

Katja Stepec

## EINLEITUNG TEIL 1:

### DER BEGRIFF DER INTELLIGENZ

Der vorliegende Sammelband hat den Titel „Künstliche Intelligenz – die große Verheißung“. Mögliche Konsequenzen und Bewertungen der künstlichen Intelligenz bedürfen einer Auseinandersetzung, was der Begriff der Intelligenz – und der künstlichen Intelligenz – im Allgemeinen und für uns bedeutet. Die Aufsätze im ersten Teil des Sammelbandes wollen genau dies: die unterschiedlichen Bedeutungen des Begriffs der Intelligenz und der künstlichen Intelligenz beleuchten.

Hilfreich für eine begriffliche Analyse ist die Aufschlüsselung der Voraussetzungen und der Konsequenzen des Begriffsgebrauchs. In der analytischen Sprachphilosophie wird auch von der inferentiellen Rolle eines Satzes gesprochen. Diese umfasst die Umstände sowie die Folgen einer Behauptung, also die Menge ihrer Prämissen und Konsequenzen. (vgl. z.B. Brandom 2000) Entsprechend ergibt sich die Bedeutung des Begriffs ‚Intelligenz‘ in Sätzen aus den Prämissen, unter denen man von Intelligenz sprechen kann, und die Folgen, die sich daraus ergeben. In Abhängigkeit der unterschiedlichen Voraussetzung und Konsequenzen kann ein Begriff unterschiedliche Gehalte aufweisen.

Ich möchte drei Bereiche von Voraussetzungen vorstellen, die alltagssprachlich mit Intelligenz verbunden werden und jeweils bestimmte Voraussetzungen von Intelligenz beinhalten: analytische Intelligenz, soziale Intelligenz und emotionale Intelligenz. Den drei Konzepten ist gemeinsam, dass Intelligenz als Fähigkeit verstanden wird. Meine Beschreibung der Voraussetzungen der drei Bereiche ist exemplarisch und keinesfalls abschließend, wie sich später auch aus den Beiträgen der Kapitel noch ergibt.

#### **Drei exemplarische Voraussetzungsbereiche des Begriffs der Intelligenz**

*Analytische Intelligenz* ist die Fähigkeit, ein Problem zu lösen. Voraussetzungen dafür ist die Aufnahme von Informationen, deren Auswertung für eine Fragestellung und das Liefern eines Ergebnisses. Der Prozess beruht auf regelgeleiteten Verarbeitungsschritten, die sich z.B.

als Schlussfolgerungen oder Algorithmen darstellen. Analytische Intelligenz kann auch bereits beinhalten, dass ein Wesen mit Störungen umgehen und unerwartete Hindernisse bewältigen kann. Diese Art der Intelligenz ist zielgerichtet und ergebnisorientiert und erfordert kein Verstehen der Aufgabe und deren Lösung.

Ein zweiter Bereich fügt den Voraussetzungen für Intelligenz soziale Aspekte hinzu. *Soziale Intelligenz* umfasst generell die Kooperation mit anderen Wesen, worunter Kommunikation und gemeinsames Handeln verstanden werden kann. Diese Art der Intelligenz ist ebenfalls ergebnisorientiert, aber auch anspruchsvoller, denn sie umfasst nun die Abstimmung mit anderen Wesen zur Erreichung eines gemeinsamen Ziels. Für die Zielerreichung ist daher nicht mehr nur eine Information über den Sachverhalt relevant, sondern auch Informationen über die Kooperationspartner: z.B. ihre Überzeugungen. Entsprechend ist Voraussetzung der sozialen Intelligenz nicht nur die reine Erfassung von Informationen im Sinne eines Inputs, sondern Erfassung im Sinne eines Austauschs von Informationen. Dafür ist Kommunikation notwendig. In der sozialen Interaktion erfolgt Kommunikation auf einer ‚öffentlichen‘ Ebene, also auf Basis gemeinsamer, standardisierter Normen, die sprachlicher, logischer oder mathematischer Art sein können. Kommunikation impliziert zumindest ein basales Verstehen dessen, was man tut. Für eine kommunikative Abstimmung mit Kooperationspartnern reicht es nicht aus, die internen Prozesse zu beschreiben, man muss sie auch mit den Prozessen der Beteiligten vergleichen können. Die Einnahme einer zweiten Perspektive, das Erkennen von Gemeinsamkeiten und Unterschieden sind Momente des Verstehens. Aus der Kommunikation und dem gegenseitigen Abgleich im Sinne eines Verstehens folgt eine koordinierte Aktion, die man unter Umständen bereits als gemeinsames Handeln bezeichnen kann. Voraussetzungen für soziale Intelligenz beinhalten also, dass das intelligente Wesen die Überzeugungen der anderen Wesen berücksichtigen und mit den eigenen abgleichen kann; dass es die eigene Aktion mit den Aktionen der anderen Wesen zu einem gemeinsamen Handeln koordinieren kann. Soziale Kooperation ist komplexer als das individuelle Lösen einer Aufgabe, erfordert daher mehr und andere Ressourcen, bietet aber möglicherweise auch schnellere und bessere Ergebnisse.

Schließlich können emotionale Voraussetzungen angenommen werden, die für die Qualifikation von Intelligenz relevant sind: *emotionale Intelligenz* beinhaltet das Erkennen und Bewerten von Gefühlen und Stimmungen und zwar bei sich und bei anderen. Emotionen können eine Problemlösung behindern oder beschleunigen. Unsicherheit wirkt

sich beispielsweise störend auf die Erreichung eines Ziels aus; Neugierde und die Aussicht auf Belohnung wirken sich positiv auf Problemlösungen aus. Emotionale Wesen mit der Fähigkeit, Gefühle zu reflektieren, zu bewerten und zu manipulieren, können diese zu ihrem Vorteil bei der Bearbeitung von Aufgaben einsetzen; Wesen, die darüber hinaus auch ihre Kooperationspartner positiv emotional beeinflussen, sind noch erfolgreicher. Die Kommunikation erfolgt hier in der Regel nicht auf einer öffentlichen oder expliziten Ebene, sondern bleibt subjektiv und implizit. Dazu gehören Fähigkeiten wie das Wahrnehmen von Stimmungen oder das subtile ‚Anstupsen‘ durch sprachliche Instrumente. Auf dieser Ebene kann zudem ein tieferes Verstehen erwartet werden: ein Verstehen nicht nur der Aufgabe und der Prozesse, die zu ihrer Erledigung notwendig sind, sondern auch ein Verstehen der eigenen und fremden Subjektivität.

Für die hier vorgeschlagene begriffliche Struktur verstehe ich diese Bereiche von Voraussetzungen als aufeinander aufbauend. Die Voraussetzungen aus dem analytischen Bereich gehören zum Kernbereich von Intelligenz. Ein Sprecher, der den Ausdruck Intelligenz verwendet, sollte diese Voraussetzungen immer akzeptieren. Die Voraussetzungen aus dem sozialen Bereich beinhalten zusätzliche Anforderungen aufgrund der (öffentlichen) Interaktion und Kommunikation mit anderen Wesen. Entsprechend erweitert sich die Bedeutung des Begriffs Intelligenz um die Begriffe der Sprache, des Verstehens und des gemeinsamen Handelns. Darauf aufbauend bringt der dritte Bereich subjektive Voraussetzungen in die Bedeutung von Intelligenz ein, und erweitert damit den Begriff zusätzlich um (implizite) emotionale Aspekte, sowie um ein weitergehendes reflektierendes Verstehen.

### **Konsequenzen des Begriffsgebrauchs: Mensch oder Maschine?**

Die Bedeutung des Begriffs Intelligenz hängt nicht nur davon ab, welche Voraussetzungen in den Begriff einfließen, sondern auch von den Konsequenzen, die daraus gezogen werden. Nicht aus jedem Voraussetzungsbereich dürfen die gleichen Schlussfolgerungen gezogen werden. Ich konzentriere mich hier auf zwei sehr allgemeine Schlussfolgerungen: Mensch oder Maschine.

In einer ersten Intuition scheint die Schlussfolgerung aus allen drei Bereichen auf den Menschen hinauszulaufen. Menschen wird sowohl die Fähigkeit zur Erfassung und Verarbeitung von Informationen zugeschrieben, als auch die Fähigkeiten zur Kommunikation und zum

gemeinsamen Handeln und schließlich zum Erkennen eigener und fremder Emotionen. Mit der Schlussfolgerung auf den Menschen sind weitere Konsequenzen möglich: z.B. die Annahme von Selbstbewusstsein und freiem Willen.

Für die Schlussfolgerung auf eine Maschine dagegen reichen zunächst die Voraussetzungen der analytischen Intelligenz aus. Eine solche problemlösende Maschine kommt dem nahe, was wir gemeinhin unter einem Computer verstehen: Eine Fragestellung wird durch systemimmanente Algorithmen bearbeitet und generiert eine Lösung. Hier sind weder Schlussfolgerungen auf Bewusstsein noch auf freien Willen angebracht, sondern das Vorgehen wird als rein instrumentell betrachtet.

Mensch und Maschine sind jeweils die extremen Enden einer graduellen Reihe von Schlussfolgerungen mit Zwischenpositionen, wobei sich bestimmte Kombinationen gängigen philosophischen Positionen zuordnen lassen. So weist eine *Kombination der analytischen mit der sozialen und der emotionalen Intelligenz* auf die These der starken KI hin, und erlaubt damit die Schlussfolgerung auf eine Gleichsetzung von Mensch und Maschine. Die These geht davon aus, dass Maschinen eine allgemeine Intelligenz aufweisen können, die der des Menschen gleicht oder sie übertrifft. Mit dieser These wird oft die Behauptung verbunden, dass Maschinen denken oder ein Bewusstsein haben. Eine solche Gleichsetzung von Mensch und Maschine setzt voraus, dass Maschinen auch soziale und emotionale Fähigkeiten haben können. In dem Aufsatz *“What is it like to be a Bat”* (Nagel 1974) beschreibt Nagel die subjektive Qualität als das Gefühl, wie es ist, in einem bestimmten mentalen Zustand zu sein. Um diese Art der Subjektivität nachempfinden zu können, setzt Nagel voraus, dass nur ausreichend ähnliche Wesen einen Standpunkt gegenseitigen Verstehens einnehmen können. Wesen oder Entitäten, die also alle drei Voraussetzungen erfüllen, lassen die Schlussfolgerung zu, dass es sich um einen Menschen oder ein menschenähnliches Wesen handelt, unabhängig davon, ob dieses Wesen organisch, auf Silikonbasis oder durch Algorithmen realisiert ist.

Eine *Kombination von analytischer und sozialer Intelligenz* würde dagegen in vielen Fällen ausreichen für die Schlussfolgerung, dass der Turing Test bestanden ist. In dem Artikel *“Computing Machinery and Intelligence”* (Turing 1950) schlägt Turing ein Imitationsspiel vor, in dem eine Person Fragen an zwei verborgene Kandidaten stellt und aufgrund der Antworten entscheiden muss, ob einer der beiden Kandidaten eine Maschine ist. Die Maschine besteht den Test, wenn sie so erfolgreich kommuniziert, dass die Person nicht bemerkt, dass es eine Maschine

ist. Eine solche Maschine weist sprachliche Fähigkeiten in einem begrenzten Kontext auf, wobei zusätzlich zu beachten ist, dass ein solches eingeschränktes Sprachspiel nicht immer auch schon Verstehen beinhaltet. Wird eine Maschine durch soziale Fähigkeiten angereichert, sind dennoch Schlussfolgerungen möglich, dass es sich um mehr als ein Instrument handelt.

Reduziert sich die Voraussetzung auf *analytische Fähigkeiten*, bleibt die Annahme instrumenteller Intelligenz, die beispielsweise durch Searles Beschreibung des chinesischen Zimmers illustriert wird (Searle 1980). Hier kombiniert ein in einem Zimmer befindlicher Mensch mit Hilfe eines Regelbuches Kärtchen mit chinesischen Zeichen und reagiert damit auf eine Frage in chinesischer Sprache. Aus Sicht des außerhalb stehenden Fragenden mag der Mensch eine zufriedenstellende Antwort gegeben haben, gucken wir jedoch in das chinesische Zimmer hinein, werden dort lediglich Zeichen gemäß einer formalen Regel, einer Syntax, angeordnet – ähnlich einem Computer, der Symbole gemäß einer mathematischen Regel, eines Algorithmus, kombiniert. Der Mensch ist weder der chinesischen Sprache mächtig, noch hat er ein Verständnis von Frage und Antwort.

Die hier hilfswise vorgeschlagene Struktur des Begriff der Intelligenz ist durch diese einfache Kombinatorik längst nicht umfassend beschrieben. So kann durchaus in Frage gestellt werden, ob es weitere Voraussetzungen gibt – z.B. aus dem Bereich der Kunst, Literatur oder Musik. Und auch die zunächst einleuchtende Behauptung eines Kernbereichs von Intelligenz kann bestritten werden. Beispielsweise ist ein Wesen vorstellbar, dem lediglich soziale und emotionale Intelligenz zugeschrieben werden kann – oder auch nur einer dieser Bereiche. Schließlich sind auch graduell unterschiedliche Kombinationen vorstellbar – beispielsweise ein Wesen ausgestattet mit geringfügigen analytischen Fähigkeiten, dafür aber mit überdurchschnittlicher sozialer und / oder emotionaler Intelligenz.

Im ersten Teil dieses Bandes stellen Autoren aus unterschiedlichen Bereichen ihre Überlegungen zum Begriff der Intelligenz, seinen Voraussetzungen und seinen Konsequenzen – also auf seine Anwendungsbereiche – vor.

Zu Beginn stellt der Aufsatz von **Daniel Dennett** "Turings seltsame Umkehrung der Argumentation. Was uns Darwins Evolutionstheorie über Künstliche Intelligenz verrät." die provokante Frage, ob wir nicht die Perspektive wechseln sollten, wenn es um das Verhältnis von Mensch und Maschine geht. Vielleicht nehmen wir intuitiv an, dass so etwas wie Intelligenz oder Bewusstsein nur durch eine höhere Intelli-

genz oder ein überlegeneres Bewusstsein in die Welt kommen kann. Mit Hilfe einer Analogie zwischen den Thesen Darwins und Turings zeigt Dennett allerdings auf anschauliche Weise, dass fortgeschrittene Organismen oder Fähigkeiten nicht notwendig auf einem spezifischen und sehr komplexen Funktionsschema beruhen müssen, sondern stattdessen in vielen Fällen auf sehr einfachen Prozessen aufbauen. Es ist das Prinzip der einfachen Bausteine, welches nicht nur das Leben und die Evolution erklärt, sondern für Turing auch das menschliche Denken. Konstruiert man auf dieser Basis eine intelligente Maschine, zeigt sich eine Art der ‚Pseudo-Intelligenz‘ oder ein ‚Als-ob‘-Verständnis, welches zumindest als Annäherung an ‚echte‘ künstliche Intelligenz oder Verständnis gewertet werden kann. Dennett drückt damit eine mögliche Erwartung an die Entwicklung der künstlichen Intelligenz aus.

Eine instrumentelle ‚Als-ob‘-Perspektive schlagen auch Schumann und Du vor. Der Aufsatz „Grenzgänge: Von Menschen zu smarten Maschinen – und zurück?“ von **Nadine Schumann & Yaoli Du** konzentriert sich auf die Genese der Schlussfolgerungen auf Mensch / Natur und Maschine / Technik. Der Aufsatz erzählt die Geschichte der Versuche, das eine über das andere zu verstehen. Ausgehend von einem ganzheitlichen Ansatz des Aristoteles, vollzieht sich nach Descartes eine Trennung mit der Tendenz, Natur unter technischen Voraussetzungen zu verstehen. Der Versuch einer Formalisierung der Mathematik durch die Logik und die allgemeine Systemtheorie führen zur Analogie des menschlichen Denkens mit den funktionalen Berechnungen am Modell der Turingmaschine. Auf Basis dieses Erbes kämpft die aktuelle Forschung mit einer *petitio principii*: Zum einen soll der reduktionistische Intelligenzbegriff in der KI-Forschung menschliche Fähigkeiten nachahmend modellieren, zum anderen werden Modelle der KI-Forschung zum Verständnis menschlicher kognitiver Prozesse herangezogen. Am Ende wird eine Sichtweise vorgeschlagen, die nicht auf die Reduktion von Mensch auf Maschine oder des Denkens auf technische Informationsprozesse hinausläuft, sondern mit einer ganzheitlichen ‚Als-ob‘-Betrachtung ein Simulationsmodell des menschlichen Zusammenlebens einbringt, mit dem auch menschliche Subjektivität erfasst werden kann.

Während Schumann und Du einen Weg in Richtung einer sozialen Intelligenz von Maschinen aufzeigen, diskutiert der Aufsatz „Künstliche Subjektivität? – Eine Analyse anhand der neurowissenschaftlichen Subjektbeschreibung Antonio Damasio“ von **Hans Zillmann** die grundsätzliche Möglichkeit einer sozialen und emotionalen Intelligenz

von Maschinen. Zillmann vertritt dabei die These, dass Subjektivität immer eine Frage der sozialen und kulturellen Zuschreibung ist. Auf Grundlage des naturwissenschaftlich orientierten Ansatzes Damasio wird Subjektivität als Auseinandersetzung des Gehirns mit dem eigenen Körper und den Objekten der Umwelt beschrieben und von Zillmann auf zwei Voraussetzungen von Subjektivität konzentriert. Für die erste Voraussetzung unterscheidet Zillmann zwischen Emotionen als physiologischen Reaktionen des Organismus, und Gefühlen als qualitativen Bewertungszustände. Subjektivität besteht entsprechend darin, dass Wesen die emotionale Erfassung ihrer Umwelt über Gefühle wahrnehmen. Zweitens steht diese Art der Subjektivität in einem kulturhistorischen Kontext und dieser Bereich ist von einer physikalistischen Sichtweise nur schwer zu erfassen. Auch wenn Menschen dazu tendieren, Maschinen Subjektivität zuzuschreiben, und darin sogar Gefahren sehen, bleibt die Vermutung, dass Maschinen eher auf das deduktive Schließen beschränkt sein werden und damit auf den Bereich der analytischen Intelligenz.

Emotionale Intelligenz muss nicht die stärkste Voraussetzung für Intelligenz sein. In dem Beitrag „Die Frage nach der Weltoffenheitmaschine. Drei Problemstellen für die Erzeugung einer künstlichen Intelligenz“ von **Michael Meyer** werden drei weitere Voraussetzungen präsentiert, die dem Begriff der Intelligenz eine anspruchsvolle existenzialistische Bedeutung geben. Auf Basis des Ansatzes McDowells wird Intelligenz vorbereitend als Ko-Konstitution von Denken und Wahrnehmen erläutert. In einem ersten Schritt wird dann diese epistemische Perspektive durch den phänomenologischen Gedanken von Hermann Schmitz im Sinne einer körperlichen Kommunikation mit der Welt angereichert. Mit Hilfe der hermeneutischen Perspektive Gadamers zeigt sich die zeitliche und geschichtliche Dimension der menschlichen Reflexivität, die der vorwärts gerichteten Linearität des algorithmischen Problemlösens ein interagierendes Gespräch mit der Tradition gegenüberstellt. Beides mündet in den Begriff der Weltoffenheit als eine leiblich erfahrene und geschichtliche Reflexion der Intelligenz. Meyer fügt mit Leiblichkeit und sinnhafter Geschichtlichkeit sowie existentieller Weltoffenheit dem Begriff von Intelligenz nicht nur weitere wichtige Voraussetzungen hinzu, sondern stellt am Ende auch die provozierende Frage nach ethischen Bedenken bei der Erschaffung von künstlichen Intelligenzen.

Eine wichtige Voraussetzung für soziale Intelligenz diskutiert der Beitrag „Künstliche Intelligenz als Kommunikationsproblem“ von **Jan Tobias Fuhrmann**. Er argumentiert dafür, dass Kommunikation nicht

notwendig mit Bewusstsein und Verstehen verbunden sein muss; Intelligenz zeigt sich stattdessen darin, dass die Kommunikation am Laufen gehalten wird. Kern der Kommunikation ist eine strukturelle und sinnhafte Unbestimmtheit, die wesentlich die Möglichkeit kommunikativer Anschlüsse beinhaltet. So sind auch durch Algorithmen gesteuerte Bots, d.h. quasi-umgangssprachliche Kommunikationsprogramme, in der Lage, eine Unterhaltung zu führen und innerhalb eines thematisch begrenzten Bereichs den Turing-Test zu bestehen, ohne dass sie Sinn konstituieren oder sich selbst reflektieren. Tatsächlich stellt sich die fehlende Reflexion als Effizienzvorteil bei der Verarbeitung großer Datenmengen heraus. Das überraschende Ergebnis der Überlegungen ist schließlich, dass gerade die Eigenschaft von Maschinen als nicht-originell und strukturell konservativ dazu führt, dass sie erfolgreich an Kommunikation teilnehmen und diese fortsetzen können.

Mit dem Begriff des Wissens schlägt der Beitrag „Verheißung, Verdammung oder einfach ein Selbstmissverständnis? Sprachkritische Überlegungen zum Umgang mit KI und ihren Beschreibungen“ von **Mathias Gutmann, Marie-Claire Haag und Christian Wadehul** einen weiteren alternativen und vielversprechenden Bereich von Voraussetzungen vor. Die Autoren untersuchen die begrifflichen Zusammenhänge zwischen Wissen, Information und Daten mit dem Ziel, einen tätigkeitstheoretischen Wissensbegriff zu etablieren, ein Wissen um den Zweck einer Verwendung. So wird der operative Wissensbegriff – wie funktioniert eine Maschine – abgegrenzt vom situativen, ‚umgänglichen‘ Wissen um den Zweck einer Maschine, was erst die Beurteilung ihrer Funktionsfähigkeit erlaubt. Unterstützt wird die Betrachtung durch die Unterscheidung eines Funktionierens ‚gemäß einer Regel‘ und dem Handeln ‚nach einer Regel‘ oder in Kenntnis der Regel. Auch der Informationsbegriff enthält bereits eine Zwecksetzung und unterscheidet sich damit von dem Begriff der bloßen Daten. Situatives Wissen aber kann schon in Richtung eines reflexiven Wissens gedeutet werden – nämlich als ein Wissen von sich selbst. Die Einordnung des Begriffs der künstlichen Intelligenz in das begriffliche Umfeld eines zweckorientierten Wissens vermeidet so zwar zunächst die problematische Schlussfolgerung auf Bewusstsein, führt aber zu nicht weniger anspruchsvollen Schlussfolgerungen auf selbständige Zwecksetzung, Rationalität und Verantwortung.

Der Beitrag „Künstliche Intelligenz und kognitive Leistung“ von **Christian Freksa** führt die Idee der zweckorientierten Verwendung als ökologisches Problemlösen weiter aus. Er legt den Fokus auf eine der Voraussetzungen für soziale Intelligenz: Handlungen als



situative Interaktion mit der Welt und Selektion von Informationen im Zusammenhang mit Problemlösungen. Im Bereich der analytischen Intelligenz sind Maschinen zunächst Instrumente zum Lösen von Problemen. Das dafür benötigte umfassende Wissen über die Welt stellt sich als eine Art digitaler Zwilling dar, welcher aus riesigen Datenmengen besteht, die formal und sequenziell organisiert sind. Es handelt sich um Repräsentationen oder Beschreibungen, die die Grundlage für ein Verstehen des Problems im Sinne des Nachvollzuges einzelner Funktionsschritte darstellen. Für die Lösung eines Problems ist die Kenntnis der einzelnen Funktionsschritte jedoch nicht hilfreich; ein handelnder Organismus dagegen erkennt die notwendigen Informationen in der Interaktion mit der Umwelt als "Affordanzen", d.h. als Verwendungsangebote. Für Freksa ist diese Art des 'ökologischen' Problemlösens, das sich in der handelnden Interaktion zeigt, eine Anforderung an künstliche Intelligenz, prägnant gefasst als Intelligenz pro Kilowatt, die deren bisheriges Primat, das logische Schließen auf Basis von Repräsentationen, in Frage stellt.

Logisches Schließen ist nicht nur die Grundlage für die Algorithmen, aus denen die Software einer KI besteht, sondern auch die Grundlage für die Mathematik. In dem Beitrag „Künstliche Intelligenz in mathematischen Beweisen und das Problem der Erklärbarkeit“ von **Daniel Wenz** wird daher einleitend gefragt, ob nicht die Nähe der künstlichen Intelligenz zur Mathematik diese für mathematische Beweise besonders qualifiziert (siehe hierzu auch den Beitrag von Dieter Mersch). Warum das nicht der Fall ist, weist auf ein aktuelles Problem der KI hin: der Erklärbarkeit oder besser Nicht-Erklärbarkeit. Komplexe mathematische Beweise müssen nicht nur formal nachvollzogen, sondern auch konzeptuell verstanden werden, indem der Beweis in einen größeren theoretischen Kontext eingeordnet wird. Damit wird nicht nur das eigentliche Problem, sondern auch der mathematische Kontext neu beleuchtet. Die historische Entwicklung der automatisierten Beweisführung dagegen zeigt eine Trennung der Verbindung von formaler und konzeptueller Beweisführung auf. Waren automatisierte Systeme mit dem Ziel der Beweisführung zunächst nach den Fähigkeiten des Menschen modelliert, und sollten primär Aufschluss über den menschlichen Geist geben, wurde diese Ähnlichkeit später zugunsten einer Problemorientierung aufgegeben, mit dem Ziel schnellerer und effizienterer Ergebnisse. Diese führen zu Beweisen, die für den Menschen nicht mehr verstehbar sind und somit auch nicht nutzbar zur Erweiterung mathematischer Kenntnisse. Mit Hilfe einer Art des expressiven Verstehens soll die konzeptuelle Ebene neu erschlossen werden. Dafür

wird die Interaktion zwischen Mensch und Maschine als gemeinsames Handeln gleichwertiger Akteure verstanden, das nicht nur von konstitutiven Regeln bestimmt wird, sondern in dem sich auch implizite Normen entwickeln. Dieses praktische *Know-How* der Teilnehmer, d.h. der konzeptuelle Überschuss, wird explizit gemacht, um undurchsichtige Beweisführung verstehbar und damit nutzbar zu machen.

Der Artikel „Künstliche Moral – Können Maschinen moralisch handeln?“ von **Christoph Merdes** verweist bereits auf die Thematik des zweiten Teils des Sammelbandes. Merdes bejaht nicht nur die Möglichkeit eines künstlichen moralischen Akteurs, sondern schlägt im Wesentlichen drei Kriterien vor, mit denen Moralität von Maschinen beurteilt werden kann. Als erste Bedingung für moralisches Handeln versteht der Autor die Freiheit von Handlungen, und kommt zu dem Ergebnis, dass Maschinen zumindest nicht als völlig fremdbestimmt aufzufassen sind. Auch die zweite Bedingung für moralisches Handeln, die moralischer Rechtfertigung, wird durch die Übertragung von maschinellem Lernen auf eine maschinelle Ethik teilweise bejaht. Mit einem angepassten Turing-Test wird schließlich die dritte Bedingung untersucht: moralische Kompetenz. Hier stellt sich als eigentliches Problem die Auswahl der Beurteilenden heraus – letztlich gibt es also keinen Maßstab für moralische Befähigung, auch der des Menschen selbst.

Zum Ende des ersten Teils wird das Moment der „Verheißung“ aufgebrochen in die Begriffe des Hypes und des Mythos. **Walther Ch. Zimmerli** beschäftigt sich in seinem Beitrag „Künstliche Intelligenz und postanaloges Menschsein. Entstehung, Entwicklung und Wirkung eines realen Mythos“ mit der geschichtlichen Entwicklung der Voraussetzungen für künstliche Intelligenz. Im Sinne der genealogischen Methode Nietzsches, durch Aufarbeitung der historischen Entwicklung eines Begriffs dessen Voraussetzungen zu beleuchten, erzählt Zimmerli die Geschichte der KI zwischen Mythos und Hype, also zwischen der versuchten Bändigung der Entwicklung mit Hilfe eines sinnstiftenden Narratives und dem Verselbständigen der Entwicklung im Sinne eines Hypes. Die ‚Erfindung‘ des Begriffs anlässlich der Dartmouth-Conference ist der Beginn des Mythos der ‚denkenden Maschine‘, der durch die Generierung überzogener Erwartungen auch bereits den ersten Hype auslöst. Die Fortsetzung des Hypes kulminiert im endzeitlichen Mythos der Singularität, also der Vorstellung, dass die künstliche die menschliche Intelligenz übertreffen wird. Diese eschatologische Grenze soll überwunden werden durch die Idee des ‚Mensch-Maschine-Tandems‘ was letztlich zu der anthropologischen Frage führt: Trans- oder Posthumanismus – oder: wie transformiert sich der Mensch zwi-

schen dem Analogen und dem Digitalen? So katapultiert uns der Hype der ersten Stunde in ein Zeitalter, in der wir die alten Mythen zugunsten einer neuen Konzeption aufgeben sollten: dem post-analogen Menschen.

Während die Beiträge im ersten Teil bestimmte Voraussetzungen für die Zuschreibung von Intelligenz im Allgemeinen und künstlicher Intelligenz im Speziellen untersuchen und damit das Spektrum der Bedeutung dieser Begriffe ausreizen, werden im zweiten Teil des Sammelbandes nicht nur die Konsequenzen ‚Mensch oder Maschine‘ betrachtet, sondern eine Vielzahl weiterer Folgen, die sich aus dem Begriffsgebrauch ergeben können.

