise folgt ihm Philippa Foot, die zusammen mit Thompson zu den Hauptvertreterinnen und Hauptvertretern des aristotelischen Naturalismus zählt (siehe Foot 2004).¹⁵ Beide sind davon überzeugt, die Begriffe „Leben“ bzw. „Lebensform“ erlangten ihren Gehalt nicht durch eine direkte oder indirekte Verbindung mit der Erfahrung, sondern umgekehrt seien es diese Begriffe, mit deren Hilfe wir bestimmte Erfahrungen strukturieren und ihnen ihren spezifischen Sinn verleihen. Entscheidend dabei sei, dass dem Begriff der Lebensform eine inhärente Normativität eigne (siehe Thompson 2011, S. 96–99). Sogenannte naturhistorische Urteile (siehe Thompson 2011, S. 83–88) liefern einen Standard zur Beurteilung einzelner Individuen einer Lebensform und damit gleichzeitig auch die „Kategorie des ‚natürlichen Defekts‘“ (Thompson 2011, S. 104). Die humane Lebensform zeichne sich speziell dadurch aus, dass sie über die „Fähigkeit, Handlungsgründe anzuerkennen und entsprechend zu handeln“ (Foot 2004, S. 43) verfügt. Menschen, die die moralische Dimension der menschlichen Lebensform nicht anerkennen und umsetzen, sind demnach als defiziente Exemplare ihrer Spezies anzusehen. Gegen Foots Argumentation hat John McDowell allerdings den schwerwiegenden Einwand formuliert, dass

¹⁵ Vgl. Foot 2004. Zu nennen ist zudem noch Rosalind Hursthouse (1999). Als fraglich zu bewerten ist, inwieweit der von Martha Nussbaum verfochtene *Capabilities Approach* ebenfalls diesen Leitlinien folgt. Nussbaums frühere Arbeiten zum *Capabilities Approach* weisen noch eine starke neoaristotelische Prägung auf (siehe Nussbaum 1992), während sie in jüngerer Zeit weitestgehend dem Neokantianismus Rawls'scher Prägung folgt und für ihren Ansatz einen fundierenden Bezug auf eine *human nature* ausdrücklich zurückweist (siehe Nussbaum 2011, S. 28).

mit dem Auftreten der Rationalität jede inhaltliche Festlegung auf normative Standards hinfällig werde (siehe McDowell 1996, vgl. dazu Heinrichs 2017). Es zeichne die rationale Lebensform gerade aus, dass der Verweis auf externe Standards, die der ‚ersten Natur‘ entstammen, nicht mehr als Begründung hinreichen. Stattdessen müssen rationale Wesen eine ‚zweite Natur‘ ausbilden, die allerdings in die erste Natur eingebettet bleibt. Die rationale Lebensform ist eben nicht mehr auf die normativen Standards der Lebensform festgelegt. Sie muss sich ihre normativen Standards vielmehr selbst geben. Diesen Gedanken hat Kant mit dem Begriff der Autonomie zum Ausdruck gebracht (siehe Kant, GMS, 433).

Die Zurückweisung des Einwands, dass der ethische Rekurs auf anthropologische Bestimmungen grundsätzlich invalide sei, weil er auf einen Argumentationsfehler hinauslaufe, ist jedoch nur ein erster Schritt. Der zweite Typus von Einwänden richtet sich gegen die Möglichkeit anthropologischen Wissens. Eine ethische Bezugnahme auf anthropologische Bestimmungen kann schließlich auch daran scheitern, dass es so etwas wie *die* Natur des Menschen oder *den* Menschen gar nicht gibt und folglich auch keine anthropologischen Tatsachenbehauptungen verfügbar sind. In einem schon älteren Beitrag hat Jürgen Habermas diese Kritik am anthropologischen Wissen auf den Punkt gebracht: „[D]en‘ Menschen gibt es sowenig wie ‚die‘ Sprache. Weil Menschen sich erst zu dem machen, was sie sind, und das, den Umständen nach, je auf eine andere Weise, gibt es sehr wohl Gesellschaften und Kulturen, über die sich [...] allgemeine Aussagen machen lassen; aber nicht über ‚den‘ Menschen.“ (Habermas 1973, S. 105-106) Dieser Einwand besitzt auch eine kritische epistemologische Pointe, der zufolge anthropologische Bestimmungen unter einer systematischen kontextuellen Verzerrung leiden: „Wer Anthropologie treibt, [...] läßt seinen Begriff vom Menschen anleiten durch die objektiven Interessen, die aus den geschichtlichen Tendenzen der gesellschaftlichen Entwicklung hervorgehen.“ (Habermas 1973, S. 110, vgl. Schnädelbach 1992b, S. 32) Nicht zuletzt werden derartige Argumente oftmals gegen die Formulierung von Menschenrechten ins Feld geführt¹⁶ und haben gerade in interkulturellen Debatten nach wie vor großen Einfluss. Der Mensch wird dabei als „nicht festgestelltes Tier“ (Nietzsche) betrachtet, das sich in unterschiedlichste Richtungen entwickeln kann, so dass die Rede von einer singulären menschlichen Lebensform substanzlos sei und man in deskriptiver Hinsicht nicht über eine Pluralität von menschlichen Lebensformen hinausgelangen könne. Wo in normativer Absicht anthropologische Bestimmungen dennoch ins Spiel gebracht werden, stehen sie vor diesem Hintergrund unter einem in der Regel anti-essenzialistisch motivierten Ideologieverdacht.

Derartige Einwände lassen sich nicht einfach abweisen. Anthropologische Reflexionen in der Philosophie nehmen keine rein objektsprachliche Form an, sondern enthalten stets einen Selbstbezug, der interpretatorisch und somit theoretisch-praktisch angelegt ist. Die philosophisch-anthropologische Frage „Wer sind *wir*?“ besitzt immer auch eine normative

¹⁶ Ein *locus classicus* für eine solche Argumentation ist die Stellungnahme der American Anthropological Association zum Entwurf der UN-Menschenrechtserklärung von 1948, siehe Executive Board, American Anthropological Association 1947.

Dimension, woraus allerdings keine Festlegung auf eine aristotelische Anthropologie folgt, aus der ethische Prinzipien einfach ableitbar wären.

Die interpretative Struktur anthropologischer Reflexion eröffnet indes den Raum für eine relativistische Skepsis. Das „Wir“ in der philosophisch-anthropologischen Frage „Wer sind wir?“ kann sich immer auch lediglich auf eine partikulare Herkunftsgemeinschaft beziehen, worüber sich die jeweilige Sprecherin oder der jeweilige Sprecher täuschen kann. Doch die bloße Möglichkeit einer partikularistischen Verkennung schlägt anthropologische Ansätze noch nicht aus dem Feld. Es bleibt stets eine offene Frage, inwieweit sich mutmaßlich anthropologische Bestimmungen tatsächlich im Diskurs kontextübergreifend als etwas Allgemeinmenschliches erhärten lassen. Es ist allerdings nicht erkennbar, weshalb die Beweislast in dieser Frage einseitig bei den Verfechterinnen und Verfechtern anthropologischer Bestimmungen liegen sollte. Allgemeine Charakterisierungen der menschlichen Lebensform wie die, dass der Mensch ein soziales Wesen ist, dass er ein rationales und zwecksetzendes Wesen ist oder dass er psychophysisch konstituiert ist, sind jedenfalls nicht von vornherein dogmatismusverdächtig, zumal, wenn sie hinreichend formal ausbuchstabiert werden. Die Bestimmung des *animal rationale* als „Bewohnerinnen und Bewohner des Raums der Gründe“ vermeidet beispielsweise eine Vielzahl inhaltlicher Engführungen dieser klassischen anthropologischen Definition.

Ein weitere Herausforderung der Anthropologie weist eine auffällige Ähnlichkeit mit der gerade zuvor diskutierten Kritik an anthropologischen Ansätzen auf. In beiden Fällen richtet sich die Kritik darauf, dass es durch anthropologische Festlegungen sowohl in theoretischer als auch in praktischer Hinsicht zu Exklusionen komme. Der anthropologische Blick lasse Entwicklungsmöglichkeiten des Menschen entweder ganz unberücksichtigt oder klassifiziere sie auf eine auch ethisch problematische Weise als defizitär. Während zuvor vor allem die synchrone Diversität menschlicher Lebensformen gegen einen mutmaßlich homogenisierenden Zugriff der Anthropologie verteidigt werden sollte, steht nun – in diachroner Perspektive – die evolutionäre Offenheit des Menschen im Fokus. Letztere Form der Anthropologiekritik wird von Vertreterinnen und Vertretern trans- und posthumanistischer Positionen verfochten, die insbesondere in den Bereichen von Bio- und Technikethik aktiv sind.

Die Diagnose, dass das anthropologische Denken einen einengenden Effekt habe, wird in transhumanistischer Perspektive auf die evolutionäre Dynamik bezogen, der auch die menschlichen Lebensform ausgesetzt ist. Gegenstand der philosophischen Anthropologie sei stets der evolutionäre Status quo der menschlichen Lebensform, der essenzialistisch gedeutet werde, so dass wünschenswerte Möglichkeiten der Weiterentwicklung der Gattung aus dem Blick gerieten, wobei Transhumanistinnen und Transhumanisten diese Weiterentwicklung als einen aktiv betriebenen Prozess technologischer Überschreitungen der aktuellen Grenzen menschlicher Lebensform betrachten.¹⁷ Ihre Opponentinnen und Opponenten sehen Transhumanistinnen und Transhumanisten wie Bostrom in denjenigen, die die

¹⁷ Siehe etwa Bostrom 2005a, S. 4: „Transhumanists view human nature as a work-in-progress, a half-baked beginning that we can learn to remold in desirable ways.“ Vgl. auch Loh 2020, S. 41–64.

Schutzwürdigkeit der menschlichen Lebensform verteidigen. Deren Position wird als „biokonservativ“ bezeichnet (siehe Bostrom 2005b). Während auf der sogenannten biokonservativen Seite die Grundzüge der menschlichen Lebensform als bewahrenswert gelten, sieht der Transhumanist in erster Linie deren einschränkenden Charakter, der eine Überwindung wünschenswert erscheinen lässt.¹⁸

Bei genauerer Betrachtung wirkt die transhumanistische Position allerdings weniger radikal, als es die Rhetorik mancher ihrer Vertreterinnen und Vertreter vielleicht nahelegt. Dass der Transhumanismus keine grundsätzliche Abkehr vom Humanismus anstrebt, wird bereits deutlich, wenn seine Vertreter angeben, ihre Position habe „roots in secular humanist thinking“ (Bostrom 2005a, S. 4, vgl. dazu auch Loh 2020, S. 17–30). Ihre Forderung nach weitreichenden technologischen Optimierungen und Erweiterungen der menschlichen Lebensform sehen Transhumanistinnen und Transhumanisten ihrerseits anthropologisch fundiert, nämlich in einem menschlichen Streben nach Selbsttranszendenz (siehe Loh 2020, S. 33). Doch diese anthropologische Begründung offenbart eine Schwäche der transhumanistischen Position: Selbst wenn man das Streben nach Selbsttranszendenz als ein allgemeines Charakteristikum menschlichen Lebens ansieht, lässt sich doch kaum die Frage nach weiteren Charakteristika und deren Gewicht für die Fundierung des transhumanistischen Projekts abweisen. Der Verweis auf das menschliche Streben nach Selbsttranszendenz, das das technologische Optimierungsprojekt der Transhumanistinnen und Transhumanisten stützen soll, erscheint willkürlich und wenig überzeugend,¹⁹ wenn er nicht in eine differenzierte anthropologische Reflexion eingebunden wird, die gegebenenfalls auch gegenstrebige anthropologische Merkmale einbezieht und diese nicht einfach als ‚biokonservative‘ Ideologie abtut.

Die transhumanistische Verwurzelung im Humanismus und die Inanspruchnahme anthropologischer Argumente führt manche Posthumanistinnen und Posthumanisten dazu, vom Transhumanismus Abstand zu nehmen. Bei aller Heterogenität dieser Theorieströmung scheint eine entschiedene Kritik am Humanismus – aber nicht notwendig dessen vollständige Zurückweisung – ein einigendes Band darzustellen. Die Kritik ist anti-essenzialistisch formiert, das heißt, es wird die Möglichkeit bestritten, bestimmte Lebewesen aufgrund eines oder mehrerer Wesensmerkmale als Menschen identifizieren zu können. Die Überwindung des „unreflektierten Essenzialismus“ im systematischen Zentrum herkömmlichen anthropologischen Denkens führe langfristig, so die Erwartung vieler Vertreterinnen und Vertreter des Posthumanismus, zu einer „Überwindung der philosophischen Anthropologie als Disziplin“ (Loh 2020, S. 149). Sicherlich kann man der vorgebrachten Essenzialismuskritik insoweit folgen, als unterschiedlichste Wesensbestimmungen des Menschen immer wieder zu ethisch nicht hinnehmbaren Exklusionen beigetragen haben. Fraglich ist jedoch, inwieweit diese Diagnose hinreichende Gründe liefert, den Wert anthropologischer Reflexionen grundsätzlich skeptisch einzuschätzen. Verwendungen von anthropologischen Argumenten in

¹⁸ Siehe Bostrom 2005a: „There is no reason to think that the human mode of being is any more free of limitations imposed by our biological nature than are those of other animals.“

¹⁹ Loh spricht in diesem Zusammenhang sogar von einer transhumanistischen „Trivial-Anthropologie“ (Loh 2020, S. 82).

ethischen Evaluationen ließen sich mittels kritischer Begleitdiskurse durchaus sensitiv gestalten. Umgekehrt stellt sich – mindestens für die ethische Anwendung – die Frage nach gangbaren anti-essenzialistischen Alternativen. Nicht zuletzt greift die moralische Statuszuweisung traditionell auf die Bestimmung von Weseneigenschaften der möglichen Subjekte und Objekte von Normen moralischer Rücksichtnahme zurück.

Insgesamt lässt sich festhalten, dass die Einwände gegen das Vorhaben, anthropologische Argumente in ethischer Absicht zu verwenden, keine Ausschlusswirkung entfalten. Jeder der diskutierten Einwände kann auf triftige Argumente zurückgreifen, die dazu zwingen, den Einsatz anthropologischer Einsichten in ethischen Evaluationen immer wieder neu zu justieren. So ist die schlichte Ableitung von ethischen Normen aus naturalistischen anthropologischen Bestimmungen obsolet. Der Ort für die Einbeziehung der Anthropologie ist nicht die Fundierung ethischer Prinzipien, sondern deren Anwendung. Auch eine philosophisch-anthropologische Reflexion, die gegenüber eigenen schädlichen Essenzialismen blind und für empirische anthropologische Erkenntnisse taub ist, verbietet sich von selbst. Doch der bloße Umstand, dass zum Beispiel frühere Ausformulierungen des traditionellen anthropologischen Konzepts des *animal rationale* mit euro- oder androzentristischen Verzerrungen befrachtet waren, die auch schwerwiegende normative Konsequenzen nach sich zogen, beweist nicht die grundsätzliche Unbrauchbarkeit oder moralische Verwerflichkeit dieses Konzepts. Vielmehr bedarf es Reformulierungen, die die berechtigten Einwände konstruktiv aufnehmen.

4. Medizinische KI für den Menschen

Folgt man den zuvor angestellten Überlegungen, dann verfangen die Argumente gegen anthropologische Argumente lediglich teilweise. Richtig bleibt, dass ein Verweis auf *die* Natur des Menschen nicht ohne Weiteres herangezogen werden kann, um bestimmte Techniken oder Praktiken in ethischer Hinsicht zu bewerten. Dessen ungeachtet gibt es Bestimmungen des Menschen, die allen kritischen Vorbehalten standzuhalten vermögen. Dazu gehört insbesondere die Bestimmung des Menschen als sinnliches Vernunftwesen (Kant) oder eben als Wesen, das zwar bestimmte biologische Eigenschaften mit anderen Lebewesen teilt – vor allem natürlich mit den sogenannten Großen Menschenaffen –, das aber gleichzeitig als diskursives Wesen im Raum der Gründe agiert. Menschliches Handeln ist jederzeit in die diskursive Praxis des Gebens und Nehmens von Gründen eingebettet. Dies gilt auch für solche Handlungsweisen, die primär auf die Sinnlichkeit und Leiblichkeit des Menschen gerichtet sind. Stets kann man Handlungen auf ihre Gründe hin befragen und der Umstand, dass sie ein Grundbedürfnis – wie etwa das nach Nahrung oder Gesundheit – befriedigen ist nicht unbedingt geeignet, eine solche Befragung abschließend zu beantworten. Das Spiel des Gebens und Nehmens von Gründen bleibt vom Ansatz her offen. Menschen sind Wesen, die ihr Leben im Modus dieses offenen Spiels führen.

Um ‚menschlich‘ zu sein, muss Technik sich in die diskursive Praxis einbinden lassen. Das bedeutet insbesondere, dass ihre Nutzung das Fragen nach Gründen nicht grundsätzlich

ausschließen darf. Das Befragen eines göttlichen Orakels – das natürlich nur in einem sehr weiten Sinne als Technik gelten mag – kann dies versinnbildlichen: Ein Orakel gibt auf Anfragen hin Auskunft, entzieht sich aber jeglicher Rückfrage. In dieser Hinsicht hat es etwas ‚Unmenschliches‘ und mit dem menschlichen Maß Inkompatibles.

Moderne KI-Techniken können orakelhafte Züge annehmen, beispielsweise wenn sie für den Menschen verborgene Muster in sehr großen Datenmengen zutage fördern, sich die Mustererkennung selbst *ex post* aber nicht rational rekonstruieren lässt. Das oben beschriebene System zur Erkennung von Hautkrebs fällt in diese Kategorie. Welche Zuordnungen in den tieferen Schichten des CNN schließlich dazu führen, dass die im Bild präsentierte Hautveränderung als bösartig charakterisiert wird, bleibt – vollständig oder zumindest teilweise – unzugänglich. Eine Befragung nach Gründen ist nicht möglich. Allerdings gibt es neuere Ansätze, die darauf abzielen, genau dies zu ändern, indem sie „a qualitative understanding between the input variables and the responses“ liefern (Ribeiro et al. 2016, Abs. 2). Derzeit kommen solche Ansätze aber noch nicht verbreitet zum Einsatz. Die Ärztin oder der Arzt, die bzw. der ein System wie das zur Erkennung von Hautkrebs verwendet, wird seinerseits bzw. ihrerseits auf geringe Fehlerraten hinweisen und das Ergebnis des Systems auf diese Weise als verlässlich ausweisen. Es bleibt aber dabei, dass im konkreten Fall unklar ist, welche Gründe für eine Diagnose sprechen und welche dagegen. Mehr noch, es bleibt unklar, wie konfligierende Gründe, die dem Ergebnis des Systems widersprechen, zu gewichten sind. Wir können uns trotzdem entscheiden, solche Techniken einzusetzen, wenn ihr Nutzen zu überwiegen scheint. Gleichwohl kann man darauf beharren, dass ihr Einsatz dem zutiefst menschlichen Bedürfnis nach Begründung zuwiderläuft.

Die KI-Entwicklung hat bereits vor einiger Zeit auf die Kritik der Intransparenz und Opakheit von Algorithmen reagiert und begonnen, Methoden zu entwickeln, die es zumindest teilweise ermöglichen sollen, die Ergebnisse von KI-Systemen im Nachhinein verständlich zu machen (Ribeiro et al. 2016, Samek et al. 2019). Ob dies gelingt und welche Einbußen an Performanz damit einhergehen, ist bislang bestenfalls ansatzweise klar. Immerhin kann man diese Bemühungen so deuten, dass es so etwas wie ein menschliches Maß gibt, auf das bei der Bewertung von künstlicher Intelligenz zurückgegriffen werden kann. Auf eine Formel gebracht könnte man sagen: Das menschliche Maß besteht in der Möglichkeit der diskursiven Einbindung. Daran muss sich (Medizin-)Technik und insbesondere auch KI messen lassen, wenn es eine menschliche Technik sein soll.

Diese Forderung nach diskursiver Einbindung – oder besser: diskursiver Einbindbarkeit – von KI ist nicht gleichbedeutend mit der Forderung nach Lebensbedingungen, die nicht durch Technik dominiert werden und in denen zwischenmenschlicher Kontakt nicht zu kurz kommt, und sollte daher nicht mit letzterer verwechselt werden. Speziell eine „technisierte Medizin“ wird immer wieder als Herausforderung gesehen (vgl. Aurenque und Friedrich 2014). Dessen ungeachtet zeichnet sich der Mensch nicht zuletzt dadurch aus, dass er Technik und Techniken entwickelt, um sein Leben zu gestalten. Zwar weiß man seit Langem, dass es sich dabei um kein Alleinstellungsmerkmal handelt – auch nicht-menschliche Tiere verwenden Werkzeuge – , aber es gibt doch Gründe, die dafür sprechen, dass der Werkzeuggebrauch des Menschen

Besonderheiten aufweist (siehe Leroi-Gourhan 1988, S. 296). Es wäre vor diesem Hintergrund problematisch, die Verwendung von Technik im Allgemeinen und KI im Besonderen *per se* als mit einem menschlichen Maß unvereinbar zu charakterisieren. Richtig ist aber wohl, dass die Forderung nach diskursiver Einbindung nicht ausschließlich rationalistisch verstanden werden sollte. Erfolgreiche Kommunikation hat viele Ebenen, von denen einige die kognitiven und andere die emotiven Seiten des Menschen ansprechen. Im Pflegekontext können die emotiven Aspekte von Kommunikation besonders wichtig sein, weil sie das wechselseitige Beziehungsgefüge zwischen Personen betonen. Sally Dalton-Brown bemerkt zu Recht: „Carers are enmeshed within reciprocal, dependent interrelations.“ (Dalton-Brown 2019, S. 118) Entscheidend ist, dass Menschen unabhängig von ihrer jeweiligen Situation und Rolle ‚verstrickt‘ sind in diskursive Beziehungen mit anderen (und mit sich selbst) und sich daraus ein Maß für gute, menschliche Technik ergibt.

Die Vorbehalte gegen anthropologische Argumente in der Philosophie bleiben damit bestehen. Es bleibt dabei, dass eine abschließende Antwort auf die Frage „Was ist der Mensch?“ nicht möglich ist. Alle Versuche dieser Art münden in problematischen Festlegungen, die nicht selten ideologisch motiviert sind. Hinter diesen kritischen Befund gibt es kein Zurück mehr. Dies gilt in noch stärkerem Maße für alle Versuche, aus anthropologischen Bestimmungen unmittelbar normative Folgerungen zu ziehen. Moralische Fragen lassen sich nicht unter Verweis auf *die* Natur des Menschen beantworten. Gleichzeitig bleibt die anthropologische Frage aber nach wie vor virulent. Mit Kant könnte man von einer Frage sprechen, durch die wir belästigt werden, die wir aber nicht beantworten können (vgl. Kant, KrV, A, 7). Eine – sicher nicht neue – Lehre aus den langen Bemühungen um überzeugende Antworten kann sein, dass das Stellen der Frage selbst wichtiger ist als das Formulieren von Antworten. Auf einer höheren Stufe mag man dies dann sogar als eine Art von Antwort deuten: Der Mensch ist demnach dasjenige Wesen, das sich selbst unweigerlich fragwürdig bleibt, oder eben das Wesen, das fragt: „Wer sind wir?“ Dies wiederum deutet darauf hin, dass Menschen diskursive Wesen sind, die ihr Leben im Modus des offenen Spiels des Gebens und Nehmens von Gründen führen. Diese Auskunft nimmt zwar Bezug auf die anthropologische Frage, ohne sie aber endgültig zu beantworten.

Literaturangaben

- Anderson, Michael L. 2003. Embodied Cognition: A field guide. *Artificial Intelligence* 149: 91–130.
- Aurenque, Diana, und Orsolya Friedrich, Hrsg. 2014. *Medizinphilosophie oder philosophische Medizin? Philosophisch-ethische Beiträge zu Herausforderungen technisierter Medizin*. Stuttgart: Frommann-Holzboog.
- Barkhaus, Annette, Matthias Mayer, Neil Roughley, und Donatus Thürnau. 1996. Einleitung: Zur Wiederkehr anthropologischen Denkens. *Identität, Leiblichkeit, Normativität*.

- Neue Horizonte anthropologischen Denkens*, Hrsg. Annette Barkhaus, Matthias Mayer, Neil Roughley und Donatus Thürnau, 11–25. Frankfurt am Main: Suhrkamp.
- Basile, Pierfrancesco. 2014. Die analytische Tradition von Frege bis Quine. *Die Philosophie des ausgehenden 19. und des 20. Jahrhunderts 1*, Hrsg. Pierfrancesco Basile und Wolfgang Röd, 184–326. München: Beck.
- Beaney, Michael. 2013. What is Analytic Philosophy? *The Oxford Handbook of the History of Analytic Philosophy*, Hrsg. Michael Beaney, 3–29. Oxford: Oxford University Press.
- Bohlken, Eike. 2009: Von der Natürlichkeit zur künstlichen Künstlichkeit? *Das technisierte Gehirn. Neurotechnologien als Herausforderungen für Ethik und Anthropologie*, Hrsg. Oliver Müller, Jens Clausen, Giovanni Maio, 421–437. Paderborn: Mentis.
- Bostrom, Nick. 2005a. Transhumanist Values. *Ethical Issues for the Twenty-First Century*, Hrsg. Frederick Adams, 3–14. Charlottesville, VA: Philosophy Documentation Center.
- Bostrom, Nick. 2005b. In defense of posthuman dignity. *Bioethics* 19(3): 202–214. doi:10.1111/j.1467-8519.2005.00437.x
- Brandom, Robert B. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.
- Bringsjord, Selmer, und Naveen Sundar Govindarajulu. 2020. Artificial Intelligence. *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), Hrsg. Edward N. Zalta. <https://plato.stanford.edu/archives/sum2020/entries/artificial-intelligence/> Zugegriffen: 17. Juni 2022.
- Dalton-Brown, Sally. 2019. The Ethics of Medical AI and the Physician-Patient Relationship. *Cambridge Quarterly of Healthcare Ethics* 29: 115–121.
- Esteva, Andre, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, und Sebastian Thrun. 2017. Dermatologist-level classification of skin cancer with deep neural net-works. *Nature* 542: 115–118.
- Executive Board, American Anthropological Association. 1947. Statement on Human Rights. *American Anthropologist*. New Series 49(4): 539–543.
- Fingerhut, Joerg, Rebekka Hufendick, und Markus Wild. 2013. Einleitung. *Philosophie der Verkörperung. Grundlagentexte zu einer aktuellen Debatte*, Hrsg. Joerg Fingerhut, Rebekka Hufendick, und Markus Wild, 9–102. Berlin: Suhrkamp.
- Fischer, Joachim. 2006. Der Identitätskern der Philosophischen Anthropologie (Scheler, Plessner, Gehlen). *Philosophische Anthropologie im 21. Jahrhundert*, Hrsg. Hans-Peter Krüger und Gesa Lindemann, 63–82. Berlin: Akademie.
- Foot, Philippa. 2004. *Die Natur des Guten*. Frankfurt am Main: Suhrkamp.
- Gallagher, Shaun. 2010. Philosophical antecedents to situated cognition. *Cambridge Handbook of Situated Cognition*, Hrsg. Philip Robbins und Murat Aydede, 35–52. Cambridge: Cambridge University Press.

- Glock, Hans-Johann. 2008. *What Is Analytic Philosophy?* Cambridge: Cambridge University Press.
- Habermas, Jürgen. 1973. Philosophische Anthropologie (ein Lexikonartikel) – 1958. *Kultur und Kritik. Verstreute Aufsätze*, Jürgen Habermas, 89–111. Frankfurt am Main: Suhrkamp.
- Hartung, Gerald. 2018. *Philosophische Anthropologie*. 2., durchgesehene und erweiterte Auflage. Stuttgart: Reclam.
- Haugeland, John. 1998. Mind Embodied and Embedded. *Having Thought. Essays in the Metaphysics of Mind*, John Haugeland, 207–237. Cambridge, Mass: Harvard University Press.
- Heaven, Douglas. 2019. Why deep-learning AIs are so easy to fool. *Nature* 574: 163–166. <https://www.nature.com/articles/d41586-019-03013-5>. Zugegriffen: 26. Mai 2022.
- Heinrichs, Bert. 2017. Aristotelischer Naturalismus und der Begriff der Person. *Aristotelischer Naturalismus*, Hrsg. Martin Hähnel, 314–330. Stuttgart: Metzler.
- Heinrichs, Bert. 2019. *George Edward Moore zur Einführung*. Hamburg: Junius.
- Heinrichs, Bert. 2022. Moral gestalten – Zur Bezugnahme auf natürliche Tatsachen in der Ethik. *Natur, Ethik und Ästhetik*, Hrsg. Dieter Sturma, 173–186. Paderborn: Mentis.
- Heinrichs, Bert, Jan-Hendrik Heinrichs, und Markus Rüter. 2022. *Künstliche Intelligenz*. Berlin: de Gruyter.
- Herder, Johann Gottfried. 1997. *Abhandlung über den Ursprung der Sprache*. Hrsg. Hans Dietrich Irscher. Stuttgart: Reclam.
- Hursthouse, Rosalind. 1999. *On Virtue Ethics*. Oxford: Oxford University Press.
- Jaworska, Agnieszka, und Julie Tannenbaum. 2021. The Grounds of Moral Status. *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), Hrsg. Edward N. Zalta. <https://plato.stanford.edu/archives/spr2021/entries/grounds-moral-status/> Zugegriffen: 23. Mai 2022.
- Kant, Immanuel. 1968. Grundlegung zur Metaphysik der Sitten. *Kant's Gesammelte Schriften, Band IV (Akademieausgabe)*. Hrsg. Königlich Preußische Akademie der Wissenschaften. Berlin: de Gruyter. [= GMS]
- Kant, Immanuel. 1968. Kritik der reinen Vernunft (1. Aufl. 1781). *Kant's Gesammelte Schriften, Band IV (Akademieausgabe)*. Hrsg. Königlich Preußische Akademie der Wissenschaften. Berlin: de Gruyter. [= KrV]
- Kant, Immanuel. 1968. Die Metaphysik der Sitten. *Kant's Gesammelte Schriften, Band VI (Akademieausgabe)*. Hrsg. Königlich Preußische Akademie der Wissenschaften. Berlin: de Gruyter. [= MdS]
- Kant, Immanuel. 1968. Logik. In: *Kant's Gesammelte Schriften, Band IX (Akademieausgabe)*. Hrsg. Königlich Preußische Akademie der Wissenschaften. Berlin: de Gruyter. [= Logik]

- Komorowski, Matthieu, Leo A. Celi, Omar Badawi, Anthony C. Gordon, und Aldo Faisal. 2018. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine* 24: 1716–1720.
- Krizhevsky, Alex, Ilya Sutskever, und Geoffrey E. H. Hinton. 2012. ImageNet Classification with Deep Convolutional Neural Networks. <https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>. Zugegriffen: 26. Mai 2022.
- Krüger, Hans-Peter. 2006. Die Fraglichkeit menschlicher Lebewesen. Problemgeschichtliche und systematische Dimensionen. *Philosophische Anthropologie im 21. Jahrhundert*, Hrsg. Hans-Peter Krüger und Gesa Lindemann, 15–41. Berlin: Akademie.
- Leroi-Gourhan, André. 1988. *Hand und Wort. Die Evolution von Technik, Sprache und Kunst*. Frankfurt am Main: Suhrkamp.
- Loh, Janina. 2020. *Trans- und Posthumanismus zur Einführung*. 3., korrigierte Auflage. Hamburg: Junius
- Marr, Bernard. 2019. The 10 Best Examples of How AI Is Already Used in Our Everyday Life. <https://www.forbes.com/sites/bernardmarr/2019/12/16/the-10-best-examples-of-how-ai-is-already-used-in-our-everyday-life>. Zugegriffen: 26. Mai 2022.
- McDowell, John. 1996. Two Sorts of Naturalism. *Virtues and Reasons. Philippa Foot and Moral Theory*, Hrsg. Rosalind Hursthouse, Gavin Lawrence, und Warren Quinn, 149–179. Oxford: Clarendon.
- Moore, George E. 1993. *Principia Ethica*. Revised Edition. Cambridge: Cambridge University Press.
- Narla, Akhila, Brett Kuprel, Kavita Sarin, Roberto Novoa, und Justin Ko. 2018. Automated Classification of Skin Lesions: From Pixels to Practice. *Journal of Investigative Dermatology* 138: 2108–2110.
- Newen, Albert. 2007. *Analytische Philosophie zur Einführung*. Hamburg: Junius.
- Nussbaum, Martha C. 1992. Human Functioning and Social Justice. In Defense of Aristotelian Essentialism. *Political Theory* 20(2): 202–246.
- Nussbaum, Martha C. 2011. *Creating Capabilities. The Human Development Approach*. Cambridge, MA/London: Belknap Press.
- Ribeiro, Marco Tulio, Sameer Singh, und Carlos Guestrin. 2016. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*. Association for Computing Machinery, New York, NY, USA, 1135–1144. <https://doi.org/10.1145/2939672.2939778>.

- Robbins, Philip, und Murat Aydede. 2010. A Short Primer on Situated Cognition. *Cambridge Handbook of Situated Cognition*, Hrsg. Philip Robbins und Murat Aydede, 3–10. Cambridge: Cambridge University Press.
- Samek, Wojciech, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, und Klaus-Robert Müller, Hrsg. 2019. *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*. Charm: Springer.
- Schnädelbach, Herbert. 1992a. Die Philosophie und die Wissenschaften vom Menschen. *Zur Rehabilitierung des animal rationale. Vorträge und Abhandlungen 2*. Herbert Schnädelbach, 116–136. Frankfurt am Main: Suhrkamp.
- Schnädelbach, Herbert. 1992b. Zur Rehabilitierung des *animal rationale*. *Zur Rehabilitierung des animal rationale. Vorträge und Abhandlungen 2*. Herbert Schnädelbach, 13–37. Frankfurt am Main: Suhrkamp.
- Sellars, Wilfrid. 1997. *Empiricism and the Philosophy of Mind*. Cambridge, MA/London: Harvard University Press.
- Shapiro, Lawrence, und Shannon Spaulding. 2021. Embodied Cognition. *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Hrsg. Edward N. Zalta. <https://plato.stanford.edu/archives/win2021/entries/embodied-cognition/>. Zugegriffen: 23. Mai 2022.
- Singer, Peter. 2013. *Praktische Ethik*. Dritte, revidierte und erweiterte Auflage. Stuttgart: Reclam.
- Taylor, Charles. 1985. *Human Agency and Language. Philosophical Papers 1*. Cambridge: Cambridge University Press.