

Zwischen System und Verantwortung

Ethische Überlegungen zum Begriff der Autonomie angesichts der Entwicklung autonomer Systeme

Elisabeth Gräß-Schmidt

- › In der philosophischen Tradition steht die Autonomie des freiheitlichen Subjekts in einer Wechselbeziehung von Freiheit und Bindung an Gesetze. Autonomie bedeutet nicht Unabhängigkeit, sondern versteht sich als Selbstbegrenzung aus Freiheit, die Ordnung durch Gesetze anerkennt.
- › Mit der Entwicklung autonomer Systeme und ihrer Überlegenheit gegenüber manchen menschlichen Fähigkeiten verschiebt sich das Verständnis der Autonomie von der Selbstgesetzgebung des Menschen auf eine Selbststeuerung technologischer Systeme.
- › Wenn im Verhältnis von Technik und Ethik das Prinzip der Funktion zum Maßstab gemacht wird, entspringt das Handeln nicht mehr einer verantwortlichen Entscheidung, sondern der Steuerung des Systems.
- › Eine Maschinenethik, die die Verantwortung an die Maschine abgibt, ist auszuschließen, weil die Freiheit der Verantwortung nicht maschinell herstellbar ist.
- › Ethische Überlegungen müssen vorausschauend technische Entwicklungen begleiten und dem Menschen in ihrer Gestaltung die frei verantwortete Entscheidung zuweisen.

Inhaltsverzeichnis

1. Begriff, Gegenstand, und Reichweite der autonomen Systeme	2
2. Was ist Autonomie angesichts der Entwicklung sogenannter autonomer Systeme?	3
3. Herausforderung des Verständnisses der Autonomie des Menschen durch autonome Systeme	4
4. Autonome Systeme als Anfrage an das traditionelle Verständnis von Ethik und Technik	6
5. Kann es eine Maschinenethik geben?	7
6. Autonomie zwischen Personsein und Maschine – autonome Systeme und ihre ethische Verantwortung	8
Impressum	11

1. Begriff, Gegenstand, und Reichweite der autonomen Systeme

Unter dem Begriff des autonomen Systems verbirgt sich ein nahezu unüberschaubares Feld neuer technologischer Anwendungen. Die Feuilletons und Wissenschaftsteile der großen Zeitungen informieren nahezu täglich über neue Entwicklungen und die Chancen solcher Technologien, aber auch über ihre Gefahren. Diese liegen nicht nur in der möglichen Verdrängung der Arbeitskraft des Menschen, sondern in ihrer sämtliche Arbeits-, aber auch Lebensbereiche bestimmenden Macht. Sie bieten eine revolutionierende Entfaltungsmöglichkeit im Bereich der gesamten Kultur, vom Internet der Dinge über selbstfahrende Autos bis hin zum Einsatz von Robotern in der Medizin und im Pflegebereich. Aufgrund ihrer Potentiale sind diese Technologien daher konstruktiv, aber auch kritisch aufzunehmen. Die Konsequenzen für das Verständnis des Menschen, aber auch für Wirtschaft, Politik und Recht und Ethik müssen angesichts solcher Entwicklungen bedacht und gesamtgesellschaftlich diskutiert werden.

Digitalisierung bietet
Entfaltungspotentiale.

Zu denken wäre bei solchen Entwicklungen etwa an die Big Data-Technologie oder an Assistenzroboter in der Chirurgie – zur Optimierung von Operationsprozessen durch computer-gestützte technische Assistenz – oder an eine Verbesserung medizinischer Diagnosen. Zu denken ist auch an Verbesserungen in der Pflege im Krankenhaus und in Altenpflegeheimen, wenn durch automatisierte Technik der Alltag nicht nur der Pflegenden, sondern auch der Kranken besser bewältigt werden kann.¹ Solche technisch-intelligenten Hilfssysteme, die im modernen Roboter auch menschliche Hilfsdienste begleiten und teilweise ersetzen können, bedeuten einen Rückgewinn von Lebensqualität im Alter.²

Anwendungs-
möglichkeiten auto-
nomer Systeme

Eine weitere Leistungsmöglichkeit dieser Systeme ist die Entwicklung Künstlicher Intelligenz in Konkurrenz zum Menschen. So kann mithilfe künstlicher neuronaler Netze eine Künstliche Intelligenz das Denken des Menschen nachahmen und gegebenenfalls überbieten. Man spricht dabei von sich von menschlicher Intelligenz emanzipierenden, smarten Systemen³, die zur größeren Effizienz und Verlässlichkeit menschlichen Handelns beitragen können. Solche *Deep-learning*-Systeme sind selbstlernend, d. h. sie werden mithilfe von umfangreichen Datensätzen nicht mehr nur programmiert, sondern trainiert. Das führt zur Eigenständigkeit von Lernprozessen und umfasst nicht nur kognitive, sondern auch sensomotorische, emotionale und soziale Intelligenz. Während auf kognitiver Ebene die Künstliche Intelligenz weit fortgeschritten ist, ist sie jedoch auf der sensomotorischen Ebene von der menschlichen Leistung noch weit entfernt.

Potentiale Künst-
licher Intelligenz

Die mit der Digitalisierung einhergehende Transformation kultureller Selbstverständlichkeiten sind daher vielfältig. Sie bergen Chancen, aber auch Gefahren. So sind mögliche Gefahren zu bedenken, wenn es um die Entwicklung Künstlicher Intelligenz sogenannter smarterer

Systeme in der Kriegsführung geht. Aber auch die ethischen und rechtlichen Bedingungen der Verantwortung und Haftung müssen geklärt werden. Wer ist haftbar zu machen für Fehlentscheidungen? Liegt die Verantwortung weiterhin beim Menschen oder kann sie an ein autonomes System abgegeben werden? Kann von einer Zurechenbarkeit autonomer Systeme ausgegangen werden, etwa bei Handlungen intelligenter Kriegssoftware oder auch bei autonom fahrenden Autos?⁴

Offene Fragen bei
der Entwicklung
Künstlicher Intelligenz

Zu klären ist daher zunächst der Begriff der Autonomie. Lässt er sich auf ein System oder auf Formen Künstlicher Intelligenz übertragen? Erst dann kann der Anspruch und die Reichweite von sogenannten autonomen Systemen hinsichtlich ihrer Autonomie bestimmt werden. Dabei stellen sich weitere Fragen: Welche Konsequenzen ergeben sich aus den Autonomieansprüchen von Systemen bzw. Künstlicher Intelligenz für das Verständnis des Menschen und seiner Autonomie? Diese Frage berührt das Verhältnis von Technik und Ethik. Bedeutet die Entwicklung autonomer Systeme bzw. Künstlicher Intelligenz eine grundlegend neue Technik, und ist daher gar eine „Maschinenethik“ zu fordern, oder bleibt ethische Verantwortung an menschlich-personale Autonomie gebunden?

2. Was ist Autonomie angesichts der Entwicklung sogenannter autonomer Systeme?

Autonomie ist in der Moderne Ausweis der Vernunftnatur, der Selbstzwecklichkeit und inneren Freiheit des Menschen. Künstliche „Autonome Systeme“ – diese spannungsvolle begriffliche Zusammenstellung zweier sich auch widersprechender Konzeptionen, der der Autonomie und der des Systems –, stellen uns nun gegenwärtig vor große Herausforderungen. Sie betreffen nicht nur das Verhältnis von Technik und Ethik, sondern berühren die Frage nach dem Menschsein als solchem. Sah man die Selbstbestimmung des Menschen in antiker und christlicher Tradition vor dem Hintergrund einer inneren Freiheit, die es dem Menschen ermöglicht, selbst zu urteilen und Verantwortung zu übernehmen, so steht dieser Kern der Entscheidungsfähigkeit des Menschen durch die Übertragung des Begriffs auf ein System zur Disposition. Der Aspekt der Freiheit des Selbst ist entscheidend auch für den modernen Begriff der Autonomie. Selbstbestimmung als Entscheidungs- und Urteilskraft sind sein wesentliches Kennzeichen. Mit der Entscheidungsfähigkeit ist die Freiheit des Menschen angesprochen. Die Bestimmung eines Systems als „autonom“ wird sich daher an diesem Kriterium der Entscheidungs- oder Urteilsfähigkeit erweisen.

Selbstbestimmung als
Urteilskraft und Ent-
scheidungsfähigkeit

Autonomie bedeutet in der Tradition und in der Aufklärung Selbstbestimmung des Menschen als Auszeichnung menschlicher Freiheit. Der Ursprung des Begriffs der Autonomie in der griechischen Antike war mit der politischen Selbstbestimmung der Stadtstaaten verbunden. Autonomie (αὐτονομία), die Eigengesetzgebung, wurde als äußere politische Freiheit an Stelle von Fremdherrschaft verstanden.⁵ Spätestens seit der Bedrohung der politischen Unabhängigkeit durch die Perserkriege trat der Gedanke der Autonomie auch im Sinne einer inneren Freiheit auf. Dieser Gedanke zeigte sich ansatzweise bereits bei Herodot in der Bestimmung von Autonomie als Aufgabe, den Gesetzen zu gehorchen, die für alle gleich sind.⁶ In dieser inneren Freiheit entspricht er nicht zuletzt dem paulinischen Verständnis einer Freiheit in Christus. Unabhängig von äußeren Zwängen ist jeder Mensch aufgrund seiner inneren Freiheit gleich und frei (Galater 3, 28) durch und in der Bindung an Jesus Christus als der inneren Bestimmung des (göttlichen) Gesetzes. Christlich ist Autonomie daher weniger als Selbstgesetzgebung denn als freie, selbstbestimmte Entsprechung gegenüber dem vorgegebenen Gesetz verstanden.

Autonomie
als Ausweis der
inneren Freiheit zur
Selbstbestimmung

Explizit finden wir den Gedanken der Freiheit und Selbstzwecklichkeit als Ausdruck der Autonomie bei Aristoteles (Metaphysik 982b, 25 f.): „[E]in Mensch ist frei, wenn das Ziel des Handelns in ihm selbst liegt, wenn er um seiner selbst willen und nicht um eines anderen willen ist.“ Dieses Autonomieverständnis hat nachgewirkt und bei Thomas von Aquin sowie in der Neuzeit bei Spinoza im Gedanken der *causa sui* Ausdruck gefunden.

Autonomie als
Freiheit und Selbst-
zwecklichkeit

Die Bezogenheit des Selbst (αὐτός) im Begriff der Autonomie auf ein Gesetz (νόμος) setzt dieser dem Verständnis von Selbstbestimmung des Menschen als Willkür Grenzen. Dieser Bezug des Selbst zu einem ihm selbst vorgegebenen Gesetz ist mithin auch für das mündige, aufgeklärte, selbstbestimmte Subjekt konstitutiv. Autonomie ist daher nicht zu verwechseln mit bloßer Unabhängigkeit. Vielmehr bildet die Autonomie des freiheitlichen Subjekts eine Wechselbeziehung von Freiheit und Gesetz ab. Somit wohnt der Autonomie als Selbstbestimmung selbst die Autorität eines Anderen inne. Dem Selbst (αὐτός) steht der Begriff des Gesetzes (νόμος) gegenüber. Diese Autorität des Gesetzes darf die Freiheit nicht missachten, will sie sich nicht selbst aufheben.

Dieser spannungsvolle Zusammenhang ist in Kants Begriff der Autonomie festgehalten, wenn er Freiheit und eigene Gesetzgebung in der Autonomie als „Wechselbegriffe“ bezeichnet. Das Im-Blick-haben der Bezüge von Gesetz und Freiheit führt in der Beachtung der Gesetze des Zusammenhangs zur Selbstbegrenzung aus Freiheit. Dies ist der geniale Schachzug der Bestimmung von Freiheit als Autonomie bei Kant. Freiheit ist bei Kant eben nicht Willkürfreiheit und Beliebigkeit ihrer Ausdrucksformen, sondern Selbstbindung an ein Gesetz. Ein mögliches Missverständnis der Autonomie liegt daher in der Missachtung des Gesetzes als Bezug der Freiheit. Es entsteht aus dem Verkennen der Bedingungen der Freiheit, die als Ordnungen und Gesetze dem Leben dienen und es erhalten. Werden hingegen diese Bedingungen erkannt, führt dies zur Selbstbegrenzung aus Freiheit.

Autonomie als
Wechselbegriff von
Freiheit und Gesetz

Vor dem Hintergrund dieser Bestimmung von Autonomie als Ausdruck innerer Freiheit der Selbstverpflichtung stellt sich nun verstärkt die Frage: Ist ein solcher Begriff auf intelligente Maschinen oder Systeme anwendbar? Ist es möglich, dass ein Verständnis von Autonomie als innere Freiheit der Selbstentsprechung zum Gesetz durch ein System bzw. eine Künstliche Intelligenz in Anspruch genommen wird? Oder wird damit nicht die Errungenschaft der Differenzierung von äußerer, materieller Bestimmung von Freiheit und innerer Freiheit im Sinne der Selbstzwecklichkeit eingegeben? Mit anderen Worten: Bricht sich nicht der Vergleich von Mensch und Maschine am Verständnis des Menschen in seiner Würde und Personalität, wie es dem antiken und dem modernen, an der Aufklärung orientierten Verständnis von Personsein zugrunde liegt?

Anwendung des
Autonomiebegriffs
auf Systeme möglich?

3. Herausforderung des Verständnisses der Autonomie des Menschen durch autonome Systeme

Ein Verständnis des Menschen in seinem Geschaffensein entzieht sich einer Vorstellung des Menschen, die von seinem technischen Hergestelltsein ausgeht. Die Digitalisierung stellt uns jedoch mit ihrer Entwicklung autonomer Systeme auch vor Herausforderungen anthropologischer Natur. Die transhumanistische Idee der Verschmelzung von Mensch und Maschine, die über den Ansatz der Künstlichen Intelligenz hinausgeht, hat Folgen für die Bestimmung dessen, was als Menschsein wahrgenommen wird.

Ändert Künstliche
Intelligenz das
Verständnis vom
Menschen?

Können Maschinen Personen sein? Zukunftsforscher wie Michio Kaku und Ray Kurzweil denken, dass das möglich ist. Künstliche Intelligenz sei die „letzte Erfindung“, die wir als Menschen noch machen müssen, so Michio Kaku.⁸ Der Weg zur Künstlichen Intelligenz erweist sich nach

ihm durch Schaffung eines Ich-Bewusstseins und eines Sinnesempfindens, die selbst Kreativität ermöglichen. Alle Fähigkeiten werden nachgemacht werden können. Wird so der Mensch durch Selbstoptimierung über sich hinaus geführt im Sinne des von Nietzsche propagierten Übermenschen oder sogar abgeschafft durch Übersteigerung des Menschen in einem Transhumanismus, so stellt sich die Frage, ob dann durch diese Technologie das Verständnis des Menschen und seiner Selbstbestimmung nicht verändert wird. Ändern sich mit der Künstlichen Intelligenz die Parameter Vernunft und Freiheit des Menschen?

Heraufbeschworen wird diese Transformation im Bild einer *Brave New World*, die Technik nicht mehr nur als Dienst am Menschen, sondern als Selbstinszenierung des Menschen bis hin zum technologisch frisierten Übermenschen avisiert. Dieser Überschritt des Menschen zur Künstlichen Intelligenz vollzieht sich im Namen jener sogenannten „Singularität“, mit der der Übergang des Menschen zu einem selbstlernenden Intellekt erreicht ist, durch den man das autonome System ausgezeichnet sieht. Dies wäre ein Paradigmenwechsel, der schwer vorstellbar ist. Der von Yuval Harari prognostizierte *Homo Deus* löst den *Homo Faber* ab, ohne jedoch die bereits diesen treffende Kritik eines verengten Freiheits- und Autonomieverständnisses überwunden zu haben und überwinden zu können.

Übergang des Menschen zu einem selbstlernenden Intellekt

Die Künstliche Intelligenz in der Entwicklung autonomer Systeme wird so zu tief greifenden Veränderungen der Wahrnehmung des Menschseins führen können, die nicht erst das Menschsein in seiner natürlichen Ausstattung, wie in den gentechnischen Eingriffen, oder in der Gefährdung seiner Existenz, wie in der Kernenergie, berühren, sondern das Menschsein in seiner Bestimmung gefährden. Denn es geht jetzt nicht mehr nur um technische Errungenschaften, die wir in ihren Fortschritten als Zivilisation auch genießen können, sondern es geht jetzt um die technische Überformung der Natur des Menschen.

Dabei verändert sich nicht nur der Begriff des Selbst, des „Autos“, sondern auch der des Gesetzes, des Nomos. Beides hängt zusammen. Das Gesetz, der Nomos, hat seinen Ort nicht mehr als Gegenüber des Selbst, sondern das Selbst ist reduziert worden auf die technischen Funktionsprinzipien des Systems. Damit ist es nicht mehr rückgebunden an eine Freiheit, die sich selbst Rechenschaft zu geben und zur Rechenschaft gezogen weiß. Damit aber entschwindet der Begriff der Verantwortung und damit die Bedeutung, die der Begriff der Autonomie für den Menschen in seiner Selbstbestimmung hat. Autonomie verkörpert nicht mehr die Freiheit des Subjekts noch schließt sie seine Verantwortungsfähigkeit gegenüber sich selbst ein, sondern im Wandel vom Gesetz zur Funktion wird auch das Selbst des Menschen auf ein System reduziert. Aus der Selbstgesetzgebung des freiheitlichen Subjekts wird jetzt eine Selbststeuerung technologischer Systeme.

Von der Selbstgesetzgebung des Subjekts zur Selbststeuerung von Systemen

Mit der Entwicklung autonomer Systeme tritt daher eine Verschiebung des Verständnisses des Menschen und seiner Autonomie ein. Indem nicht mehr dem Subjekt, sondern einer selbst lernenden Maschine Autonomie zugesprochen wird, berührt das die Dimension von Selbstbestimmung und Freiheit als Ausweis der Autonomie.

Soll die Autonomie des Menschseins in seinem Personsein nun auf sein technisches Gemachtsein in Form von Künstlicher Intelligenz übergehen, dann stellt sich grundsätzlich die Frage nach dem Humanum. Bemessen wir Menschsein an seinen Fähigkeiten, seiner kognitiven Leistungskraft, dann wird es sich gegenüber der Künstlichen Intelligenz nicht behaupten können und wollen. So ist die Frage, was bestimmt den Menschen und was bestimmt seine Autonomie? Um einen Übergriff der Technik auf die „Naturwüchsigkeit des Menschen“ (Jürgen Habermas) zu verhindern, wird jetzt aufs Neue die Frage nach einer ethischen Verantwortung der neuen Technologien und damit das Verhältnis von Ethik und Technik drängend. Ist diesen mit einer herkömmlichen Ethik zu begegnen oder brauchen wir eine neue Ethik?

Unterscheidung zwischen Menschsein und Künstlicher Intelligenz

4. Autonome Systeme als Anfrage an das traditionelle Verständnis von Ethik und Technik

Kennzeichen von Ethik ist es, sich in Verantwortung verschiedenen Instanzen gegenüber verpflichtet zu wissen, den Instanzen des Selbst, der Welt sowie den Voraussetzungen von Selbst und Welt. Zu dieser Verantwortung des Menschen gehört auch die Technik. Wie es die philosophische Anthropologie herausgearbeitet hat, ist der Mensch dadurch bestimmt, dass er im Unterschied zum Tier keine instinktgeleitete Verhaltensweise und keine ökologische Nische hat, auf die er sich einfach verlassen kann. Vielmehr ist er in seinem Handeln auf die Entschlusskraft seiner Freiheit angewiesen. Er muss sich seinen eigenen Lebensraum bauen und gestalten. Technik gehört zum Menschsein dazu. Dann aber ist technischen Entwicklungen nicht einfach ein Verzicht zuzumuten, sondern sie sind in ihren Leistungen wahrzunehmen¹⁰ und als Verbesserung des Lebens zu begreifen. Die Technologie autonomer Systeme ist daher zunächst wie jede Technik in ihrer Ambivalenz wahrzunehmen. Sie kann ein sinnvolles Hilfsmittel sein, sie kann aber auch missbraucht werden.

Technik kann hilfreich sein, aber auch missbraucht werden.

Dennoch fragt es sich, ob nicht eine neue Stufe erreicht ist, die die Technik nicht mehr als neutral einzustufen erlaubt, weil sie das Menschsein in seiner Individualbestimmtheit tangiert, indem sie etwa den Menschen zum Übermenschen stilisieren möchte. Dies ist der Fall, wenn die Technik selbst ihr Gegenüber, sei es die Natur, sei es ein Gesetz als eine äußere Ordnung, in sich einzieht im Dienst der Funktion. Denn dann ist keine Gestaltung der Technik mehr möglich. Wenn das Prinzip der Funktion zum Maßstab der Autonomie gemacht würde, kann Autonomie sich nicht mehr als Verantwortung verstehen. Handeln bzw. Verhalten entspringt dann nicht einer verantwortlichen Entscheidung, sondern es fügt sich einem Modus der Steuerung. Eine Unterscheidung von ethischer und technischer Freiheit ist hinfällig. Indem ethische Gestaltung selbst nur noch nach technischen Maßstäben erfolgt, ohne den Blick auf die Dienlichkeit des Lebens aufrecht zu erhalten, hat sie ihr Eigenrecht verspielt.

Nivellierung der Unterscheidung von Technik und Ethik

Der Unterschied zeigt sich in den ethischen Überlegungen zur Autonomie. Während sie für das freiheitliche Subjekt Ausfluss und Reflexion von dessen Autonomie selbst waren, aus der sich die Prinzipien der Verantwortung ergaben, beziehen sie sich jetzt allenfalls auf ein Objekt. Das, was die Selbstbestimmung wesentlich ausmacht, die Reflexivität wird dadurch eingezogen. Das im Begriff der Autonomie das Autos konstituierende Andere des Nomos ist ihm abhandengekommen. Damit wird der Begriff der Autonomie entkernt. Er ist nicht mehr Ausdruck von Freiheit und Selbstbestimmung, sondern Ausdruck systemischer Funktionalität. Eine Rückbindung des Subjekts in seiner freiheitlich schöpferischen Kreativität an ein ihm Anderes ist damit ausgeschlossen. Dieser Ausschluss aber eliminiert ein kritisches Gegenüber zur Technik, als das die Ethik immer gelten konnte.

Damit aber haben sich die Verhältnisse zur Bestimmung von Natur und Ordnung, Freiheit und Gesetz, wie sie das Autonome in der christlichen und philosophischen Tradition gekennzeichnet haben, verkehrt. Jetzt steht Autonomie nicht mehr für Selbstgesetzgebung, sondern allein für die Steuerung eines Prinzips der Machbarkeit. Das Verständnis des Menschen, wie es in der christlich-philosophischen Tradition gesehen wurde, wird aufgegeben. Denn dieses weiß sich in seiner Verantwortung seiner Endlichkeit ebenso wie seinem Ursprung verpflichtet. Beides entzieht sich einem Verständnis von Autonomie eines autonomen Systems, das aus einer technischen Machbarkeit resultiert. Wird so der Begriff der Autonomie seines freiheits- und verantwortungsbestimmenden Kerns beraubt, wird er zum Wegweiser einer Gesellschaft, die sich nicht mehr als eine Formation aus freien Individuen, sondern aus funktionsgesteuerten Systemen versteht.

Von der Selbstbestimmung zur funktionalen Systemsteuerung

5. Kann es eine Maschinenethik geben?

Die für das Zeitalter autonomer Systeme naheliegende Frage, ob es eine Maschinenethik geben kann, rückt daher in greifbare Nähe.¹¹ Das Neue einer solchen Ethik bestünde darin, dass das Subjekt der Ethik gewechselt hat. Dieses wäre nicht mehr der Mensch, sondern die Maschine. Können programmierte Systeme, die nach moralischen Regeln, Sitten und Gebräuchen handeln, tatsächlich als ethische Subjekte gelten?

Die öffentliche Diskussion über ethische Fragen wie die, wie sich ein selbstfahrendes Auto im Fall eines unvermeidbaren Unfalls verhalten soll, zeigt die Problematik. Die Forderung allgemeinverbindlicher Regeln für das Verhalten eines selbstfahrenden Autos, etwa bei einem Unfall, wird nur von eingeschränkter Reichweite sein können. Übersehen wird bei solchen Vorstellungen der Programmierung moralischen Verhaltens, dass solche nicht allein nach Mustern im Sinne von Sitten, Gebräuchen und Konventionen verallgemeinert werden können. Verantwortliche Entscheidungen müssen als zeit-, situations- und erlebnisorientierte Entscheidungen wahrgenommen werden. Die vom Altbundesverfassungsrichter Udo Di Fabio geleitete Kommission über ethische Fragen des autonomen Fahrens konstatiert daher auch zu Recht: Entscheidungen wie die von Leben gegen Leben seien „nicht eindeutig normierbar und auch nicht ethisch zweifelsfrei programmierbar“. Damit hat man sich in der Ethikkommission „automatisiertes und vernetztes Fahren“¹² klar nach einer Ethik Kants im Sinne der Selbstzwecklichkeit des Menschen ausgerichtet. Sie sperrt sich gegen Abwägungen unter dem Gesichtspunkt des Nutzens und damit gegen die utilitaristische Programmierungsmöglichkeit.

Wenn das aber so ist, dann stellt sich weniger die Frage, ob man Moral an eine Maschine abgeben möchte, hier an das selbstfahrende Auto, sondern wie mit der Tatsache umzugehen ist, dass menschliche Moral nicht „maschinisierbar“ ist. Denn diese bleibt immer an eine Selbstbestimmung im Sinne der Selbstzwecklichkeit des Subjekts gebunden. Daher wird eine Maschinenethik in starkem Sinne, die das Abgeben der Verantwortung an eine Maschine bedeutet, in Frage zu stellen sein. Ein verantwortliches Subjekt ist gerade dadurch von jeder Künstlichen Intelligenz unterschieden – selbst wenn diese selbstlernende Algorithmen vorsieht –, dass es eben nicht künstlich, d. h. durch Poiesis hervorgebracht ist. Es hat seine Wurzel in der Praxis, die auf Freiheit basiert.

Die aristotelische Unterscheidung von Poiesis und Praxis gewinnt daher angesichts der Herausforderung der Ethik durch autonome Systeme erneut an Gewicht. Denn Verantwortung entzieht sich der Herstellbarkeit. Das heißt letztlich, dass von Verantwortung autonomer Systeme nicht gesprochen werden kann im Sinne der Übertragung der Verantwortung an das System. Autonome Systeme bedeuten dann lediglich eine Steigerung des Handlungsspielraums des Menschen. Mit diesem geht eine Steigerung der Verantwortungspflicht einher, aber nicht, wie es eine Maschinisierung der Moral vermuten lassen könnte, deren Abgabe selbst an das System. Autonom ist das System dann allenfalls in seiner Funktion der Selbststeuerung, nicht aber als Ausdruck von deren Freiheit. Diese Freiheit bleibt jedoch Voraussetzung von Verantwortung. Diese aber ist nicht machbar. Menschliche Verantwortung bleibt es daher, die Freiheit nicht selbst einer Machbarkeit unterzuordnen, sondern alle Machbarkeit in Freiheit zu gestalten. Autonomie im Rahmen von Künstlicher Intelligenz und solcher im Rahmen von Personen lassen sich daher nicht aufeinander abbilden, denn die Autonomie des Menschseins ist – wie seine Freiheit – nicht machbar.

Eine Künstliche Intelligenz, die einem Transhumanismus zuarbeitet, wird insofern nicht nur den Menschen zur Abhängigkeit verdammen, sondern sie droht, das Verständnis von Freiheit des Menschen in seinem Verhältnis zur Technik und Verantwortung zu verkehren. Insgesamt gilt es daher zu überlegen, wie eine Künstliche Intelligenz zu entwickeln ist, und

Ethische Entscheidungen programmierbar?

Zweifel an der Abgabe von Verantwortung an eine Maschine

Verantwortung für autonome Systeme?

ob wir eine schwache oder eine starke Künstliche Intelligenz wollen. Unter einer starken Intelligenz, die aber nach wie vor noch ein unrealistisches Zukunftsszenarium darstellt, sind autonome Systeme zu verstehen, bei denen Steuerung und Entscheidung dem System allein obliegt. Bei einer schwachen Intelligenz hingegen wird die Steuerung und Entscheidung weiterhin dem Menschen in Letztinstanz überlassen. Für eine starke Künstliche Intelligenz gilt nicht nur, dass in ihrem Rahmen die Ethik hinterherhinkt, sondern dass sich die Ethik nicht mehr von der Technik unterscheiden lassen wird. Freiheit versteht sich dann allein als Machbarkeit. Das aber bedeutet in der Tat die Erschaffung einer neuen Spezies, die in Konkurrenz zum Menschen träte. Sie ließe sich weder unter die Maschine noch unter das Humane einordnen. Sie wäre trans-human und sie wäre zu fürchten, denn sie strebte die Unterwerfung des Humanum an. Für eine schwache Künstliche Intelligenz bleibt das Verhältnis der Verantwortung und Steuerbarkeit der Technik klar. Es sind Menschen, die für die Technik, für die Programmierung intelligenter Systeme verantwortlich sind. Insofern eignet solchen Systemen dann allenfalls eine Teilautonomie.

Die große Herausforderung ist es daher, die Bestimmungen unserer eigenen Freiheit zu reflektieren und die in ihr zu Tage tretende Ambivalenz, die immer auch das Verhältnis von Technik und Ethik betrifft, im Gleichgewicht zu halten. Schließlich war es immer die Gefahr der Technik, dass sie durch Übersteigerung umzuschlagen droht, indem die vermehrte Freiheit, die die Technik schafft, in eine umso größere Abhängigkeit von ihr führt. Die schon bei Goethe im Zauberlehrling apostrophierten Überwältigungsmechanismen des Menschen durch die von ihm selbst entwickelte Technik scheinen sich daher auch mit der Künstlichen Intelligenz auf Dauer zu bewahrheiten. Gefordert ist daher eine Entscheidung, ob wir unsere Freiheit zu handeln behalten wollen oder nicht. Denn ist es zum Menschen 2.0¹³, zur Künstlichen Intelligenz als Singularität, gekommen, die das Menschsein umfasst und übergreift, dann können wir nicht mehr entscheiden, was wir als Menschen sein wollen. Denn dann ist das Wesentliche des Menschen, sein Entscheidungszentrum, okkupiert durch ein Programm, dessen Grundlage und letzter Kern die Programmierung selbst ist. Indem sie mittels Technologie selbst die Ursprünglichkeit des Menschen eliminiert, hat sie dem Menschen den Entscheidungsspielraum genommen. Sie lässt nur noch die Berechenbarkeit von Verhalten übrig. Damit verwirkt sie das eigentliche Potenzial menschlicher Freiheit, das in der Ambivalenz ihrer Spannungen menschliche Entscheidungen reifen lässt.

Notwendigkeit der
Entscheidung über
die Freiheit zur
Entscheidung

6. Autonomie zwischen Personsein und Maschine – autonome Systeme und ihre ethische Verantwortung

Bedeutet die Kritik an den modernen Technologien autonomer Systeme nun ein Plädoyer für einen Technikverzicht? Ein solcher Verzicht entspricht der antiken und christlichen Tradition. Das Christentum sieht den Menschen im Auftrag der Herrschaft über die Erde. Sie gilt es zu bebauen und zu bewahren. Mit dieser Kulturbestimmung ging ein Gleichheits- und Freiheitsversprechen einher, zu dem die Technik wesentlich hinzugehört. Gerade die Entwicklung dieser Kulturtechnik sollte der Inklusion der bisher aus der Kommunikation und ökonomischen Interaktion Ausgeschlossenen sowie der politischen Freiheit angesichts der Unterdrückung von Meinungsfreiheit dienen.

Technik steht im
Dienst des Menschen.

Zu diesem Kulturauftrag gehört nicht nur die Umsetzung von Können und Wissen in der Technik, sondern auch deren Verantwortung. Diese bleibt selbst dann allein dem Menschen übertragen, wenn er versucht, Elemente von Autonomie auf ein System zu übertragen. Selbst wenn in dieses System Selbststeuerungsregeln implementiert werden, liegt deren verantwortliche Steuerung weiterhin in der Hand des verantwortlichen Menschen. Dessen Verantwortung ist nicht programmierbar und daher auch nicht technisch machbar.

Verantwortung ist
nicht technisch
machbar.

Das intrinsische Problem einer Ethik der Technik, die sich angesichts der Fragen der Autonomie bewährt, entscheidet sich daher an der Frage der Machbarkeit bzw. der Reichweite des Steuerungspotentials. Verantwortung erfordert eine bleibende Unterscheidung von Technik und Ethik, bzw. Poiesis und Praxis, insofern die Ethik der Praxis und ihrer Freiheit, die Technik der Poiesis, der Kunst und der Machbarkeit, zugeordnet bleibt.

Auseinanderklaffen
von technischen
und ethischen
Entwicklungen

Eine angemessene Zuordnung erfordert die Klärung des ethischen Leistungsvermögens. Kritisch betrachtet werden muss hier das Auseinanderklaffen von technischen und ethischen Entwicklungen, d.h. das Voranschreiten der Technik, bei gleichzeitigem Hinterherhinken der ethischen Reflexionen dieser neuen Techniken. Die ethischen Begleitüberlegungen können in der Regel mit den neuen Errungenschaften der Technik nicht mithalten. Dementsprechend kommt – ein bereits klassisch gewordener Slogan – die Ethik immer zu spät.

Poetisch gefasst, in Anlehnung an Hegels Eule der Minerva, versuchte man bisher diese Aporie als zur Ethik gehörig zu bestätigen. Befriedigt hat diese Haltung aber nie – vor allem hat sie die Ethik in Misskredit ihrer möglichen Leistungsfähigkeit gebracht. Spätestens angesichts der sich immer beschleunigenden technischen Entwicklungen der Künstlichen Intelligenz kann eine solche Haltung nicht mehr ausreichen. Vielmehr ist angesichts der Vision eines Homo Deus eine proaktive ethische Begleitung technischer Entwicklung notwendig. Diese aber erfordert weiterhin die Unterscheidung von Ethik und Technik bzw. von Selbstbestimmung und Ordnung, wie sie der Begriff der Autonomie in der Tradition zusammengefasst hat.

Nach Kant muss Autonomie als Selbstbestimmung an der Einhaltung der Dialektik unter Berücksichtigung der gegenseitigen Interpretation dieser polaren Zusammengehörigkeit von Gesetz und Freiheit orientiert sein. Eine Gesetzgebung durch die Gesetzmäßigkeiten der sich selbst regulierenden Systeme, die ihre Steuerung nicht der Vernunft ihrer Autoren verdanken, sondern den in sie implementierten Steuerungsmechanismen, würde die technische Freiheit zum Maßstab auch der Ethik machen. Technik wäre dann sich selbst überlassen, da die für die Verantwortung notwendige Selbstreflexivität der Freiheit verloren ginge. Dieser Verlust aber hebt den Begriff der Verantwortung auf. Autonomie kann daher nie das Abgeben von Freiheit und Verantwortung an ein externes System bedeuten. Insofern kann es auch keine Maschinenethik geben.

Autonomie kann nicht
das Abgeben von
Freiheit bedeuten.

Die Akzeptanz autonomer Systeme lässt sich daher nicht ohne die Frage der Autonomie des Menschen vornehmen. Diese aber bleibt an die Anerkennung ihrer Grenzen gebunden. Wie bei der Gen-, Kern- und Informationstechnologie wird auch bei der Technologie autonomer Systeme dann eine Grenze überschritten werden, wenn der Autonomiebegriff seiner Dialektik von Gesetzgebung und Freiheit verlustig geht. Diese Gefährdung besteht in einem Umschlag der Autonomie auf eine Totalität der Freiheit ebenso wie auf eine Totalität der Gesetzgebung. Beiden Gefahren ist das autonome System ausgesetzt, indem hier entweder von der Machbarkeit der Autonomie oder von der Selbstgenügsamkeit des Systems ausgegangen wird. Eine solche Autonomie wäre keinem Selbst verpflichtet und keinem Gesetz. Sie drückt sich allein in jenen Steigerungsformen des Wissens und des Könnens aus, ja der Machbarkeit des Menschen, die dann in der Tat zur „letzten Erfindung des Menschen“, zum Transhumanismus, führten.

Transhumanismus
entspricht nicht dem
christlich-antiken
Verständnis des
Menschen.

Hingegen bleibt es Ausweis des Verständnisses jenes Menschen, der sich als freies autonomes Subjekt im Sinne der christlich-antiken Tradition versteht, im Ursprung und am Ende unverfügbar zu sein, und gerade dies ist Ausweis seiner Verantwortungsfähigkeit.

- 1 Vgl. Karl-Heinz Pantke (Hg.), *Mensch und Maschine: Wie Brain-Computer-Interfaces und andere Innovationen gelähmten Menschen kommunizieren helfen*, Frankfurt 2010.
- 2 Lucienne Rey, *Unser Freund der Roboter*, Zentrum für Technologiefolgen-Abschätzung (Hg.), Bern 2012. https://www.ta-swiss.ch/RoboComFactsheet_d.pdf (letzter Aufruf 06.09.2018).
- 3 Vgl. Robin Hanson, *The Age of Em: Work, Love and Life when Robots Rule the Earth*, Oxford 2016.
- 4 Vgl. Susanne Beck (Hg.) *Jenseits von Mensch und Maschine: ethische und rechtliche Fragen zum Umgang mit Robotern, Künstlicher Intelligenz und Cyborgs*, Baden Baden, 2012; Jan Schuhr, *Willensfreiheit, Roboter und Auswahlaxiom*, in: ebd., 43–75. Siehe auch Anmerkung 12.
- 5 S. dazu Herodot, *Historia* I, 96; vgl. Rosemarie Pohlmann, *Autonomie*, in: Ritter et al. (Hg.): *Historisches Wörterbuch der Philosophie* Bd. I, Darmstadt 1971, 701 ff.
- 6 (Herodot 7,103). Spätestens bei Sophokles wird mit der Figur der Antigone – allerdings kritisch – ein solcher auf das Individuum bezogene Begriff der Autonomie sichtbar. „Du lebst nach eigenen Gesetzen, drum allein ins Land der Toten gehst du.“ (Sophokles, *Antigone*, Vers 821).
- 7 Immanuel Kant, *Grundlegung zur Metaphysik der Sitten* in: Immanuel Kant, *Gesammelte Schriften* (= Akademieausgabe), Abt. I Bd. 4, Berlin 1911, 450. Vgl. dazu auch Elisabeth Gräb-Schmidt, *Autonomie*, in: ZEE 59 (2/2015), Gütersloh, 133–137.
- 8 Vgl. Michio Kaku: *Die Physik des Bewusstseins. Über die Zukunft des Geistes*, Reinbek 2014.
- 9 Ray Kurzweil, *Menschheit 2.0. Die Singularität naht*, Berlin 2014.
- 10 Vgl. Arne Manzeschke, Fabian Karsch (Hg.), *Roboter, Computer und Hybride: Was ereignet sich zwischen Menschen und Maschinen?*, Baden-Baden 2016.
- 11 Vgl. Catrin Misselhorn, *Maschinenethik und „Artificial Morality“: Können und sollen Maschinen moralisch handeln?*, in: *Künstliche Intelligenz, Aus Politik und Zeitgeschichte (APuZ)*, Februar 2018, 29–33 ; Oliver Brendel, *Überlegungen zur Disziplin der Maschinenethik*, in: *Künstliche Intelligenz, Aus Politik und Zeitgeschichte (APuZ)*, Februar 2018, 34–38.
- 12 Ethik-Kommission, *Autonomes und Vernetztes Fahren*, Bericht Juni 2017. Bundesministerium für Verkehr und digitale Infrastruktur, Berlin 2017.
- 13 Vgl. Ray Kurzweil, a. a. O. Anm. 9.

Impressum

Die Autorin:

Prof. Dr. theol. Elisabeth Gräß-Schmidt hat den Lehrstuhl für Systematische Theologie und Ethik inne und ist Direktorin des Instituts für Ethik an der Evangelisch-theologischen Fakultät der Eberhard Karls Universität Tübingen. Sie ist Mitglied im Wissenschaftsrat des Internationalen Zentrums für Ethik in den Wissenschaften (IZEW) und Mitglied in der Zentralen Ethikkommission der Bundesärztekammer (ZEKO). Darüber hinaus ist sie Mitglied im Rat der Evangelischen Kirche in Deutschland (EKD) und Vorsitzende des Evangelischen Hochschulbeirats der EKD.

Konrad-Adenauer-Stiftung e. V.

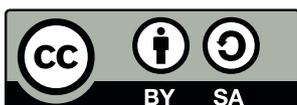
Dr. Karlies Abmeier

Leiterin Team Religions-, Integrations- und Familienpolitik
Hauptabteilung Politik und Beratung
T +49 30 / 26 99 63-374
karlies.abmeier@kas.de

Postanschrift: Konrad-Adenauer-Stiftung, 10907 Berlin

Herausgeberin: Konrad-Adenauer-Stiftung e. V. 2018, Sankt Augustin/Berlin
Gestaltung & Satz: yellow too Pasiek Horntrich GbR

ISBN 978-3-95721-474-4



Der Text dieses Werkes ist lizenziert unter den Bedingungen von „Creative Commons Namensnennung-Weitergabe unter gleichen Bedingungen 4.0 international“, CC BY-SA 4.0 (abrufbar unter: <https://creativecommons.org/licenses/by-sa/4.0/legalcode.de>).

Bildvermerk Titelseite
© Syda Productions/AdobeStock