BORIS VOIGT

Musizierende Maschinen mit poetischem Gemüt
Was lässt sich von E.T.A. Hoffmanns Kritik an Automaten über Künstliche Intelligenz lernen?

E.T.A. Hoffmann ist ein herausragender Diagnostiker der Folgen, die sich aus einem unhinterfragten Umgang des Menschen mit maschineller Technologie ergeben. War es zu seiner Zeit die Mechanik, welche die industrielle Revolution vorantrieb, und auf gesellschaftlicher Ebene die mit den Stein-Hardenberg'schen Reformen in Preußen eingeführten effizienten bürokratischen Apparate, in denen der Mensch idealiter neutral und – quasi mechanisch – nach strikten Regeln agierte, so sind es heute die Veränderungen durch digitale Technologien, die wiederum sowohl das Verhältnis des Menschen zur materiellen Welt wie menschliches Handeln und gesellschaftliche Strukturen betreffen. Mit der Entstehung regelbasierter Bürokratien, mit ihren nicht zu leugnenden Vorzügen und Nachteilen kannte sich Hoffmann als preußischer Kammergerichtsrat aus, ebenso mit mechanischen Automaten, mit denen er sich intensiv auseinandersetzte. Als Komponist und Dichter konnte er ungebunden an praktische Erfordernisse mit abstrakten, quasi maschinenhaften Verfahren ästhetisch experimentieren und die möglichen Auswirkungen eines nicht reflektierten Umgangs mit Maschinen in Gestalt von Gedankenexperimenten ausloten, als die seine Erzählungen ohne weiteres verstanden werden können. Hoffmann als Diagnostiker auch für die Folgen einer nicht ausreichend reflektierten Nutzung von Künstlicher Intelligenz (KI) heranzuziehen bietet sich insofern an, als er sich vor allem auf die Konsequenzen konzentrierte, die sich im Umgang mit Automaten für das menschliche Denken, Wahrnehmen und Fühlen, für das Verständnis des menschlichen Geistes grundsätzlich ergeben können. Gerade das (Selbst-)Verständnis des menschlichen Geistes ist wesentlich von der Durchsetzung von KI betroffen. Wie bei Hoffmann mag im Folgenden die Sphäre von Musik und Kunst als gedankliches Experimentierfeld dienen, von dem ausgehend auf andere Bereiche menschlichen Denkens und Handelns extrapoliert werden kann.

Nun scheint es prima facie anachronistisch, einen vor zweihundert Jahren verstorbenen Autor zu KI zu befragen. Gegen den Vorbehalt, kein Romantiker habe die Entwicklung der digitalen Technologien auch nur erahnen können, daher gebe es für ein solches Unterfangen keine guten Gründe, lässt sich nicht leicht etwas einwenden. Zu bedenken ist jedoch, dass bestimmte Vorbedingungen für eine auch

das Denken und die Subjektivität erfassende Technologie im frühen 19. Jahrhundert längst gegeben waren. Schon Thomas Hobbes begriff Denken als Rechnen. Im Neujahrsbrief von 1697 erläuterte Gottfried Wilhelm Leibniz seinem Landesherrn Rudolf August Herzog von Braunschweig-Wolfenbüttel die Binärzahlen, denen er großes Gewicht für die Entwicklung des Wissens zumaß. Er betrachtete die Binärzahlen als einen elementaren Schritt für die Systematisierung des Wissens über Zahlen und Rechenvorgänge: »Omnibus. ex. nihilo. ducendis. SUFFICIT. UNUM [Um alles aus dem Nichts herzuleiten, GENÜGT EINES].«[1] 1748 veröffentlichte der französische Arzt Julien Offray de La Mettrie – freilich anonym – sein Traktat *L'homme machine*, in dem er Descartes' These von der Maschinenförmigkeit des Lebens auf das Geistige, die Seele, ausweitet. Der Mensch sei nichts weiter als eine ihre eigenen Triebfedern aufziehende Maschine, konstatiert er in metaphorischer Zuspitzung.[2] In all diesen Überlegungen kündigt sich eine Denkweise über Geist, Mensch und Maschine an, die für die Entwicklung von und den Umgang mit KI entscheidend ist. Wäre Hoffmann nicht mit 46 Jahren schon gestorben, hätte er Zeitgenosse der Entwicklung der ersten Rechenmaschine werden können, die die entscheidenden Merkmale eines Computers aufwies. An der programmierbaren »Analytical Engine«, die wie ein heutiger Computer über eine Eingabe, eine Arbeitsrecheneinheit, einen von dieser getrennten Speicher und eine Ausgabeeinheit verfügte, arbeitete der 15 Jahre nach Hoffmann geborene Mathematiker Charles Babbage seit den 1830er Jahren bis an sein Lebensende, ohne den Entwurf vollständig bauen lassen zu können. Wie inzwischen rekonstruiert wurde, hätte die beinahe hausgroße dampfmaschinengetriebene Maschine einwandfrei funktioniert.[3] Dass die Anwendung der »Analytical Engine« potentiell weit über den Bereich der Mathematik herausreichte, erkannte die mit Babbage zusammenarbeitende, von Maschinen faszinierte Mathematikerin Ada Lovelace. Es ließe sich ein Gespräch imaginieren, das Lovelace, Babbage und Hoffmann über Maschinen und ihre Auswirkungen auf das menschliche Leben, Denken und Empfinden geführt haben könnten. Vielleicht wären sie dabei auf den Gedanken gekommen, die Annahme zu diskutieren, es gäbe Maschinen, die den menschlichen Geist simulieren.

---

1 Gottfried Wilhelm Leibniz an Herzog Rudolf August, 2.1.1697, in: Gottfried Wilhelm Leibniz: *Sämtliche Schriften und Briefe*, Reihe 1: *Allgemeiner politischer und historischer Briefwechsel*, Bd. 13: *August 1696–April 1697*, hrsg. v. der Leibniz-Forschungsstelle Hannover der Akademie der Wissenschaften zu Göttingen beim Leibniz-Archiv der Gottfried-Wilhelm-Leibniz-Bibliothek Hannover (Bearb.: Gerda Utermöhlen / Sabine Sellschopp), Berlin 1987, S. 116–121, hier S. 119. **2** Julien Offray de La Mettrie: *L'homme machine. Die Maschine Mensch* (frz./dt.), übers. und hrsg. v. Claudia Becker, Hamburg 1990, S. 34f. **3** Charles Babbage: *Passages from the Life of a Philosopher*, London 1864, S. 112–141; Doron Swade: *The Difference Engine. Charles Babbage and the Quest to Build the First Computer*, New York 2001, S. 91–133, 155–171.

Romantik und Maschine, eine kritische Affinität

Gewöhnlich wird der Romantik eine gegenüber naturwissenschaftlich-technischer Rationalität mit ihrem Anspruch auf Naturbeherrschung ablehnende Haltung unterstellt. Die Romantik zeichne sich durch den Rückzug in die Subjektivität, durch Feindschaft der Rationalität gegenüber, durch Gegenwartsferne und Mittelaltersehnsucht aus. Doch beruht die zugegebenermaßen karikierend dargestellte Annahme von einer einfachen Entgegensetzung von Romantik und Technik auf einer Reihe von Irrtümern. Etwa dürfen die ästhetisch in Dichtung, Bild oder Musik umgesetzten Empfindungen angesichts der Natur nicht als Aussage über die Natur selbst missverstanden werden, ebensowenig waren den Romantikern technisches Denken und Konstruieren fremd. Friedrich von Hardenberg war als Bergwerksassessor – neben Jura in Jena hatte er in Freiberg Montanwissenschaften studiert und daher gründliche Kenntnisse der Mathematik, der Chemie und des Bergbaus erworben – unter anderem für die Erschließung von Braunkohlevorkommen in der Nähe von Profen zuständig. Diese Kenntnisse prägen deutlich auch seine literarischen und philosophischen Texte. So begreift er das Romantisieren als qualitatives Potenzieren des Niederen und Logarithmieren des Höheren, das Niedere wird in dieser Operation mit dem Höheren identifiziert. Das Niedere, Gewöhnliche ist – mit romantischem Blick betrachtet – das Höhere, Unbegreifliche. »Indem ich dem Gemeinen einen hohen Sinn, dem Gewöhnlichen ein geheimnißvolles Ansehn, dem Bekannten die Würde des Unbekannten, dem Endlichen einen unendlichen Schein gebe so romantisire ich es.«[4] Gewöhnliche Gegenstände beleihen die Romantiker mit Qualitäten, die ihnen nicht tatsächlich zukommen, in der Phantasie jedoch als wirklich erscheinen. Solcherweise wird die innere Vorstellungswelt auf äußere Gegebenheiten projiziert, als sei sie in diesen selbst vorhanden. Innerhalb von Dichtung, Malerei oder Musik ein legitimes Verfahren, wäre es außerhalb der Kunst schlicht ein vitium subreptionis, wie Kant es genannt hätte. Ein subjektiver Vorstellungsgehalt würde fälschlich als etwas Objektives, tatsächlich Gegebenes behauptet. An solcher Subreption partizipiert in erheblichem Ausmaß auch der heutige Umgang mit KI, angefangen mit der Bezeichnung selbst. Da sich das romantische Denken generell mit der Problematik der Subreption intensiv auseinandersetzt, und insbesondere E.T.A. Hoffmann sie am Verhältnis des Menschen zur Maschine darstellt, bieten sich dessen Überlegungen geradezu an, sie diagnostisch und versuchsweise auf die Thematik der KI anzuwenden.

---

4 Novalis: *Vorarbeiten zu verschiedenen Fragmentsammlungen*, in: ders.: *Werke, Tagebücher und Briefe Friedrich von Hardenbergs*, Bd. 2: *Das philosophisch-theoretische Werk*, hrsg. v. Hans-Joachim Mähl, München / Wien 1978, S. 334 (Nr. 105).

Im romantischen Denken besteht eine komplexe Affinität zwischen Maschine und Geist, die sich eminent in der Diskussion musikalischer Themen äußert. Bereits Wilhelm Heinrich Wackenroder führt die »wunderbare Bedeutsamkeit«, die der in Musik verwandelte Ton besitze, auf das »gelehrte System« zurück, zu dem der in der Natur nur als roher Stoff gegebene Schall geordnet ist. Durch die Verfeinerung innerhalb eines »kunstreichen Systems« seien Schall und Ton zum Abbild der »Verfeinerung und harmonischen Vervollkommnung des heutigen menschlichen Geistes« geworden.[5] »[Z]wischen den einzelnen mathematischen Tonverhältnissen und in den einzelnen Fibern des menschlichen Herzens« habe sich »eine unerklärliche Sympathie offenbart, wodurch die Tonkunst ein reichhaltiges und bildsames Maschinenwerk zur Abschilderung menschlicher Empfindungen geworden ist«.[6] Maschinenförmigkeit von Musik und Kunst steht nicht in Gegensatz zur Musikästhetik Wackenroders, allerdings versteht er die Wirkung der Musik auf das Gefühl nicht lediglich kausalmechanisch. Das musikalische Maschinenwerk funktioniere nicht gleich »einem künstlichen Weberstuhle für gewirkte Zeuge«. Ein köstliches Werk entstehe vielmehr dann, wenn »Gefühl und Wissenschaft so fest und unzertrennlich in einander hangen, wie in einem Schmelzgemählde Stein und Farben verkörpert sind«.[7] Gefühl und Rationalität bezeichnen zwei verschiedene Ebenen in der Musik, die Wackenroder offenbar als operativ geschlossene Systeme begreift. Das im Musikwerk abgeschilderte Gefühl könne nur vom inneren Gefühl erfasst und ergriffen werden. Freilich lässt Wackenroder die Gegenstände aus, auf die sich das Gefühl intentional richtet; die resultierende Schwierigkeit, nachfühlen oder bemerken zu können, welches Gefühl musikalisch abgeschildert werden soll, sei einmal hintangestellt. Zwar ist die Abschilderung des Gefühls im Musikwerk maschinenförmig hervorgebracht und insofern auch das innere eigene Gefühl Resultat der Mechanik. Der menschliche Geist, das Subjekt, setzt das äußerlich Mechanische jedoch in seine eigenen Kategorien um, die nicht auf das ihnen Äußere rückführbar sind.

> »[…] alle die tönenden Affekten werden vo[n] dem trocknen wissenschaftlichen Zahlensystem, wie von den seltsamen wunderkräftigen Beschwörungsformeln eines alten furchtbaren Zauberers, regiert und gelenkt. Ja, das System bringt, auf merkwürdige Weise, manche wunderbar neue Wendungen und Verwandlungen der Empfindungen hervor, wobey das Gemüth über sein eignes Wesen erstaunt […].«[8]

Aus romantischer Sicht ist Kunst – insbesondere Musik – ein Medium, in dem sich das Subjekt selbst begegnet und seine Innenwelt artikuliert; ein Medium, das ein Gegengewicht gegen die widerständige äußere Natur und die widerständige Alltagswelt bildet. Um dieses Potential der Kunst entfalten zu können, bedarf es aller-

5 Wilhelm Heinrich Wackenroder: *Das eigenthümliche innere Wesen der Tonkunst, und die Seelenlehre der heutigen Instrumentalmusik*, in: ders.: *Sämtliche Werke und Briefe*, Bd. 1: *Werke*, hrsg. v. Silvio Vietta, Heidelberg 1991, S. 216–223, hier S. 216. **6** Ebd., S. 216f. **7** Ebd., S. 217 u. 218. **8** Ebd., S. 220.

dings – dessen waren sich die Romantiker vollkommen bewusst – der rationalen Naturbeherrschung, der Wissenschaft und Technik. Die Übersetzung des rationalen Systems, dem in der Musik die physikalischen Schallereignisse unterworfen werden, in das Innere des Subjekts vollzieht sich demnach zwingend, jedoch entzieht sich die Wirkungsweise der rationalen Erklärung. Ähnlich fassen Novalis und Hoffmann die subjektive Umsetzung der kausalmechanischen Wirkungen von Sinneseindrücken; die Reizung der Nervenfasern – bei Hoffmann vorzugsweise mittels der Klänge einer Glasharmonika – bewirkt ihrer Auffassung nach das absonderlichste Erleben und die unheimlichsten Seelenzustände.

Die Rede vom Wunderbaren und von Beschwörungsformeln sollte nicht von der prinzipiell gültigen Einsicht ablenken, dass sich das in erster Person innegehabte Erlebnis – sei es vom Klang, von Farben oder anderen Eindrücken – weder in objektivierende Erklärungen in dritter Person übersetzen lässt noch selbst aus diesen erklärt werden kann. Außerdem ist – die Romantiker bleiben diesbezüglich vage – eine vollständige kausale Rückführung auf physikalische Vorgänge ebensowenig möglich. Was gehört wird, hängt nicht einfach von einer zwingenden Kraft ab, sondern von den Fertigkeiten eines Menschen, das zu Hörende in einer bestimmten Weise zu hören, auf bestimmte Aspekte des zu Hörenden die Aufmerksamkeit zu lenken, genauso wie die Offenheit für die im Klingenden angelegten Möglichkeiten, es zu hören.

Wichtig für die Frage nach dem Verhältnis des Menschen zur KI ist zunächst die Behauptung Wackenroders, nicht unmittelbar in der rohen, ungestalteten Natur finde der Mensch seine Entsprechung, sondern in dem von ihm selbst geschaffenen Maschinenwerk; in ihm gelange das Innere des Menschen zum Ausdruck oder begegne sich allererst selbst.[9] In der Technik begeistert sich der Mensch demnach vor allem für sich, indem sie in ihrer Zweckhaftigkeit auf ihn bezogen ist, seinen Handlungsspielraum erweitert und er in ihrer Gestalt seinen eigenen konstruktiven Fähigkeiten begegnet. Technik, zumal solche, deren Bewegungen oder Verrichtungen sich von den Handlungen des Menschen zu einem gewissen Grad verselbständigen, spiegelt ihm in veräußerlichter Gestalt etwas genuin ihm Zugehöriges zurück. Diesen Vorgang, in dem das Niedere, zu dem das mechanisch Äußerliche un-

9 Das romantische Denken schlägt damit eine Richtung ein, die sich im späteren 19. und vor allem im 20. Jahrhundert in den Grundüberlegungen der Technikanthropologie wiederfindet, so etwa in der von Ernst Kapp vertretenen These von der Technik als Organprojektion, nach welcher sich der Mensch in der Technik als einem Modell bestimmter Aspekte seiner selbst nicht zuletzt selbst vergegenwärtigt (*Grundlinien einer Philosophie der Technik. Zur Entstehungsgeschichte der Cultur aus neuen Gesichtspunkten*, Braunschweig 1877). Arnold Gehlen sieht in der Faszination durch die scheinbare Selbstbewegtheit von Automaten, den rhythmisch wiederholenden Ablauf ihrer Bewegungen ein weit ursprünglicheres Interesse an der Technik, als es die rationalen und praktischen Interessen an ihr sind. Das Faszinierende des Automatismus einer Maschine sei ganz unabhängig von der Leistung. In höchstem Grade sieht er die Idee vom Perpetuum mobile von dieser Faszination motiviert, der Idee von einer Maschine, die funktionslos ganz in der eigenen Selbstbewegtheit aufgeht (*Die Seele im technischen Zeitalter. Sozialpsychologische Probleme in der industriellen Gesellschaft*, hrsg. v. Karl-Siegbert Rehberg, Frankfurt a. M. 2007, S. 15f.).

zweifelhaft gehört, mit dem Höheren, dem Geistigen und Inneren identifiziert wird, spricht Novalis mit dem Verb des Romantisierens an. Jedoch überlassen sich die Romantiker nicht unvermittelt den artifiziellen Entsprechungen zu ihrem Inneren, sondern reflektieren zugleich auf sie und auf die Konsequenzen, die ein solches Sich-Überlassen für das eigene Selbst und das eigene Handeln in der Welt hätte. Kunst wird in der Romantik auch zu einem Medium, in welchem das Subjekt eine durchaus kritische Distanz zu sich, aber auch zur Alltagswelt aufbauen kann. Darin bestand die aufklärerische Kritik der Romantiker an der Aufklärung; behauptete diese vereinfacht gesagt das menschliche Subjekt als eine allem anderen zugrundeliegende Instanz, wiesen die Romantiker auf eine grundsätzlich bestehende Inkommensurabilität von Innen- und Außenwelt hin. Das Subjekt konstituiert nicht die Welt, es stößt in ihr – sowohl in der ›rohen Natur‹ wie in der Gesellschaft – auf Widerstand, findet sich in ihr nicht unmittelbar wieder. Ebensowenig fügen sich die Maschinen umstandslos dem Subjekt, das sie geschaffen hat oder sie anwendet.

Maschinelle Verfahren in der romantischen Ästhetik und den musikalischen und literarischen Werken E.T.A. Hoffmanns

Trotz seiner großen Begeisterung für Automaten untersuchte Hoffmann das Verhältnis des Menschen zu den mechanischen Maschinen äußerst kritisch. Die Drastik, mit der er etwa in den Erzählungen *Der Sandmann* oder *Die Automate* dieses Spannungsfeld durchdringt, verleiht seinen Überlegungen ein analytisches Potential, das für das Verstehen des menschlichen Umgangs mit KI aktualisiert werden kann. Dass Hoffmann in seiner Vorliebe für Automaten kritisch blieb, ist entscheidend; ohne die ambivalente Verbindung von Faszination und kritischem Bewusstsein – in der KI-Forschung ist eine solche Ambivalenz nicht ausnehmend weit verbreitet – hätte er die Probleme, die in dem Verhältnis zwischen Mensch und komplexer Maschine bestehen, nicht angemessen durchdenken und veranschaulichen können.

Hoffmann setzt sich zum einen inhaltlich mit den Automaten als Gegenstand auseinander, zum anderen reflektiert er sie in der formalen Gestaltung seiner literarisch-essayistischen Texte und musikalischen Kompositionen. Nicht allein sind die Folgen der Integration technischen Denkens in die Gesellschaft Gegenstand seiner Werke, ebenso bildet eine quasi technische Operationalität bereits ein Konstituens seiner Musik und seiner Texte und dient ihrer Konstruktion. Hoffmanns Kompositionen beruhen wesentlich auf kombinatorischen Verfahren. Darin allein liegt freilich nichts Ungewöhnliches. Bereits im frühen 18. Jahrhundert gehörte die mathematische Kombinatorik zum festen Bestandteil der Musiktheorie und -didaktik. Die abstrakte Systematisierung von kombinatorischen Möglichkeiten mit einer definierten Menge von Elementen, ganz gleich worin diese Elemente bestehen, mit-

tels mathematischer Operationen ist selbst musiktheoretischen Ursprungs. Marin Mersenne entwickelte die mathematische Kombinatorik in der *Harmonie universelle* (1636). Verbreitung im Bereich der Musik dürfte die von Mersenne geschaffene Kombinatorik vor allem über Athanasius Kirchers *Musurgia universalis* (1650) gefunden haben. Zu Hoffmanns Zeit unfraglich ins musikalische Denken integriert, erfüllte die Kombinatorik jedoch andere Funktionen als bei ihrer Entstehung. Während Mersenne die Unzulänglichkeit des menschlichen Verstandes betonte, angesichts der schon bei einer geringen Anzahl von Tönen ausufernden Kombinationsmöglichkeiten etwa in der Melodiebildung kaum zu einer rationalen Auswahl der Tonverbindungen gelangen zu können – nur die göttliche Vernunft gebiete über solche ins Unendliche weisenden Verhältnisse –, heben die Romantiker den Zufall als poetisches, also schöpferisches Prinzip hervor. »Der Dichter betet den Zufall an«,[10] und die gesamte Poësie beruhe »auf thätiger Idéenassociation – auf selbstthätiger, absichtlicher, idealischer *Zufallproduktion* – (zufällige – freye *Catenation*.) (Casuïstik – Fatum. *Casuation*.) *(Spiel.)*«,[11] behauptet Novalis. Auch für Friedrich Schlegel liegt im kombinatorischen Anordnen des Stoffs die Idee des Kunstwerks begründet,[12] und wie Novalis gelangt er zu dem Ergebnis: »Jede K[unst] hat μουσ[ikalische] Princ[ipien] und wird vollendet selbst Musik«.[13] Anders als Dichtung oder Malerei betrachten die Romantiker Musik als von vornherein selbstreferentiell, die übrigen Künste müssten die Freiheit von äußeren Bezügen erst mühsam erringen. Diese Freiheit verbürgt das poietische Potential der Kunst, indem sie nicht genötigt ist, sich an den Verhältnissen und Beschaffenheiten der Außenwelt, an dem, was außerhalb ihrer als sinnvoll gilt, orientieren zu müssen; eine Auffassung, die freilich Probleme aufwirft, zumal, wenn nicht allein die eigene Innenwelt in die Außenwelt projiziert, sondern ebenso die Welt des ästhetischen Scheins – der Simulation – als Wirklichkeit erachtet wird.

Hoffmann formuliert keine komplexen Theoriemodelle wie die Frühromantiker, nutzt jedoch seine Musik und seine erzählerisch-essayistischen Texte als Experimentalanordnungen, um die romantischen Vorstellungen und Haltungen zu erproben und zu untersuchen. Musikalisch mag sein *Grand Trio* in E-Dur (1809) als Beispiel dafür dienen, in welcher Weise Hoffmann kombinatorische Verfahren zur Herstellung eines quasi in sich geschlossenen musikalischen Ganzen nutzt. Keines der von ihm verwendeten thematischen und motivischen Elemente besitzt ein gestalthaftes Eigengewicht, auch mangelt es ihnen an semantischer Bestimmtheit. Ihre Funktionalität erhalten sie ausschließlich durch die Konfigurationen, in die sie gebracht werden. Infolge dieser Beschaffenheit eignen sich die Themen und Motive

10 Novalis: *Das Allgemeine Brouillon*, in: ders.: *Werke, Tagebücher und Briefe*, Bd. 2 (wie Anm. 4), S. 691 (Nr. 940). **11** Ebd., S. 692 (Nr. 953). **12** Friedrich Schlegel: *Kritische Friedrich-Schlegel-Ausgabe*, Bd. 18: *Philosophische Lehrjahre 1796–1806 nebst philosophischen Manuskripten aus den Jahren 1796–1828 I*, hrsg. v. Ernst Behler, Paderborn u. a. 1963, S. 341 (V, Nr. 224). **13** Ders.: *Kritische Friedrich-Schlegel-Ausgabe*, Bd. 16: *Fragmente zur Poesie und Literatur I*, hrsg. v. Hans Eichner, Paderborn u. a. 1981, S. 213 (VII, Nr. 120).

zu Veränderungen und immer neuen Kombinationen auf engem Raum, und zwar sowohl in ihrer horizontalen Abfolge wie zu unterschiedlichsten polyphonen Kombinationen in der Vertikale. Beide Möglichkeiten nutzt Hoffmann in großem Umfang. Die Zusammensetzung der Themen aus denkbar einfachen Motiven zieht eine weitgehende Aufhebung des Unterschieds von thematisch Relevantem und Begleitfiguren nach sich. Was in einem Takt der Begleitung dient, mag im nächsten schon thematisch werden.[14] Nahezu jedes Motiv, jede Floskel rückt im Verlauf des Satzes an irgendeiner Stelle mit thematischem Gewicht in den Vordergrund, wobei sich an diesen Konfigurationsprozessen die Instrumente gleichberechtigt beteiligen. Auf diese Weise variiert Hoffmann in seinem *Grand Trio* beständig Konfigurationen und klangliche Eigenschaften der verschiedenen thematischen Elemente und erzielt zugleich eine hohe formale und strukturelle Dichte der Komposition.

Auf ähnliche Weise strukturiert Hoffmann seine Erzählungen. Schon auf den ersten Blick kombinatorisch verfährt er in den *Lebensansichten des Katers Murr*. In dem Roman geraten scheinbar zwei Bücher durcheinander; genauer, eines wird immer wieder durch Seiten eines zweiten Buchs unterbrochen, dessen Blätter beim versehentlichen Einfügen zudem in der Reihenfolge vertauscht sind.[15] Hoffmann inszeniert Zufallskombinationen. In anderer Weise folgt die Erzählung *Der Sandmann* kombinatorischen Verfahren, indem ihr Aufbau an der Sonatenform orientiert ist.[16] Anstelle von Hauptsatz, Seitensatz und Überleitung treten in dem Analogon zur Exposition drei Briefe, denen vergleichbare Funktionen zukommen wie jenen in der Musik. Die Inhalte der Briefe werden durch Vertauschung, umgekehrte Sichtweisen und Zuordnungen von Eigenschaften in ganz unterschiedliche Relationen zueinander gebracht, sprich durch Kombinatorik.

Kombinatorische Verfahren gehören zu den grundlegenden Merkmalen auch des automatischen Komponierens, beginnend mit Kirchers Komponierkästchen bis hin zu ›komponierenden‹ KI-Systemen. Grob gesagt benötigt eine Komponiermaschine, ganz gleich wie sie aufgebaut ist, Elemente, mit denen sie komponiert und, soll das Komponieren nicht erratisch verfahren, Regeln, nach denen die Elemente miteinander kombiniert werden können. Nun ergibt das Einhalten von Regeln beim Kombinieren und Re-Kombinieren musikalischer Elemente noch keine gelungene Musik. Ein Musikstück kann sämtlichen der einmal festgelegten Regeln gerecht werden und gleichwohl eine fade Angelegenheit sein. Um auf maschinellem Wege der bloßen Regelerfüllung zu entkommen und abweichende, inter-

14 Vgl. Anja Pohsner: *»Wenn ich von mir selbst abhinge, würd' ich Componist …«. Die Umwege des Musikers E. T. A. Hoffmann. Wechselwirkungen innerhalb seines musikalischen und literarischen Werkes*, Dissertation Ruprecht-Karls-Universität Heidelberg 1999, S. 202f., http://archiv.ub.uni-heidelberg.de/volltextserver/2218/1/Dissertation_Anja_Pohsner.pdf [30.9.2025]. **15** E. T. A. Hoffmann: *Lebens-Ansichten des Katers Murr*, in: ders.: *Sämtliche Werke*, Bd. 5: *Lebens-Ansichten des Katers Murr. Werke 1820–1821*, hrsg. v. Hartmut Steinecke unter Mitarbeit von Gerhard Allroggen, Frankfurt a. M. 1992, S. 9–458, hier S. 12. **16** Dies deutet an, ohne es auszuführen, Friedhelm Auhuber: »Der Sandmann. Aspekte zu E. T. A. Hoffmanns Erzählkunst«, in: *E. T. A. Hoffmann: Der Sandmann. Faksimile-Ausgabe als Gesamtkunstwerk*, hrsg. v. Johannes Häfner, Berlin 1998, S. 93–109, hier S. 104.

essante kompositorische Ergebnisse zu erreichen, bietet sich der Zufall an. Wie die Romantiker annehmen, können mittels des Zufalls ungewöhnliche Kombinationen entstehen, genauso allerdings allzu Übliches. Vor allem wird selbst den Romantikern nicht jedes ungewöhnliche kompositorische Zufallsergebnis als ästhetisch und inhaltlich angemessen vorgekommen sein. In ihren Werken verhält es sich eher so, dass tatsächlich arrangierte Strukturen als zufällig entstanden erscheinen sollen. Der Grundgedanke, den Zufall als quasi schöpferisches Moment ins Verfahren zu integrieren, findet sich bereits in den ersten Entwürfen dessen wieder, was heute als Künstliche Intelligenz bezeichnet wird. Vermutlich sei es klug, in eine lernende Maschine ein Zufallselement zu integrieren, postuliert Alan Turing in seinem grundlegenden Essay *Computing Machinery and Intelligence* (1950).[17] Ein reines Zufallselement genüge jedoch nicht, um Kreativität hervorzubringen, mutmaßen die Autoren des *Proposal for the Darthmouth Summer Research Project on Artificial Intelligence* (1955), vielmehr erfordere ihre technische Realisierung zugleich dessen Kontrolle in einem geordneten gedanklichen Rahmen.[18] In dem Antrag findet der Begriff der künstlichen Intelligenz überhaupt erstmals Verwendung.

Zwar gibt es seit den 1920er Jahren ästhetische Verfahren, die den Zufall als ästhetisch mitentscheidendes Moment integrieren, im Surrealismus oder später in der Aleatorik. Sie zeigen, dass der Zufall in ganz unterschiedlicher Weise in Komposition oder Aufführung integriert werden kann; das Bestehen eines Zufallsaspekts sagt für sich genommen also noch nicht viel darüber aus, welche Funktion ihm zukommt. Entscheidender ist die Frage, ob der Zufall – abgemildert, Kontingenz – notwendigerweise innerhalb des Kompositionsverfahrens verortet werden muss und damit eine gewissermaßen syntaktische Funktion erfüllen würde, oder ob der Zufall ungeplant und von außen zum Komponieren hinzukommen und ästhetisch bedeutsame Folgen haben kann. Letzteres ist selbstverständlich zu bejahen. Das vom Komponisten nicht durchweg planbare Umfeld seines Komponierens, von kosmologischen Ereignissen über das politische Weltgeschehen bis zum Umkippen der Teetasse beim Schreiben, all dies kann Eingang finden in seine Komposition. Fehlt ein solcher Bezug zur Außenwelt, zu der vom Musikwerk aus betrachtet ebenfalls der Komponist gehört, wäre das Kunstschaffen eine sterile Angelegenheit ohne verstehbaren Sinn und jede Bedeutung. Hoffmann inszeniert diesen Welt- oder Lebensweltbezug der Kunst in den *Lebensansichten des Katers Murr* geradezu. Eine Unachtsamkeit des Katers sorgt für das Durcheinander der zwei Texte und so letztlich für die Gestalt des Buchs; nicht tatsächlich versteht sich, die Vermengung der Texte ist von Hoffmann sorgfältig arrangiert. Als Problem gibt Hoff-

17 Alan M. Turing: *Computing Machinery and Intelligence. Können Maschinen denken?* (engl./dt.), übers. und hrsg. v. Achim Stephan / Sven Walter, Ditzingen 2021, S. 24/25 u. 100/101. **18** John McCarthy / Marvin L. Minsky / Nathaniel Rochester / Claude E. Shannon: »A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence«, in: *AI Magazine 27*, 2006, No. 4, S. 12–14, hier S. 14, https://doi.org/10.1609/aimag.v27i4.1904.

mann seinem Lesepublikum mehr zum Denken auf als das Verstehen lediglich eines abstrakten Kompositionsverfahrens. Wer letzteres begreift, begreift nicht auch schon den Roman.

Kein Schriftsteller schreibt, um bestimmte Regeln für den Aufbau von Erzählungen zu erfüllen, sondern benutzt sie idealiter, um die Dinge, von denen er erzählen möchte, mit der literarischen Form zu verbinden, zu der er sie gestaltet. Kein Komponist komponiert zur Erfüllung von Satzregeln und dergleichen, sondern um musikalisch etwas mit ihnen zum Ausdruck zu bringen. Über solche Zwecke kann man nicht durch das Beherrschen von Kompositionsverfahren verfügen, und sie erschließen sich umgekehrt auch nicht allein aus der jeweiligen Anwendung der Verfahren und dem daraus entstandenen Ergebnis. Letzteres ist ein heikler Punkt, da Leser und Hörer nicht in die Köpfe der Künstler schauen können, um etwas über die Zwecke ihrer musikalischen, literarischen, bildnerischen Kompositionen zu erfahren. Verbunden sind sie – Menschen, Werke, Aufführungen – durch eine Welt, in der sie erfahren (werden), in der sie hervorbringen und hervorgebracht werden, in der sie aufführen und aufgeführt werden. Ohne diese Verbundenheit bliebe Kunst leeres Spiel. Dies lässt sich als Frage auf die sogenannte KI wenden, auf Komposition mittels automatischer Verfahren schlechthin: Wie stellt sich in ihnen ein solcher Außenweltbezug her? Mögen die generierten Muster noch so komplex sein, welchen Stellenwert hätten Kunst- und Musikverstehen beim Entfallen der Außenweltbezüge, der sie vom Kaffeesatzlesen unterschiede?

Das Subjekt als Maschine, die Maschine als Subjekt in Hoffmanns Erzählungen *Die Automate* und *Der Sandmann*

Die maschinenförmigen Verfahren ästhetischer Hervorbringung bilden den einen zentralen Aspekt in der romantischen Auseinandersetzung mit dem Verhältnis von Kunst und Maschine, der subjektive Bezug auf Maschinen und Automaten den anderen. Die wesentlichen Einsichten Hoffmanns zu der komplexen Übereinanderlagerung der naturwissenschaftlich-technischen Seite von Musik und ihrer das Gemüt bewegenden, verzaubernden Seite finden sich in den Erzählungen *Der Sandmann* und *Die Automate*. In letzterer besuchen die beiden Protagonisten, der Musiker Ludwig und der Dichter Ferdinand, eine Ausstellung, in der ein als Türke verkleideter Automat Antworten auf Fragen erteilt, die ihm gestellt werden. Ferdinand gibt sich vollkommen der Faszination an dem Automaten hin, während Ludwig ruhig abwägt und sich bisweilen spöttisch über das Spektakel äußert. Er weist Ferdinand, der auf seine Fragen von dem Automaten äußerst beunruhigende Antworten zu erhalten glaubt, darauf hin, dass er Inneres und Äußeres verwechsle, indem er selbst erst den Sinn in die Antworten des Automaten hineinlege, jedoch dem Automaten das Verfügen über Kenntnisse unterstelle, die ihm gestattet hätten, sinn-

volle Antworten auf die Fragen zu geben.[19] Mit dem sprechenden Türken, den die beiden Freunde letztlich nicht für einen tatsächlich rein mechanischen Automaten halten, überbietet Hoffmann erzählerisch Wolfgang von Kempelens schachspielenden Türken, einen Pseudoautomaten, der gegen menschliche Schachspieler antrat – und gewann. Dass ein Trick an der Sache war, hatte Kempelen nicht verheimlicht, freilich jedoch, worin dieser bestand; in einer Spiegelvorrichtung nämlich, mittels welcher ein im inneren des Automaten verborgener Gehilfe Kempelens das Schachbrett überblicken und entsprechend über eine Mechanik gesteuert die Türkenfigur die Figuren entsprechend setzen lassen konnte. Trotz aller Kritik wegen Scharlatanerie, trotz einiger gelungener Entlarvungen des Bluffs – der Automat blieb eine Sensation; das Publikum wollte an die technische Illusion glauben. Scheinbar hatte die Mechanik in Gestalt eines schachspielenden Automaten auf die menschliche Verstandestätigkeit übergegriffen.[20] Ein Automat, der sprachlich und inhaltlich sinnvoller Erwiderungen fähig ist, wie Hoffmann ihn in der Fiktion vorführt, steigert das Ausgreifen der Mechanik auf genuin menschliches Terrain noch einmal, daraus resultiert in der Erzählung die Unheimlichkeit des sprechenden Türken.

Im Fortgang der Erzählung besuchen Ludwig und Ferdinand den undurchsichtigen, über ein großes Automatenorchester verfügenden Professor X. Wiederum verwechselt Ferdinand die Ebene seines Inneren mit den mechanisch von den Automaten erzeugten Klängen und schreibt der mechanischen Musik des Ensembles Eigenschaften zu, die erst in seinem Inneren entstehen und nirgends sonst zu finden sind. Wiederum bemüht sich Ludwig, seinem Freund die komplexen Wechselbeziehungen zwischen innerem Erleben und äußerem Geschehen auseinanderzusetzen. Dabei kommt er auf Grundsätzliches. Für eine wichtige Aufgabe der Mechanik hält er es, die Musikinstrumente weiter zu vervollkommnen, nicht aber die von ihm missbilligte technische Nachahmung des Musizierens, die zu der von Ferdinand vollzogenen Ebenenverwechslung einlädt. Scheinbar lebendig und beseelt legt die äußere Gestalt, die in den musikspielenden Automaten tote Mechanik annimmt, die von Ludwig kritisierte Ebenenverwechslung manipulativ nahe. Hingegen lehnt er Maschinen, deren Aufgabenbereich sich nutzenorientiert auf die Bearbeitung der äußeren Natur beschränkt wie »eine gute Strumpfmaschine«, keineswegs ab.[21] Neben der Vorspiegelung von Subjektivität in Gestalt etwa automatisch musizierender Puppen – von ihm als Scharlatanerie betrachtet – und dem nicht verwerflichen alltagspraktischen Einsatz von Maschinen sucht Ludwig nach einer weiteren »höheren Mechanik«, die die Natur nicht in zweiter Ordnung nachahmt oder simuliert, sondern sie selbst zum Sprechen oder Klingen anregt. Die höhere Me-

19 E. T. A. Hoffmann: *Die Automate*, in: ders.: *Sämtliche Werke in sechs Bänden*, Bd. 4: *Die Serapions-Brüder*, hrsg. v. Wulf Segebrecht unter Mitarbeit von Ursula Segebrecht, Frankfurt a. M. 2001, S. 396–429, hier S. 403. **20** Vgl. Ernst Strouhal: »Eins sein mit allem, was tickt. Bewegungskontrolle und Zeitdisziplin am Beispiel des Schachautomaten von Wolfgang von Kempelen«, in: *Zyklus und Prozess. Joseph Haydn und die Zeit*, hrsg. v. Marie-Agnes Dittrich / Martin Eybl / Reinhard Kapp, Wien 2012 (Wiener Veröffentlichungen zur Musikgeschichte 10), S. 275–291. **21** Hoffmann: *Die Automate* (wie Anm. 19), S. 418f.

chanik solle die in den verschiedensten Materialien verborgenen Klangmöglichkeiten entdecken, sie systematisieren und technisch beherrschen, um sie musikalisch nutzen zu können; der für die musikalische Umsetzung notwendigen Naturbeherrschung entgeht auch die höhere Mechanik nicht. Hoffmann führt in der Erzählung zwar das im romantischen Denken häufig als unerreichbares Ideal herangezogene Bild vom Zustand einer uranfänglichen Menschheit an, deren Geist noch von der Natur ergriffen war, statt selbst die Dinge in der Natur als handhabbare Objekte zu ergreifen. In diesem Idealzustand habe das menschliche Innere noch mit den Klängen der Natur korrespondiert; ein Nachhall davon sei die Sage von der Sphärenharmonie. In einen solchen utopischen Zustand führt die höhere Mechanik freilich nicht zurück, ihrer bedürfte es sonst gar nicht. Ihre Grundlage bleibt die Naturbeherrschung, das rationale Ergreifen der Natur für die eigenen Zwecke. Die Unmittelbarkeit, die Ludwig mittels höherer Mechanik erreichen möchte, ist keine. Mit dem Ideal der höheren Mechanik beruhigt die Erzählung ihre Leser entsprechend nicht. Professor X verfüge zwar über die nötigen Kenntnisse, sich in klingende Korrespondenz mit der Natur zu setzen, lässt die Erzählung wissen, seiner Zwielichtigkeit wird er jedoch nicht völlig ledig. Von seinen technischen und musikalischen Vorführungen jedenfalls hat sich Ferdinand ergreifen lassen, die mit ihnen verbundenen Erlebnisse und deren Deutung okkupieren sein Inneres vollkommen, und er gerät anscheinend in einen Zustand der Zerrüttung. Selbst Ludwig bleibt sich seiner zunächst nüchternen Sicht zeitweise nicht sicher.

Am konsequentesten führt Hoffmann das Thema der Verwechslung von Innerem und Äußerem im *Sandmann* aus. Ihre Drastik erhält die Erzählung nicht zuletzt durch die von Hoffmann genutzten kombinatorischen Verfahren, durch welche sich Regelhaftigkeit und scheinbare Zufälligkeit miteinander verbinden. Regelhaftigkeit erhalten die Kombinationen in der Erzählung durch die Sonatenform. Selbstverständlich lässt sich der Sonatensatz nicht vollständig in Sprache übertragen, die strukturellen Analogien reichen in der Erzählung jedoch sehr weit. Als kontrastierende Themen fungieren einerseits Nathanaels von dunklen Ahnungen der Einbildung beherrschte Perspektive und andererseits Claras Blickwinkel – wie es ihr in den Namen gelegt ist – aus der Tagwelt und Vernunft. In Briefform werden sie expositionsartig gegenübergestellt. Schon hier mischt sich jedoch der scheinbare Zufall ein, Nathanael adressiert den Brief, mit dem die Erzählung beginnt, nicht an seinen Freund Lothar, an den die Mitteilung gerichtet ist, sondern an seine Verlobte Clara. Ein dritter Brief vermittelt beide Perspektiven, bildet gleichsam die Schlussgruppe der Exposition.[22] Ein Durchführungsabschnitt hebt die Themen auf eine höhere Stufe, indem die Erzählung auf sich selbst reflektiert. Beide Themen werden wieder aufgegriffen, erneut gegenübergestellt und weiterentwickelt. Nathanaels Phantasiewelt konfligiert unmittelbar mit Claras Ratio.[23] Die knappe Reprise beginnt mit Nathanaels vom Wahn genesenem Blick auf Claras Tagwelt. Allerdings

22 E.T.A. Hoffmann: *Der Sandmann*, in: ders.: *Sämtliche Werke in sechs Bänden*, Bd. 3: *Nachtstücke. Klein Zaches. Prinzessin Brambilla. Werke 1816–1820*, hrsg. v. Hartmut Steinecke unter Mitarbeit von Gerhard Allroggen, Frankfurt a. M. 1985, S. 11–49, hier S. 11–25. **23** Ebd., S. 25–47.

bricht, indem er seine Verlobte durch ein Perspektiv – das in der Erzählung dem Blick nicht nur Entferntes als nah zeigt, sondern ihm zugleich das Nahe entrückt – betrachtet, sein Wahn wieder aus.[24] Diesmal lässt sich der Konflikt von Innen- und Außenwelt, den Nathanael austrägt, nicht mehr bewältigen. Nachdem er die Zerstörung der Automatenpuppe Olimpia beobachtet hat, stürzt er in den Tod. Woran er scheitert, ist die Verwechslung seiner subjektiven Innenwelt, die er schließlich vollkommen in den Olimpia-Automaten hineinprojiziert, mit der Außenwelt, für die er blind geworden ist.

Die strukturelle Analogie zur Sonatenform setzt sich fort in der den beiden kontrastierenden Themen zugrundeliegenden Einheit. Die Themenverwandtschaft vollzieht sich über die Verwechslungen von Einbildungen mit der Wirklichkeit, die Nathanael in seiner Phantasie vornimmt, so dass sich in den verschiedenen Erzählsträngen und Abschnitten dieselben Motive wiederfinden, die Ebenen der Einbildung und der Wirklichkeit jedoch wechseln. Was in der wirklichen, hellen Tagwelt Claras beseelt erscheint, wandelt sich aus Perspektive der inneren Welt Nathanaels ins Automatenhafte, und was in der äußeren Welt ein Automat ist, verwandelt sich ihm innerlich in Beseeltes. Die Verwechslungen vollziehen sich zwar insgesamt über Nathanaels Augen, denen der Blick geraubt wird – als Kind vom Sandmann, später vom Perspektiv. Seine innere Hingezogenheit zu Olimpia entwickelt Nathanael jedoch vor allem ausgehend von einem Konzert, in dem der Automat Klavier spielt und eine Arie singt. In der Musik – verstanden als Kunst der Innerlichkeit – äußert sich ein Subjekt. Gerade die Auffassung, Musik sei Ausdruck subjektiver Innerlichkeit, verwandelt sie in ein Medium, das Nathanael dazu verführt, Olimpia seine eigene Subjektivität als die ihre zu unterschieben. Wie der sprechende Türke in der Erzählung *Die Automate* auf Kempelens schachspielenden Türken anspielt, so dürfte Olimpia ihr Vorbild in der 1774 von Pierre und Henri-Louis Jaquet-Droz konstruierten Orgelspielerin haben, die ein funktionierender Musikerautomat ist, der fünf Stücke auf einer Orgel spielen kann und dazu ähnlich wie ein menschlicher Spieler ein allerdings nur zweioktaviges Manual bedient. Nicht allein bewegt sie zum Niederdrücken der Tasten ihre Finger einzeln, dreht ihre Arme seitlich, um das Manual beim Spielen ausschöpfen zu können. Sie wendet leicht den Kopf, nickt dem Publikum zu. Mit ihren beweglichen Augen scheint sie überdies ihrem Spiel zu folgen und wie beim Atmen den Brustkorb zu heben und zu senken. Nach beendetem Spiel führt sie eine leichte Verbeugung aus.[25] Sie verfügt also über eine Mechanik, die so etwas wie Ausdrucksgestik andeuten kann. Dieser ›Lebensmechanismus‹ weist eine Funktionsdauer von etwa anderthalb Stunden auf und gehört zu den kompliziertesten von Jaquet-Droz gebauten Mechaniken. Die langsamen, nacheinander einsetzenden Körperbewegungen der Orgelspielerin sollen den Eindruck erwecken, sie erwache vor dem Publikum zum Leben.[26] Anders als Hoffmanns Olimpia erzeugt die Orgelspielerin zwar keine stimmlichen Laute; Automaten, die eine menschliche Stimme imitierten, wurden im 18. Jahrhundert jedoch

24 Ebd., S. 47–49. **25** Roland Carrera / Dominique Loiseau / Olivier Roux: *Androiden. Die Automaten von Jaquet-Droz*, Lausanne 1979, S. 52–93. **26** Ebd., S. 76.

bereits gebaut. Bemühungen, mechanisch lebendige Körperlichkeit und gestischen sowie stimmlichen Ausdruck nachzubilden sind keine Sache allein der erzählerischen Fiktion Hoffmanns. In seiner Erzählung greift er auf tatsächlich realisierte Konstruktionen als Vorbilder zurück, auch wenn sie in seinen Schilderungen modifiziert erscheinen. Der in der Erzählung von Nathanael begangenen Subreption kommt nicht lediglich das mechanisch erzeugte Verhalten des Automaten Olimpia entgegen, sondern die Wirklichkeit des avancierten Automatenbaus im Übergang vom 18. zum 19. Jahrhundert. Zumindest bei den Automaten von Jaquet-Droz geht die mechanische Technik in eine Illusion des Lebendigen über; einige Zeitgenossen wie etwa Félicité de Genlis in ihrer pädagogischen Abhandlung *Les veillées du château, ou Cours de morale à l'usage des enfants* (1784) nahmen dies durchaus wahr.[27]

Was Hoffmann im *Sandmann* ausführt, sind die möglichen negativen Auswirkungen des Romantisierens, des Identifizierens von etwas Alltäglichem, Gewöhnlichem mit aus der Phantasie stammendem Wunderbaren. Zwar durchschaut Nathanael seine Verwechslungen selbst als solche, gibt sie aber dennoch nicht auf. In der Durchführung trägt Nathanael seiner Verlobten eine düstere Erzählung über ihrer beider Beziehung zueinander vor, in der die Verwechslungen, die die Handlung des *Sandmanns* bestimmen, auf die Spitze getrieben werden. Er erzählt, dass der frühere Freund seines Vaters, Coppelius, den er in seiner Kindheit als Sandmann imaginierte, Clara die Augen raube und sie in Nathanaels Brust springen lasse. Clara weist in der von ihm verfassten Erzählung sogar darauf hin, »[k]annst du mich denn nicht erschauen? […] das waren ja nicht meine Augen, die so in deiner Brust brannten, das waren ja glühende Tropfen deines eignen Herzbluts – ich habe ja meine Augen, sieh mich doch nur an!«[28] Tatsächlich schaut er hin, sieht aber durch Claras Augen lediglich den Tod blicken. Nachdem Nathanael die Erzählung abgeschlossen hat, Clara seine düsteren Phantasien aber nüchtern abweist, reagiert er mit einem Paroxysmus und schimpft sie ein »lebloses, verdammtes Automat«.[29] Offenbar hat Clara an seiner poetischen inneren Welt keinen Anteil, für ihn rechnet sie daher allein der kalten, rational und technisch bestimmten Außenwelt zu. In Nathanaels Phantasie werden Clara die Augen geraubt, in der Realität wird er blind für seine Verlobte.

Komplementär zur Verwechslung Claras mit einem Automaten erblickt er in dem von dem zwielichtigen Spalanzani konstruierten Automaten Olimpia eine poetisch veranlagte Frau. Eine Projektion, die eintritt, nachdem er Olimpia von seinem Zimmer aus durch sein Perspektiv betrachtet hat. Von ihr behauptet er dann, »[w]ohl mag euch, ihr kalten prosaischen Menschen, Olimpia unheimlich sein. Nur dem poetischen Gemüt entfaltet sich das gleich organisierte! – Nur *mir* ging ihr Liebesblick auf und durchstrahlte Sinn und Gedanken, nur in Olimpia's Liebe finde ich

27 Vgl. Sandrine Giradier: *L'entreprise Jaquet-Droz. Entre merveilles de spectacle, mécaniques luxueuses et machines utiles (1758–1811)*, Neuchâtel 2020, S. 117, 132, 142f. u. a. **28** Hoffmann: *Der Sandmann* (wie Anm. 22), S. 31. **29** Ebd., S. 32.

mein Selbst wieder«.[30] Die lebendige Clara stellt aus Nathanaels Innenweltperspektive einen Automaten dar, der Automat Olimpia eine lebendige Frau. Zum einen entlarvt Nathanael sein poetisches Gemüt als selbst automatenhaft – die denkbar stärksten Gefühlskontraste in der Erzählung beruhen auf einer mittels rationaler Verfahren hergestellten Einheit –, zum anderen begegnet er in der mechanischen Puppe ausschließlich seinem Selbst wieder, denn ihre Liebe zu ihm ist nichts anderes als seine eigene Projektion auf sie. »Ach, ach« sind die einzigen Äußerungen, die Olimpia zu Unterhaltungen mit Nathanael beizutragen in der Lage ist, und zum Abschied bringt sie ein »Gute Nacht, mein Lieber!« hervor.[31] Nathanael legt es ihr als Tiefe des Gemüts aus und fühlt sich von ihr vollkommen verstanden.

In der Reprise der Erzählung wiederholt sich noch einmal der Augenraub. Spalanzani und der Wetterglashändler Coppola streiten sich um die von ihnen gemeinsam gebaute Olimpia und ramponieren sie im Laufe der Auseinandersetzung. Spalanzani schleudert dem hinzukommenden Nathanael die Glasaugen gegen die Brust. Was sich innerhalb der Durchführung als ein Ereignis in der von ihm verfassten Dichtung mit den Augen Claras vollzog, wird hier mit den Augen Olimpias zu einem Ereignis in der realen Außenwelt. In seiner Geschichte verliert Nathanael den Blick für die äußere Welt. Infolge des späteren äußeren Ereignisses wird er mit Olimpia seines Mediums zur Entfaltung und zum Erleben seiner Innenwelt beraubt. Die Verhältnisse werden also gespiegelt. Mit der Zerstörung Olimpias kollabiert zugleich Nathanaels Innenwelt, hatte er diese doch völlig mit dem äußeren Medium der Puppe identifiziert. Infolge der von ihm vorgenommenen Totalisierung der Ebenenverwechslungen mündet die Erzählung beinahe zwangsläufig in den Selbstverlust Nathanaels. Der Sturz in den Tod bleibt der konsequente Ausweg.

Hoffmann untersucht in den Erzählungen *Die Automate* und *Der Sandmann* mehrere für die Auseinandersetzung mit KI und von KI generierter Musik relevante Probleme. Ausgiebig untersucht er die Rolle von Subjekt und Subjektivität im Umgang mit mechanischen Automaten. Früh widmet er sich dem manipulativen Aspekt von Technik, welche durch Form und Betriebsverhalten Konstruktionsmerkmale aufweist, die eine Subjekthaftigkeit oder zumindest Subjektähnlichkeit nahelegen. In *Die Automate* kritisiert der Musiker Ludwig dies eindringlich.

Die smarte Welt der KI

Eine aus dem *Sandmann* zu gewinnende Annahme könnte sein, dass Automaten sich der menschlichen Einbildungskraft und ihren Projektionen leichter fügen als ein menschliches oder tierisches Gegenüber; sie setzen anders als diese dem Subjekt keinen »eigenwilligen« Widerstand entgegen. Eine solche widerstandslose Einpassung scheint die Leitvorstellung für smarte Technologien zu sein. Als etwas, das widerständig sein könnte, tritt Wirklichkeit idealiter nicht in Erscheinung, die

30 Ebd., S. 42. **31** Ebd., S. 43.

Geräte organisieren unvermerkt und reibungslos etwa den Haushalt um das Leben und die Bedürfnisse der Menschen, bestellen die zur Neige gehende Milch nach, beobachten das Schlafverhalten, haben den Kaffee bereits pünktlich zum Aufstehen gekocht, erinnern an Termine und kennen die Wünsche ihrer Besitzer einschließlich ihres Musikgeschmacks besser als diese selbst. Das Ideal von einer Technik, die denen, die mit ihr umgehen, die Mühe erspart, sich mit der widerständigen Realität auseinandersetzen zu müssen, analysiert Matthew B. Crawford an dem Wandel der Mickey-Mouse-Geschichten seit den späten 1920er Jahren, als die Figur erfunden wurde, bis heute. Das Bild von Technologie, das in den Filmen anfänglich gezeichnet wurde, nämlich dass sie ihre gefährlichen Tücken hat, eine zu berücksichtigende Eigengesetzlichkeit, und sich durchaus nicht umstandslos den Wünschen derer fügt, die mit ihr hantieren, sei in der heutigen *Mickey Mouse Clubhouse*-Serie einem Bild gewichen, nach welchem, sobald eine Schwierigkeit entsteht, immer schon eine mittels eines »Toodles« genannten Device abrufbare, reibungslos funktionierende technologische Lösung bereitsteht. Das Clubhouse selbst ist eine technologisch autonome Einrichtung, die Überbietung jeder Vorstellung von Smarthome. Eine merkliche Eigengesetzlichkeit der physischen oder auch nur der technischen Welt kommt nicht zum Vorschein.[32] Im Grunde verallgemeinert die *Mickey Mouse Clubhouse*-Serie eine Situation, in der sich im *Sandmann* Nathanael angesichts seiner Selbstspiegelung in Olimpia befindet; eine Situation, in welche die Realität nicht einbrechen darf.

Mit dem Verlust der Widerständigkeit der physischen Realität und der Widerständigkeit der anderen Subjekte, die nicht in der eigenen Subjektivität aufgehen, löst sich letztlich das Subjekt selbst als eigenständige, ihrer selbst bewusste Instanz auf. Ohne den Widerstand der Welt und ohne die Widerständigkeit anderer Subjekte gerät sich das Subjekt kaum als Subjekt und erst recht nicht als ein Subjekt unter anderen Subjekten in den Blick.[33] In seiner Subjektivität kann ein Subjekt allein durch andere Subjekte wahrgenommen und als ein solches anerkannt werden. Sämtliche dieser für die Eigenständigkeit des Subjekts notwendigen Bedingungen werden untergraben durch die Projektion von Subjektivität in die Apparatur, die nur erfolgen kann, indem der Apparatur die selbst hervorgebrachten Gedanken, Interpretationen, Haltungen, Stimmungen als Äußerung und Haltung untergeschoben werden. Der Technoanimismus, der – herausgetreten aus den literarischen Erzählungen der Romantiker und den Mickey-Mouse-Filmen – vermittels

32 Matthew B. Crawford: *Die Wiedergewinnung des Wirklichen. Eine Philosophie des Ichs im Zeitalter der Zerstreuung*, übers. v. Stephan Gebauer, Berlin 2016, S. 111–116. **33** Edmund Husserl: *Husserliana. Gesammelte Werke*, Bd. 13: *Zur Phänomenologie der Intersubjektivität. Texte aus dem Nachlass*, Teil 1: *1905–1920*, hrsg. v. Iso Kern, Den Haag 1973, S. 244; ders.: *Husserliana. Gesammelte Werke*, Bd. 14: *Zur Phänomenologie der Intersubjektivität. Texte aus dem* Nachlass, Teil 2: *1921–1928*, hrsg. v. Iso Kern, Den Haag 1973, S. 170ff.; ders.: *Husserliana. Gesammelte Werke*, Bd. 8: *Erste Philosophie (1923/24)*, Teil 2: *Theorie der phänomenologischen Reduktion*, hrsg. v. Rudolf Boehm, Den Haag 1959, S. 61ff.

digitaler Apparaturen inzwischen Gesellschaft und Alltag prägt, beruht wesentlich – ebenfalls ein romantischer Topos – auf Doppelgängertum.[34] Die in der Online-Nutzung erzeugten und von KI-Systemen analysierten Daten generieren Nutzerprofile als unsere Datendoppelgänger, die uns verraten, was wir uns von den unüberschaubaren Angeboten der digitalen Welt wünschen und wofür wir uns interessieren sollten. Die digitale Welt scheint sich immer schon unseren Wünschen zu fügen, wir scheinen uns mit ihr auf Du und Du zu befinden; tatsächlich jedoch fehlt die Perspektive zweiter Person in diesem Szenario. Vom archaischen Animismus, der Relationen von erster und zweiter Person zur Grundlage hatte, unterscheidet sich der »Technoanimismus«[35] durch ein verfehltes Kurzschließen von erster und dritter Person. Personalisierte Bezeichnungen und Namen für die Funktionen, anhand derer wir uns in den smarten Umgebungen orientieren, wie persönlicher Assistent, digital Companion, Alexa oder Siri, kaschieren dies lediglich. Der objektivierte Datendoppelgänger hält uns fest in dem, was über uns festgestellt wurde. Die Offenheit in der Begegnung mit einem anderen Subjekt, das den an es herangetragenen Erwartungen nicht entsprechen muss, das überraschen kann, dessen Antwort nicht determiniert ist und sogar ausbleiben kann, hat in der Begegnung mit unserem Datendoppelgänger keinen Raum. Indem wir ihm folgen, bleibt er nicht einfach die Objektivierung unserer bisherigen Wünsche und unseres vergangenen Verhaltens, er objektiviert mittels vergangener Wünsche und vergangenen Verhaltens, was wir wünschen werden und wie wir uns verhalten werden, lange bevor wir selbst etwas wünschen. Die Rahmenbedingungen, unter denen die Wünsche geäußert werden und unter denen das Verhalten stattfindet, sind so konstruiert, dass die für die Eigentümer dieser Rahmen interessanten Verhaltensweisen und Wünsche optimal erfasst werden – etwa sofern sie auf Konsum gerichtet sind, Leistung anzeigen oder politische Einstellungen verraten. Die Datendoppelgänger entstehen nicht nach unserem Ebenbild, sondern nach den Rastern der Konzerne oder Staaten, die über die entsprechenden Technologien verfügen. Unsere Datendoppelgänger spiegeln nicht uns, sie spiegeln uns nach den Maßstäben und in den Interessen anderer. Selbst unsere potentiellen Freunde und Partner ermittelt ein Datendoppelgänger für uns; auch während wir selbst nicht sozial agieren, hält uns der Doppelgänger in der digitalen Kommunikation.[36] Gewissermaßen durch eine List, die uns Subjektivierung und Individualisierung vorspiegelt, unterliegen wir radikaler Verdinglichung in einem weit höheren Maße, als es sich Theodor W. Adorno und Max Horkheimer in ihrer *Dialektik der Aufklärung* (1947) vorstellen.[37]

Die Virulenz dieser Problematik hat, seit Hoffmann seine Erzählungen niederschrieb, erheblich zugenommen, gleichwohl wird die Rolle von Subjekt und Sub-

34 Vgl. Bruno Gransche: »Technogene Unheimlichkeit«, in: *Autonomie und Unheimlichkeit. Jahrbuch Technikphilosophie 6*, 2020, S. 33–51. **35** Ebd., S. 46f. **36** Vgl. Manfred Faßler: *Partizipation ohne Demokratie. Über die Folgen der Netz- und Geopolitik von Facebook, Google, Amazon & Co.*, Paderborn 2020, S. 225–246. **37** Theodor W. Adorno: *Gesammelte Schriften*, Bd. 3: *Dialektik der Aufklärung. Philosophische Fragmente* (zus. mit Max Horkheimer), hrsg. v. Rolf Tiedemann, Frankfurt a. M. 1981.

jektivität im Umgang mit KI-Systemen kaum angemessen diskutiert. Dies hat mehrere Gründe. Zunächst wird der menschliche Geist in Kognitions- und Neurowissenschaft selbst nach einem Computermodell betrachtet und umgekehrt basieren die Modellvorstellungen für KI auf dem Aufbau der Vernetzung von Nervenzellen. KI-Systeme sind entsprechend aufgebaut aus künstlichen neuronalen Netzen. Selbstverständlich bestehen erhebliche Unterschiede zwischen den Abläufen in Gehirnen von Lebewesen und den Prozessen in künstlichen neuronalen Netzwerken; das zu erkennen ist nicht die Schwierigkeit. Was sich infolge der schleifenartigen Verknüpfung der Auffassung vom menschlichen Geist als Maschine mit der Auffassung von KI als gehirnähnlich verliert, ist eine Grenzbestimmung zwischen beiden. Anders formuliert, die Differenz zwischen Denken und Fühlen einerseits und deren Simulation andererseits wird in solcher zirkulären begrifflichen und theoretischen Wechselbeziehung, die auch zwischen den jeweils zuständigen Disziplinen besteht, nicht sorgsam gezogen oder bewusst verwischt.[38]

Behavioristische und funktionalistische Auffassungen von Intelligenz in der KI-Forschung

Entscheidend für das Funktionieren der Äquivokation von menschlichem Geist und den Operationen von KI-Systemen ist zunächst eine grundsätzlich behavioristische Ausrichtung der KI-Forschung. Die Betrachtung des menschlichen – oder tierischen – Geistes erfolgt ausschließlich in Perspektive der dritten Person, also als Beobachtung von außen. Verstehen lassen sich solche Erklärungen freilich nur innerhalb des Geistes, der erklärt werden soll, und sie beruhen auf dem Geist, den sie erst erklären sollen. Überdies ist nicht einmal geklärt, wie das Denken von Gedanken, das Wahrnehmen von Wahrnehmungen und das Fühlen von Gefühlen aus den empirischen Beobachtungen herleitbar sein könnten. Das Problem, dass Denken nur durch Denken verstanden werden kann, findet – von Ausnahmen wie Thomas Nagel abgesehen – kaum Berücksichtigung, da die denkende, wahrnehmende und fühlende Instanz, das Subjekt, vorweg ausgeklammert wird. Übrig bleibt im Auffassen des Geistes nach dem Modell eines Computers allein das Beobachten des Verhaltens, um etwas über den Geist eines Gegenstandes zu erfahren. Ein aus dem Verhalten abgeleitetes Verständnis von Intelligenz, nach dem alles irrelevant ist, was sich nicht empirisch in Perspektive der dritten Person beobachten lässt, zieht gravierende Probleme nach sich. Bereits die Reduktion auf das Verhalten spricht gegen die Absicht, Geist, Denken, Wahrnehmen, Fühlen etc.

38 Vgl. Sarah Spiekermann: »Zum Unterschied zwischen künstlicher und menschlicher Intelligenz und den ethischen Implikationen der Verwechslung«, in: *Philosophisches Handbuch Künstliche Intelligenz*, hrsg. v. Klaus Mainzer, Wiesbaden 2024, S. 1–17, hier S. 15, https://doi.org/10.1007/978-3-658-23715-8_35-1 (online publiziert: 19.4.2022).

analysieren oder erklären zu wollen. Verhalten und Verhaltensdispositionen (die nur aus dem Verhalten erschlossen werden können) sind Termini auf einer rein äußeren Ebene, die frei ist von Semantik, von Zwecken, von Intentionen. Angemessen wäre das Beobachten von Handlungen, die freilich ohne ein Subjekt, das handelt, nicht verstehbar sind. Ohne Subjekt lassen sich Handlungen schon deshalb nicht in den Blick nehmen, da sie schon rein empirisch nicht mit beobachtbarem Verhalten in Deckungsgleichheit zu bringen sind; Handlung und Verhalten sind nicht dasselbe. Die Handlung kann sein das Grüßen, das sich ausdrücken kann durch Winken, leichtes Handheben, Kopfnicken. Das nach Maßstäben beobachtbaren Verhaltens selbe Winken kann heißen »Hier bin ich«, dasselbe Kopfnicken »Es geht los«, dasselbe Handheben »Stopp, warte!«; ein Umstand, an dem in der Soziologie der Behaviorismus scheiterte. Doch schon die Beschreibung von Verhalten ergibt sich nicht einfach aus empirischen Beobachtungen, denn am tatsächlichen Verhalten können unabschließbar viele Merkmale beobachtet werden. Erforderlich ist also ein Begriff des jeweiligen Verhaltens, um genau ›dieses‹ Verhalten beobachten zu können. Zugleich muss dieser Begriff festlegen, in welcher Relation die jeweils relevanten beobachteten Merkmale zu dem Verhalten stehen, das beobachtet werden soll.

Mit diesem Wissensmodell bleiben in der KI-Forschung das bewusste Vollziehen von Denkakten, das Innehaben von Gedanken und ebenfalls die innere Vertrautheit des Subjekts mit sich selbst als irrelevant für das Verständnis von Denken und Intelligenz ausgeklammert. Ergänzt wird die Beobachtung des äußeren Verhaltens durch eine Beobachtung der mentalen Zustände. Das Innere des beobachteten mentalen Systems erfüllt Funktionen, die darin bestehen, infolge eines Inputs von außen eine Bewertung seines jeweiligen inneren Systemzustands durchzuführen, diesen zu ändern und einen durch diese Vorgänge kausal hervorgerufenen Output zu erzeugen. Für die Funktionsweise des mentalen Systems gibt es im Wesentlichen zwei verschiedene Modelle: das Computermodell, nach welchem das Gehirn quasi Programme ausführt, und das konnektionistische Modell, das auf der Theorie neuronaler Netze beruht. Die konkrete Beschaffenheit eines Systems und seiner Zustände ist belanglos, da den Prinzipien des Funktionalismus zufolge die Funktionen unterschiedlich realisiert sein können: implementiert im Gehirn, in einer rein mechanisch konstruierten »Analytical Engine«, wie Babbage sie entwarf, oder auf einem modernen Computer. Eine Schwierigkeit des Funktionalismus – abgesehen davon, dass er die eben erläuterten Probleme des Behaviorismus nicht löst – besteht in dem Fehlen des Bewusstseins, des bewussten Erlebens, Denkens und Fühlens. Es ist nicht einsichtig, wie bewusstes Erleben in diesem Modell Raum erhalten könnte. Aus den Programmabläufen oder den Verknüpfungszuständen neuronaler Netze ist der Ebenenwechsel zum Bewusstsein nicht ableitbar, zumal etwa Ned Block die Annahme von der multiplen Realisierbarkeit der Funktionen nutzt, um plausibel zu zeigen, dass die Funktionen auch in Systemen realisierbar sind, in

denen es nicht zu Bewusstsein kommen kann.[39] Anzweifeln ließe sich auch die Annahme von der beliebigen Realisierbarkeit der Funktionen. So könnten zwei Systeme sich durchaus in demselben Programmzustand befinden oder dieselben internen Verknüpfungs- und Wahrscheinlichkeitsrelationen aufweisen und doch nicht denselben Zustand haben. Es könnte sich ein menschliches Gehirn in diesem Zustand befinden, dort signalisiert er Schmerz, und ebenso eine aufwendig konstruierte mechanische Babbage-Maschine. Nur, was hieße es im Falle der vollständig aus Metall konstruierten Babbage-Maschine, sie befände sich in einem Schmerzzustand?

Denken und Sprechen in der ersten Person und damit ebenso das Anreden in der zweiten Person und das Warten auf eine Antwort von einem Gegenüber entfallen. Mit der Festlegung auf Funktionalität und Verhalten verzichtet die KI-Forschung auf so ziemlich alles, was Denken, Wahrnehmen, Fühlen etc. gewöhnlich kennzeichnet. Auch klammert sie all das aus, was die wechselseitige Bezogenheit von Subjekten aufeinander charakterisiert. Eine wesentliche Eigenschaft dieses Wissensmodells besteht in einer Verbindung eines extremen Konstruktivismus mit einem extremen Naturalismus. Subjektivität, Selbstsein, Bewusstsein, die Erfahrung des eigenen Leibes und nicht weniger der Welt werden zu Simulationen eines Biocomputers, des Gehirns, erklärt. Die Reduktion der ersten Person auf etwas, das niemand ist, gibt den Blick frei auf das, was aus naturalistischer Perspektive an physikalisch beschreibbaren Tatsachen übrigbleibt.[40] Wir seien »mentale Selbstmodelle informationsverarbeitender Biosysteme«, behauptet Thomas Metzinger,

> »und zentrieren eine kritische Teilmenge ihrer internen Repräsentate der Welt. Werden wir nicht errechnet, so gibt es uns nicht. Als geistige Wesen sind wir somit Teil einer evolutionären Strategie: Gewisse Organismen können mit unserer Hilfe ihren internen Simulationsraum zentrieren, wesentlich differenzieren und so ein beträchtlich höheres Maß an Informationen verarbeiten. Dies stellte für sie einen Vorteil im Kampf ums Überleben dar.«[41]

Anders gesagt, wir sind lediglich der Nullpunkt eines vom Gehirn hergestellten Wirklichkeitsmodells, das die chaotische Realität der eigenen Datenströme und -verteilungen kohärent ordnet. Die Wirklichkeit, in der wir uns bewegen, besteht in einer Simulation, errechnet von einem datenverarbeitenden Biosystem. Was es

39 Ned Block: »Troubles with Functionalism«, in: *Perception and Cognition. Issues in the Foundations of Psychology*, hrsg. v. C. Wade Savage, Minneapolis 1978 (Minnesota Studies in the Philosophy of Science 9), S. 261–325. **40** Beispiele wären Thomas Metzinger: »Being No One – Eine sehr kurze deutsche Zusammenfassung«, in: *Grundkurs Philosophie des Geistes*, Bd. 1: *Phänomenales Bewusstsein*, hrsg. v. dems., Paderborn 2006, S. 424–475; ders.: *Der Ego-Tunnel. Eine neue Philosophie des Selbst: Von der Hirnforschung zur Bewusstseinsethik*, 8. Aufl. München 2023; Gerhard Roth: *Aus Sicht des Gehirns*, Frankfurt a. M. 2003 oder jüngst David J. Chalmers: *Realität+. Virtuelle Welten und Probleme der Philosophie*, übers. v. Björn Brodowski / Jan-Erik Strasser, Berlin 2023. **41** Thomas Metzinger: *Subjekt und Selbstmodell. Die Perspektivität phänomenalen Bewusstseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*, 2. Aufl. Paderborn 1999, S. 284f.

außerhalb dieser Datenströme gibt, scheint irrelevant. Nun bleibt selbstverständlich die Frage, wie das etwa 24 m² große Schreibzimmer, in dem ich an diesem Aufsatz arbeite, in meinen Kopf passen könnte oder wie ich es geschafft habe, mich auf einen Stuhl in meinem Kopf zu setzen, und wie der Aufsatz, den mein Selbstmodell in meinem Kopf tippt, je von irgendeinem anderen Selbstmodell, das sich in seiner eigenen Simulation aufhält, wird gelesen werden können. Nicht einmal nur vorgestellte Gegenstände befinden sich im Kopf, weder der Kentaur, den Edmund Husserl sich vorstellte, noch der knallrote hausgroße Drachen, mit seiner allversengenden Flamme, den ich mir vorstelle. Aber nicht allein dies, den evolutionären Vorteil, mit dem Metzinger das Vorhandensein der bewussten simplifizierten Wirklichkeitskonstruktionen im Kopf begründet, gäbe es nicht, sofern der Schein des Bewusstseins nicht auch auf das Gehirn einwirken könnte, in welchem keine Information ausgetauscht wird, sondern biochemische Vorgänge stattfinden. Die Instanz, die in unserem Kopf die Simulation benutzt und die der in Rede stehenden Konzeption nach nicht wir sind, muss diese Simulation ihrerseits phänomenal hören, schauen, fühlen, erfahren können. Zunächst müsste also das »Selbstmodell«, wenn es als repräsentationales Phänomen begriffen wird, von irgendetwas als Repräsentation aufgefasst werden können. Das setzte ein weiteres Selbst voraus, dem das Selbstmodell präsent wäre, und das, wäre es wieder ein repräsentationales Phänomen, eines weiteren Selbst bedürfte, dem es präsent wäre, und weiter ad infinitum. Wäre jedes Selbst ein repräsentationales Phänomen, gäbe es auf keiner Stufe dieses Prozesses so etwas wie Bewusstsein.[42] In keiner zweistelligen Relation, gleichgültig ob hierarchisiert in reflexive Bewusstseinsprozesse erster, zweiter, dritter usw. Ordnung oder temporalisiert zu rekursiven Schleifen, in denen das Bewusstsein auf frühere eigene Zustände zurückgreift: Bewusstsein entsteht auf diese Weise nirgends. Entweder mündet der Prozess in einem infiniten Regress und gelangt nie zu etwas Bewusstem oder er endet unbewusst, mögen in den Prozess noch so viele weitere rekursive Schleifen eingebunden sein.[43] Zudem bleibt die Frage, wer die Relation zwischen den verschiedenen Zuständen des Bewusstseins herstellt. Von selbst ergeben sich die Beziehungen nicht, es sei denn, sie würden bereits für sich genommen bewusste Zustände genau desjenigen Bewusstseins sein, welches sie als seine eigenen Zustände miteinander verknüpft. Andernfalls resultiert abermals ein infiniter Regress. David Hume, selbst ein Vertreter dieses Denkens, bekannte, »that this difficulty is too hard for my understanding«.[44] Weder

42 Theorien, die das Bewusstsein mittels Repräsentationstheorien zu erfassen versuchen, sind nichts Neues. Bereits John Locke oder David Hume schlugen Vergleichbares vor. Die Schwierigkeiten sind seitdem weitestgehend dieselben geblieben. Tiefgreifende Kritiken an diesem Modell reichen von Thomas Reids *Essays on the Intellectual Powers of Man*, besonders der dritte Essay (in: *The Works of Thomas Reid*, hrsg. v. Dugald Stewart, New York 1822, Bd. 2, S. 127–162) bis beispielsweise Charles Larmores Studien in *Das Selbst in seinem Verhältnis zu sich und zu anderen* (Frankfurt a. M. 2017).
43 Franz Brentano: *Psychologie vom empirischen Standpunkt*, Bd. 1, hrsg. v. Oskar Kraus, Hamburg 1924 (Philosophische Bibliothek 192), S. 171. **44** David Hume: *The Philosophical Works*, Vol. 2: *Treatise of Human Nature II–III*, Edinburgh 1826, S. 559.

aus den im 17. und 18. Jahrhundert gepflegten mechanistischen Vorstellungen noch aus seit dem Zweiten Weltkrieg sich verbreitenden kybernetischen Konzepten lässt sich Bewusstsein herleiten.

Gibt es allerdings ein mit Bewusstsein ausgestattetes Selbst, wird dieses wohl entgegen Metzingers Behauptung, das Gehirn benutze das Selbst und nicht umgekehrt,[45] tatsächlich sein Gehirn benutzen können; andernfalls wäre das Vorhandensein von Bewusstsein vollkommen gleichgültig.[46] Falls tatsächlich, wie es die Maschinenauffassung vom Gehirn verlangt, ausschließlich das Gehirn das Selbst benutzt, wären Metzingers Theorie und verwandte Ansätze nicht durch Gründe, nicht durch den Aufbau ihrer Argumentation verstehbar. Der einzig adäquate Zugang führte über die Frage, welchen Nutzen Metzingers Gehirn von der Theorie hat. Ohne einen solchen Nutzen hätte es die Theorie nicht entwickelt oder sie wieder verworfen. Das würde auch für jede andere Theorie gelten, ob sie nun mit diesem Theorientypus übereinstimmt oder unvereinbar mit ihm ist; entscheidend wären nicht Richtigkeit oder Falschheit, sondern die Tatsache, dass ein Exemplar Information verarbeitender Biosysteme die von ihm entworfene oder angeeignete Theorie benötigt. Die Haltung zu bestehenden Theorien bemäße sich an deren Nutzen für das eigene Gehirn; Zustimmung, Modifikation oder Ablehnung aus anderen Gründen wären weder erforderlich noch möglich.

Ein Argument schon auf phänomenologischer Ebene, mit dem die Grenzen zwischen KI und menschlichem Geist aufgelöst werden, baut auf der Annahme auf, dass wir nicht zeigen könnten, uns nicht in einer simulierten Wirklichkeit zu befinden. Damit ist der kategoriale Schritt von der Wirklichkeit zu den Wirklichkeitssimulationen oder -modellen von KI-Systemen zu einem allenfalls graduellen geschrumpft. Insofern als die Wirklichkeit das »no one«, welches das Subjekt ist und daher nicht ist, immer schon als Simulation umgibt, kann auch jedem Skeptizismus in Bezug auf die Existenz der Wirklichkeit ausgewichen werden. Ob das »no one« nun als eine Computer-Simulation innerhalb einer simulierten Welt vorhanden ist oder in der Welt, die es für die materielle hält, stets ist es umgeben von etwas, das es für wirklich hält. Auf einmal erscheinen wie durch Zauberhand philosophischer Konstruktivismus und Realismus äquivalent. Das Resultat dieses Denkens ist paradox. Neben dem Selbst und seiner Wirklichkeit wäre auch das Gehirn, welches die Simulation herstellt, bereits eine Simulation. Von Gehirnen weiß nur das Selbst. Die Theorien und Annahmen über Gehirne und deren Funktionsweise sind Erzeugnisse bewussten Denkens und Forschens, nicht selbst physikalische und biochemische Tatsachen und ebensowenig durch diese zu erklären. Desgleichen weiß nur ein menschliches Selbst von der auf Computersystemen simulierten KI. Das Gehirn, dem das Konstruieren des Selbst zugeschrieben wird, gehörte demnach seinerseits zu den Konstruktionen, die es hervorbringt.

45 Metzinger: *Der Ego-Tunnel* (wie Anm. 40), S. 282. **46** Vgl. Thomas Fuchs: *Das Gehirn – ein Beziehungsorgan. Eine phänomenologisch-ökologische Konzeption*, 5. Aufl. Stuttgart 2017, S. 76–85.

Mit der Konzentration auf beobachtbares Verhalten durch die Reduktion des Geistes auf funktionale Prozesse eines Programms – etwas, das es in einem Gehirn gar nicht gibt[47] – erreicht die KI-Forschung etwas anderes als eine Erklärung und ein Verstehen von Geist und Subjektivität. Der Ertrag dieser Prämissen besteht in einer suggestiven Umkehr der Beweislast, die schon Turing in *Computing Machinery and Intelligence* (1950) als Grundlage für Erklärungen von Denken und Intelligenz deklariert. Misslich aus Sicht dieses Denkens ist zunächst, dass Menschen keine vollständig formulierten Regeln haben, an denen sie ihre Handlungen ausrichten würden, zumal eine Handlung sich nicht in der Ausführung eines bestimmten Verhaltensprogramms erschöpft. Ohne die Situation, in der eine Handlung genau diese Handlung ist, ist sie nicht als diese oder jene Handlung bestimmt. Die Situation, zu der gewöhnlich lauter kontingente Faktoren gehören, ist nicht Teil eines mental festgelegten oder programmierten Verhaltensablaufs, sie ist Teil der Außenwelt, auf die das Subjekt sich intentional richtet. Außerdem lässt sich Befolgen nicht an Handlungen beobachten. Die Umkehr der Beweislast besteht darin, von Verhaltensgesetzen auszugehen. Verhaltensgesetze sind beobachtbar, und man kann nie sicher sein, Verhalten genügend häufig und gründlich genug beobachtet zu haben, um zu behaupten, es gäbe für dieses oder jenes Verhalten keine Gesetzmäßigkeit.[48] Jedes Subjekt, das darauf beharrt, aus Gründen und Absichten heraus zu handeln, die sich aus seiner Auseinandersetzung mit anderen Subjekten und der äußeren Welt ergeben, steht unter Zugrundelegung von Turings Argumentation unter permanentem Rechtfertigungszwang, nicht doch lediglich ein Programm von Verhaltensgesetzen zu erfüllen. Die Tendenz, der das Verständnis des Menschen von sich selbst unterworfen ist, indem er seine Kompatibilität mit der Technik erhöht, fasst Konrad Paul Liessmann prägnant zusammen: »Im Bild, das der Mensch von sich entwirft, kommt er selbst immer weniger vor.«[49]

Kohärent sind die Argumente der KI-Forschung häufig nicht. Eine Beschreibung von beobachtbaren Vorgängen oder beobachtbarem Verhalten gibt keine Gründe, sich zur KI in welcher Weise auch immer zu positionieren, ihr zu unterstellen, sie könne komponieren, Gedichte schreiben oder neue Erkenntnisse gewinnen. Wenn es um die Einstufung der KI als intelligent, als denkend, als Emotionen erkennend usw. geht, werden entsprechend immer auch Argumente für die Leistungen der KI angegeben, in denen doch wieder auf Intentionales, auf Einstellungen, Bedürfnisse Bezug genommen wird. Das ergibt sich schon aus dem Umstand, dass von bloßem Verhalten nicht auf eine bestimmte Handlung geschlossen werden kann. Auch ergibt sich aus bestimmten Gesichts-, Stimm- oder Haltungsveränderungen, aus Änderungen des Blutdrucks oder der Pulsfrequenz keine definierbare Emotion oder Einstellung. Dieses Problem bildet einen zentralen Kritikpunkt etwa

47 Vgl. John R. Searle: *Die Wiederentdeckung des Geistes*, übers. v. Harvey P. Gavagai, München 1993. **48** Turing: *Computing Machinery and Intelligence* (wie Anm. 17), S. 72/73–76/77. **49** Konrad Paul Liessmann: *Lob der Grenze. Kritik der politischen Unterscheidungskraft*, Wien 2012, S. 58.

an der Emotionserkennung durch KI-Systeme.[50] Interpretierbar wird Verhalten erst, indem ihm Intentionalität zuerkannt, die Situation, in der es stattfindet, berücksichtigt und das Gegenüber, an das sich die Handlung richtet, einbezogen wird. Ein bloß zufälliges oder reflexartiges Heben der Hand sagt nichts aus, das Heben der Hand zum Gruß ist eine interpretierbare Handlung und baut eine auf Antwort hoffende Beziehung zu einem anderen Subjekt auf. Bereits das Überspringen der zweiten Person zu einer Beschreibung aus Perspektive der dritten Person ist verfehlt; an beobachtbarem Verhalten, an empirischen Tatbeständen kommt durch Handlungen nicht unbedingt etwas hinzu, aus dem sich erschließen ließe, um welche Handlung es sich jeweils handelt. Des Weiteren können allein aus beobachtbaren Merkmalen an einem Menschen keine komplexen Emotionen erschlossen werden. Das, was die Emotionen sind, ergibt sich nicht ohne das intentionale Objekt, auf das sie gerichtet sind. Diese Problematik setzt jedoch nicht erst auf der Ebene der Interpretation von Mustern ein, die ein KI-System bereits aus den ihr gegebenen Daten gewonnen hat, vielmehr ist das Erkennen von Mustern selbst bereits ein interpretativer Akt. In welchem Grad Mustererkennung interpretativ ist, hängt vom jeweiligen Gegenstandsbereich ab. Freilich lassen sich aus einem gegebenen Satz Daten unterschiedliche Muster ableiten; welche Muster relevant sind und in welcher Weise, hängt von dem Zusammenhang ab, in den sie gehören, ebenso von den an sie angelegten Begriffen, Intentionen und epistemischen, ästhetischen, moralischen Werten. Beliebig sind Musterzuschreibungen keineswegs, jedoch an Gegenständen oder Ereignissen eindeutig bestimmbar erst innerhalb der jeweiligen Konstellation der genannten Faktoren.[51] Bei Mustern im menschlichen Handeln kommt hinzu, dass dieses sich, werden entsprechende Theorien über Denken, Verhalten etc. geglaubt, den Annahmen über die Maschinenförmigkeit des Denkens und Verhaltens anpasst. Vermutlich ist Mustererkennung in Bezug auf menschliche Emotionen und Verhaltensweisen auch einer Anpassungsleistung des Menschen an die Erfordernisse der Technologie geschuldet.[52]

Freilich bemühen sich einige KI-Forscher, das Argument der Intentionalität zu umgehen, indem sie behaupten, der bloße Anschein genüge, um den Maschinen eine bestimmte Handlung zu unterstellen und die mögliche Menschenähnlichkeit der Maschinen anzuerkennen; dies allein, da Menschen gar nicht anders könnten, als bestimmtem maschinellen Verhalten Handlungscharakter zuzuschreiben, etwa das Ausstrecken des Roboterarms als das Reichen der Hand zu betrachten.[53] Was dem Anschein genügt, ist allerdings äußerst variabel. Über lange Zeiträume der Menschheitsgeschichte entzog sich nichts diesem Anschein; Flüsse, Steine, Täler, Wolken, Sterne, Pflanzen, Tiere, alles wurde als dem Menschen ähnlich angesehen

50 Catrin Misselhorn: *Künstliche Intelligenz und Empathie. Vom Leben mit Emotionserkennung, Sexrobotern & Co.*, 2. Aufl. Ditzingen 2023, S. 22f. **51** Ludwig Wittgenstein: *Philosophische Untersuchungen. Kritisch-genetische Edition*, hrsg. v. Joachim Schulte, Frankfurt a. M. 2001, bes. S. 1024–1059. **52** Vgl. Daniel Strassberg: *Spektakuläre Maschinen. Eine Affektgeschichte der Technik*, Berlin 2022, S. 388–401. **53** Martin Gessmann: »Maschinenethik – Vom Schmuddelkind zum Hoffnungsträger« (Teil 1), in: *Philosophische Rundschau 66*, 2019, S. 89–117, hier S. 92.

und als Person aufgefasst, versehen mit Geist, Intentionalität, kommunikativen Fähigkeiten, Subjektivität;[54] was die Lebewesen angeht mit weitaus mehr Recht als heute die KI-Systeme. In den Heiligenbildnissen im Mittelalter waren nach damaligem Verständnis die Heiligen real gegenwärtig. Tatsächlich bedarf es keiner besonderen Verhaltensweisen, keiner besonderen Menschenähnlichkeit, um dem Anschein folgend als Subjekt begriffen zu werden; die Spanne an Beispielen reicht vom Teddybär bis zu ELIZA, einem Programm, mit dem Menschen in natürlicher Sprache in einen einfach aufgebauten, aber illusionären Dialog treten können, entwickelt 1966 von Joseph Weizenbaum. Heute würde kaum jemand diesem Programm noch Menschenähnlichkeit unterstellen, zumal die Programmroutinen weder Notwendigkeit noch Möglichkeit von Bewusstsein implizieren.[55] Als Weizenbaum ELIZA testete, glaubten die Probanden nicht, dass sie sich tatsächlich an eine Maschine gewandt hatten. Weizenbaum hatte sein Programm nach Eliza Doolittle, dem Blumenmädchen aus George Bernard Shaws Theaterstück *Pygmalion* benannt.[56] In dem Stück erzieht der überhebliche Doktor Higgins die sich in Gossensprache artikulierende Eliza dazu, sich an den Konversationen der höheren Gesellschaft beteiligen zu können, ignoriert sie jedoch als Person, das Haben von Gefühlen spricht er ihr ab. Im Grunde behandelt er sie wie eine Maschine, nicht aus Böswilligkeit, sondern aus Unfähigkeit, in ihr ein Subjekt zu erblicken. Tatsächlich erlernt sie die gehobene Konversation, allerdings lediglich formal, sie vermag schließlich, sich in Gesellschaft zu äußern, versteht jedoch nicht, was gesprochen wird. Ähnlich ELIZA, das Programm basiert auf syntaktischen Regeln und ist eingerichtet zur Musterabgleichung, verfügt jedoch nicht über Inhalte. Gleichwohl fühlten sich Menschen von dem Programm verstanden. Zwischen ELIZA und Hoffmanns Olimpia, der Nathanael aufgrund ihrer einzigen Lautäußerung »Ach« Seelentiefe zuspricht, besteht ein gradueller, kein qualitativer Unterschied. Es ist viel schwieriger, in den Dingen, welcher Art sie auch sein mögen, keine Subjektivität zu entdecken, als sie zu finden. Das Argument, sich mit dem Anschein zu begnügen, um KI-Systemen Menschenähnlichkeit unterstellen zu können, bedeutet schlicht eine Aufforderung zum Verzicht auf jeglichen Gebrauch des menschlichen Abstraktionsvermögens. Was tatsächlich einer Erklärung bedarf, ist die sich in dieser Annahme äußernde Regression des Denkens.

Die rhetorische Strategie, mittels derer die Differenz zwischen Mensch und Maschine minimiert oder gänzlich aufgehoben wird, besteht in zwei aufeinander abgestimmten Beschreibungen, einer Beschreibung des Menschen, die ihn der Maschine annähert, und einer der Maschine, die sie vermenschlicht.[57] Dieses rhe-

54 Jüngst dazu Marshall Sahlins: *Neue Wissenschaft des verwunschenen Universums. Eine Anthropologie fast der gesamten Menschheit*, übers. v. Heide Lutosch, Berlin 2023, S. 105–122. **55** Klaus Mainzer: *Künstliche Intelligenz – Wann übernehmen die Maschinen?*, 2. Aufl. Berlin 2019, S. 55–58. **56** Joseph Weizenbaum: *Computer Power and Human Reason. From Judgment to Calculation*, New York 1976, S. 188f. **57** Die zum Teil absurden Argumentationen der KI-Forschung stellt prägnant dar Michael Funk: *Künstliche Intelligenz, Verkörperung und Autonomie*, Wiesbaden 2023 (Theoretische Probleme – Grundlagen der Technikethik 4), S. 74–76.

torischen Musters bedient sich bereits Turing; am offensichtlichsten, wenn er lernende Maschinen metaphorisch als »child machines« bezeichnet und sie als besonders schutzbedürftig schildert, da angesichts der Art und Weise, wie sie lernen würden, sich menschliche Kinder über sie lustig machen könnten.[58] Hoffmanns Nathanael verhält sich kaum anders, als es diese rhetorische Strategie nahelegt. Ihm erscheinen die Menschen, insbesondere seine Verlobte Clara, maschinenhaft, die Automatenpuppe Olimpia hingegen beseelt und menschlich. Turings Metapher von den »child machines« enthält allerdings noch einen weiteren Aspekt, den des Mängelwesens. Den Menschen als Mängelwesen zu verstehen, gehört als grundlegender Topos zur modernen Anthropologie. Prinzipiell wird der Mensch nicht dadurch, was er ist, beschrieben, sondern dadurch, was ihm fehlt. Es wird versucht, die Unvollständigkeit mittels Technik zu kompensieren.[59] Die Definition des Menschen durch Mangel lässt jedoch nicht einmal eine Bestimmung dessen zu, was kompensiert werden könne, so dass die technologische Vervollständigung des Menschen unabschließbar ist. Die Existenz eines positiv durch Eigenschaften bestimmten Menschen verschiebt sich bis in eine unendlich ferne, nie erreichbare Zukunft, hierher gehört gegen die eigene Intention Ernst Blochs Beschreibung des Menschen als ein Wesen, das »eine Aufgabe« sei, »ein riesiger Behälter voll Zukunft«.[60] Um ihr näherzukommen, muss er das aufgeben, was er schon ist, ein Subjekt, und alles, was er schon hat, insbesondere die Erste-Person-Perspektive und die Zweite-Person-Perspektive. Auf diese Weise wird er Bestandteil eines alles übergreifenden technischen Prozesses, für den bei konsequenter Weiterführung von Turings Metapher dieselbe Mängeldefinition gilt wie für den Menschen; allein, die Technik muss auf nichts verzichten, denn sie ist nicht für sich. Das »being no one« verbände sich mit einem »never being«.

Möglicherweise aber eignet der KI neben ihrem ideologischen auch ein kritisches Potential. Legt Hoffmann seiner Figur Nathanael die Worte in den Mund, das poetische Gemüt erschließe sich nur dem gleich organisierten, so lässt sich dies auch als Kritik an einer Poesie verstehen, deren Hervorbringung eben vielleicht doch einem maschinenhaften Vorgehen stärker ähnelt, als dem Dichter oder Komponisten lieb sein kann. Die Argumentationsfiguren, die die KI dem Menschen und den Menschen der Maschine annähern, erfüllen eine kritische Funktion gegenüber einer Kunst, der allzu leicht »Kreativität« attestiert wird; ein Ausdruck, der ohnedies zwischen bloßer Findigkeit und phantasievollem Gestalten changiert, seines außeralltäglichen Moments jedenfalls ledig geworden ist. Scheinbar der Rationalität, dem Kalkül enthoben, kalkuliert Kunst vielleicht doch zu umstandslos ihren Platz im Kunstbetrieb. Nicht immer werden Komponisten, Dichter, Künstler den Ansprüchen gerecht, die die Kunst idealiter von ihnen fordert. Das zu verlangen wäre

58 Turing: *Computing Machinery and Intelligence* (wie Anm. 17), S. 88/89–92/93. **59** Das vielleicht bekannteste Beispiel ist Arnold Gehlen: *Gesamtausgabe*, Bd. 3: *Der Mensch. Seine Natur und seine Stellung in der Welt*, hrsg. v. Karl-Siegbert Rehberg, 2 Teilbde., Frankfurt a. M. 1993. **60** Ernst Bloch: *Gesamtausgabe*, Bd. 5: *Das Prinzip Hoffnung*, Kap. 1–37, Frankfurt a. M. 1959, S. 135.

unbillig und setzte einen nie ermüdenden, nie nachlassenden, stets auf der Höhe befindlichen menschlichen Geist voraus, erforderte also eher eine Maschine als ein lebendiges Subjekt. Eingespannt aber in einen Musik- und Kunstbetrieb, der ständig des Neuen, sich von anderem Unterscheidenden bedarf, in dem tatsächlich nebensächlich ist, was ein Künstler in seiner Kunst aussagen könnte, befindet sich Kunst selbst eingebunden in von Strategie und Kalkül getriebene, maschinenförmige Strukturen. Eine für den maschinenähnlichen Betrieb geschaffene Kunst wird leicht selbst maschinenähnlich, indem sie sich weitgehend auf die Erzeugung von Mustern beschränkt, die sich von den von anderen Künstlern angewandten Mustern möglichst unterscheiden und zugleich selbst erkennbare Muster ergeben. Ein Hang der Kunst, die Anwendung abstrakter Verfahren zur Herstellung von Unterschieden und wiedererkennbaren Mustern in den Vordergrund zu rücken, liegt demnach nahe. Anderes unterbliebe, so etwa die Untersuchung von sozialen, gesellschaftlichen, politischen oder elementar menschlichen Problemen, die den Künstler bewegen, mit ästhetischen Mitteln. Durchaus denkbar, dass der Kunstbetrieb mehr Kunst verlangt, als die Künstler über Ausdruckskraft verfügen. Die Mittel verdrängen die Zwecke, zu deren Umsetzung sie eigentlich dienen sollten. Diese Tendenz lässt sich nicht einfach den ›bösen‹ Musikverlagen, Kunsthändlern oder Medienunternehmen, dem ›schlechten‹ Publikum, das immer dasselbe möchte, anlasten; vielmehr ist sie Folge des Umstands, dass eine größere Anzahl von Menschen die Musik, das Malen, das Schreiben, das Filmen als ihre Haupttätigkeit wählt und sich die dazu erforderlichen Institutionen etabliert haben. Die Forderung steckt gewissermaßen in der Sache, nur darf diese sich nicht durchweg in der Erfüllung dieser Forderung erschöpfen. Vor solchem Hintergrund stellt sich allerdings die Frage, inwiefern eine allein von Menschen komponierte Musik mehr aussagt, mehr ausdrückt als eine beliebige von KI erzeugte Komposition. Es mag also vorkommen, dass die Differenz zwischen menschlichem Subjekt und Maschine leicht überschritten ist; dies aber nicht, weil die KI Geist und Phantasie besäße, denken oder Sinn hervorbringen könnte, sondern weil letzteres auch Menschen unter bestimmten Bedingungen versäumen (müssen).

Weshalb funktioniert der Behaviorismus (nicht nur) in der Kunst nicht?

Die behavioristische Sichtweise, die erforderlich ist, um der KI Geist und Denken attestieren zu können, scheitert auch angesichts ästhetischer Hervorbringungen durch KI-Systeme. Mit der rein verfahrenstechnischen Auffassung vom künstlerischen Schaffen verbindet sich der Behaviorismus unschwer. Um die resultierenden Schwierigkeiten zu veranschaulichen, seien zwei Gedankenexperimente angeführt, deren erstes von Hilary Putnam stammt. Sein Buch *Vernunft, Wahrheit und Geschichte* beginnt mit der folgenden Frage:

> »Eine Ameise kriecht im Sand umher, und beim Kriechen zieht sie eine Linie in den Sand. Durch reinen Zufall verlaufen die Kurven und Überschneidungen der Linie derart, daß sie zum Schluß aussieht wie eine erkennbare Karikatur von Winston Churchill. Hat die Ameise ein Bild von Winston Churchill gezeichnet? Ein Bild, das Churchill *abbildet*?«[61]

Die mögliche Antwort auf die Frage, die Ameise hätte gar nicht Churchill zeichnen wollen, sondern irgendwelche Ameisendinge tun, ist für Putnam nicht entscheidend, Subjektivität und Intentionalität berücksichtigt er an dieser Stelle nicht. Er möchte drastischer behaupten, dass die Ameise Churchill nicht nur nicht zeichnen will, sondern dass sie dies nicht einmal wollen könnte. Der Sinn einer Zeichnung, die Karikatur von Winston Churchill zu sein, kann nur dann erfüllt sein, wenn der Zeichner weiß, wer Winston Churchill ist, wenn er ihn gesehen hat, seine Politik kennt und als guter Karikaturist seine Mimik, sein Gebaren etc. beobachtet hat. Die Ameise kennt aber Churchill gar nicht und hat vermutlich gar keine Vorstellungen von Politik. Ihrer Spur im Sand fehlt, und das ist Putnams Antwort, der Bezug auf Churchill. Sinnvolles Handeln, Sprechen, Musizieren, Malen, Filme drehen etc. ist ohne einen Bezug auf die Welt und auf das Geschehen in der Welt nicht verständlich. Das gilt auch von fiktiven Geschichten; Sherlock Holmes' Verbrecherjagden durch London wären ohne eine grundsätzliche Vorstellung und Erfahrung von Straßen, Häusern, Nebel und Nacht, von Sterblichkeit und (gewaltsamem) Tod, und was dergleichen mehr zu den Geschichten gehört, nicht verständlich. Der Ausdruck von Musik könnte nicht aufgefasst werden ohne die Fähigkeit, seelische Spannung und Entspannung zu erleben, ohne über eine Erfahrung mit bestimmten Bewegungscharakteren zu verfügen, ohne die Fähigkeit, Analogien von Musik mit Geschehen in der Welt zu bilden etc. Ohne einen Erfahrungshorizont auf die Außenwelt lässt sich keine sinnhafte Äußerung erwarten. Welchen Weltbezug hätte die KI? Wie könnte sie diesen kompositorisch umsetzen?

Ein schon vergleichsweise altes, aber beliebtes Beispiel, das die Kreativität (ein Terminus, der selbst noch einmal einer kritischen Betrachtung unterzogen werden müsste) von Maschinen belegen soll, sind von KI generierte Kompositionen, die eine Mehrheit von Hörern für Musik von Johann Sebastian Bach hält, oder zumindest zugesteht, Bach hätte genau solche Musikstücke komponieren können. Anders als manche Liebhaber der KI anscheinend annehmen, pflegte Bach keinen von Algorithmen geprägten Kompositionsstil, dazu hätte er Algorithmen kennen und interpretieren müssen. Auch programmierte er seine Kompositionen nicht,[62] denn weder gibt es im Gehirn so etwas wie Software noch ist die Notenschrift mit einer programmierten Software zu vergleichen, zumal sie die für das Musizieren erforderlichen Interpretationen und Handlungen unterbestimmt. Selbstverständlich

61 Hilary Putnam: *Vernunft, Wahrheit und Geschichte*, übers. v. Joachim Schulte, Frankfurt a. M. 1982, S. 15. **62** Vgl. Marcus du Sautoy: *Der Creativity Code. Wie künstliche Intelligenz schreibt, malt und denkt*, übers. v. Sigrid Schmid, München 2021, S. 193–197.

gibt es komplexe Regeln, die Bach beherrschte, und der Aspekt der Verfahrensrationalität ist bei Fugen- und Kanonkompositionen nicht gering zu veranschlagen – weswegen die Kanonkünste im Übrigen schon im 17. Jahrhundert bisweilen heftig kritisiert wurden –, jedoch garantiert noch die gründlichste Befolgung der Regeln keine gelingende Komposition. Die Wahl etwa des Soggettos hängt von ganz verschiedenen Aspekten, vom zugrundeliegenden Text oder vom intendierten Ausdruck, von einer Aussage ab, diese bestimmen ebenfalls die zu wählenden Kompositionsverfahren mit. Dies zu können setzt voraus, dass der Komponist Erfahrung mit dem musikalisch Ausgedrückten und Ausgesagten hat. Allein syntaktische und probabilistische Verfahren im Gehirn zu haben, würde diese Anforderung nicht erfüllen. Angenommen, die KI generierte nun ein Stück Musik, das tatsächlich dieselbe Folge von Takten und Noten, Harmonien etc. aufwiese wie eine existierende Bach-Fuge. Ließe sich das bereits als ein Kunstschaffen der KI betrachten? Welches Problem trieb die KI beim Generieren der Komposition an? Etwa die Frage nach einer von Gott sinnvoll geordneten Welt, oder die Leiden und Schmerzen Jesu am Kreuz? Hätte sie auch »Soli Deo Gloria« unter die Komposition setzen wollen? Ist KI überhaupt von einem Problem bewegt? Oder wäre man, getäuscht über den Ursprung der KI-Komposition, unfreiwillig in der Situation des Kaffeesatzlesens? Zumal, wie oben ausgeführt, geklärt werden müsste, was die KI an den Klangkombinationen in welcher Weise als Komposition meinen würde, wenn sie meinen könnte. Wichtig ist nicht die Abgeschmacktheit, Bach nachzukomponieren; es mag ja durchaus eine bewundernswerte technische Leistung sein, eine Maschine zu konstruieren, die vergleichbare Strukturen hervorbringt. Der Einwand lautet, dass musikalisches Komponieren nicht in der Ausführung von definierten Regeln nach bestimmten Wahrscheinlichkeitsgewichtungen der möglichen Entscheidungen besteht und auch nicht notwendig die Einbeziehung eines Zufallsmoments in das Kompositionsverfahren beinhaltet; Abweichungen vom Schema können auf inhaltlichen oder ästhetischen Gründen beruhen, die der Komponist hatte. Nicht die gesamte Komposition, für die KI zum Einsatz kam, beruht notwendigerweise auf KI-Systemen. KI lässt sich einsetzen auch zu präkompositorischen Materialkombinationen und -strukturierungen. Aber das Bestreben, Deep Learning so einzurichten, dass es die ästhetischen Präferenzen eines Künstlers erkennt und ihm dementsprechendes präkompositorisches Material zur weiteren Formung generiert,[63] hat Einfluss auf die weitere Werkgenese. Beispielsweise erhalten die Gewohnheiten des Künstlers dadurch ein höheres Gewicht als die Abweichung von ihnen und die KI entwickelt möglicherweise einen Datendoppelgänger des Künstlers.

Menschliche Äußerungen lassen sich nicht auf das Erfüllen eines Regelkatalogs reduzieren, selbst wenn sie unter bestimmten Bedingungen einen Hang dazu aufweisen. Im Gehirn existieren auch keine Musik- oder Kunst-Algorithmen, die

63 Jon McCormack / Andy Lomas: »Understanding Aesthetic Evaluation Using Deep Learning«, in: *Artificial Intelligence in Music, Sound, Art and Design. 9th International Conference EvoMUSART 2020. Seville, Spain, April 15–17, 2020. Proceedings*, hrsg. v. Juan Romero / Anikó Ekárt / Tiago Martins / João Correia, Cham 2020 (Lecture Notes in Computer Science 12103), S. 118–133.

eine Hervorbringung von Musik oder Kunst ermöglichen würden. Dazu müsste es eine Instanz im Gehirn geben, die Algorithmen interpretieren könnte. Nun bilden mathematische Funktionen allenfalls Strukturanalogien zu etwas Realem, sie sind es nicht. So sind auch Algorithmen nicht dasselbe wie die Gedanken, die zum Komponieren genau dieses Musikstücks führen. Ein Gedankenexperiment von Arthur C. Danto zeigt, dass die Hervorbringung von Kunst sich prinzipiell nicht allein auf Berechnungen oder andere Vorgänge im Subjekt zurückführen lässt, und nicht auf Rechenvorgänge in einer Maschine. Sein Argument ergänzt sich mit Putnams Annahmen vom Außenweltbezug des Denkens.

> »Wir wollen ein Gemälde betrachten, das einst von der dänischen Geistesgröße Sören Kierkegaard beschrieben wurde. Es ist ein Gemälde der Israeliten, die das Rote Meer durchquerten. Bei seiner Betrachtung sieht man etwas ganz anderes als das, was man bei einem Gemälde mit diesem Sujet erwarten würde […]. Statt dessen gibt es hier ein Viereck aus roter Farbe, zu dem der Künstler erklärte: ›Die Israeliten waren schon vorbeigezogen, und die Ägypter waren ertrunken.‹ Kierkegaard sagt dazu, das Ergebnis seines Lebens gleiche diesem Gemälde. […]
>
> Stellen wir nun neben das von Kierkegaard beschriebene Gemälde ein anderes, genau gleiches, von dem wir annehmen, es sei von einem dänischen Porträtisten gemalt worden, der mit ungeheurem psychologischen Einfühlungsvermögen ein Werk mit dem Titel ›Kierkegaards Stimmung‹ geschaffen hat. Und in diesem Sinne wollen wir uns nun eine ganze Reihe von roten Rechtecken vorstellen, eines neben dem anderen. Neben diese beiden stellen wir den ›Roten Platz‹ […], ein geistreiches Stück Moskauer Landschaft, das den ersten beiden ebenso ähnlich ist, wie sie ihrerseits einander genau gleichen. Unser nächstes Werk ist ein minimalistisches Musterbeispiel geometrischer Kunst, das zufällig denselben Titel trägt, ›Roter Platz‹ […].«[64]

Tatsächlich gleiche Bilder vermögen durchaus völlig unterschiedliche Bilder zu sein, ihr gleiches Rot vermag ganz unterschiedliche Inhalte und unterschiedlichen Ausdruck zu transportieren. Durch einen Algorithmus, durch eine beliebige Anzahl von Regeln kann dieser Umstand nicht eingeholt werden. Anders formuliert, die Reduktion der Hervorbringung von Kunst auf Algorithmen ist nicht triftig, da offenkundig die Ausführung eines bestimmten Algorithmus mit exakt denselben Eingabewerten und denselben Ausgabeschemata unterschiedliche Kunstwerke hervorbringen könnte, diese also durch den Algorithmus und die Eingabewerte unterdeterminiert wären. Die Schwierigkeit ästhetischer Gegenstände wie Gemälde oder Musikstücke liegt in dem voneinander unlösbaren Wahrnehmen und Interpretieren der Werke. Den Akt des Interpretierens vollführt man, indem man ein Gemälde ›so‹ sieht und eine Musik ›so‹ hört. Das Betrachten der drei Gemälde in Dantos Gedankenexperiment darf man sich nicht so vorstellen, als interpretiere der Betrachter drei in der Wahrnehmung miteinander übereinstimmende Gegenstände

64 Arthur C. Danto: *Die Verklärung des Gewöhnlichen. Eine Philosophie der Kunst*, übers. v. Max Looser, Frankfurt a. M. 1991, S. 17f.

anders, vielmehr stimmen die drei ästhetischen Gegenstände, die nicht zu verwechseln sind mit ihren materiellen Trägern, nicht überein. Anders gesagt, sich vollkommen gleichende materielle Träger können dadurch, dass sie unterschiedlich gesehen oder gehört werden können, ästhetisch differierende Werke transportieren. Ein sehr einfaches Beispiel wäre der von Ludwig Wittgenstein gründlich betrachtete H-E-Kopf, der sowohl als Bild eines Hasen wie als das einer Ente gesehen werden kann; als Hasenbild gesehen wird die Beschreibung des H-E-Kopfes völlig anders lauten, als wenn er als Entenbild gesehen wird. Weder Enten- noch Hasenbild befinden sich im Kopf dessen, der den H-E-Kopf als Enten- oder Hasenbild sieht. Für beide Sehweisen muss er sein Sehen auf den H-E-Kopf richten, der sich nicht verändert. Ein Wechsel vom Hasen- zu einem Entenbild ließe ein Rekurrieren auf den H-E-Kopf – gezeichnet mit Bleistift auf ein Blatt Papier und auf dem Tisch liegend –, der sich nicht geändert hat, nicht zu. Zu sehen ist nur der H-E-Kopf in einer spezifischen Weise, kein Bild im Kopf.[65] Grundsätzlich weist der H-E-Kopf auf Merkmale eines jeden Kunstwerks. In der Musik betrifft solche Ambiguität bereits die Töne, sobald sie auf Tonsysteme bezogen werden. Real erklingen zumindest zum Teil andere Töne als diejenigen, die gemeint sind, und die gleichen real erklingenden Töne können verschiedene gemeinte Töne repräsentieren. Solche Ambiguitäten, ob beispielsweise ein Ton als eher *dis* oder als *es* gehört werden sollte, lassen sich – das hängt vom musikalischen Kontext ab – nicht immer eindeutig entscheiden. Überdies enthalten die realen Töne ihrerseits mehr mögliche Aspekte, als sie von Tonsystemen und anderen expliziten und impliziten musikalischen Regeln festgelegt sind. So ist etwa das räumliche Verständnis der Tonordnung eine recht späte Erscheinung, in der griechischen Antike etwa unterschied man nicht hohe und tiefe Töne voneinander, sondern scharfe und schwere, zwar korrelieren ›scharf‹ und ›hoch‹ beziehungsweise ›schwer‹ und ›tief‹ weitgehend, wenngleich nicht vollkommen miteinander, doch werden sie unterschiedlich gehört. Musikalisch entscheidend ist nicht der Klang als naturalistische Gegebenheit, sondern der Sinn, in welchem er gemeint ist. Weder geht der in den gemeinten Klängen festgelegte Sinn in den realen Klängen auf – es bleibt ein ideeller Überschuss –, noch fügen diese sich vollkommen den verschiedenen musikalischen Ordnungen ein, ihnen eignet ein empirischer Überschuss.[66] Wie für jedes Erleben von Kunst gilt daher auch für Musik: »Nur von Einem, der das und das kann, gelernt hat, beherrscht, hat es Sinn zu sagen, er habe das erlebt.«[67]

Nicht alles, was an den Bildträgern oder an den musiktragenden Klängen wahrgenommen werden kann, gehört zugleich zum Werk, und was zum Werk gehört, gehört in je spezifischer Hinsicht zu ihm. Ein Musikstück zu verstehen, heißt

65 Zur Aspekt-Wahrnehmung in der Musik vgl. Boris Voigt: »Musik und Musikverstehen bei Ludwig Wittgenstein«, in: *Zeitschrift für Ästhetik und Allgemeine Kunstwissenschaft 52*, 2007, S. 119–131.
66 Vgl. Boris Voigt: »Über das Verhältnis von Musik und Klang. Eine phänomenologische Untersuchung«, in: *Musiktheorie. Zeitschrift für Musikwissenschaft 26*, 2011, S. 169–186, hier S. 176f.
67 Wittgenstein: *Philosophische Untersuchungen* (wie Anm. 51), S. 1049 (a107.3). Vgl. Voigt: »Musik und Musikverstehen bei Ludwig Wittgenstein« (wie Anm. 65).

auch, es in bestimmter Weise zu hören. Wie die Klänge zu hören sind, um ein bestimmtes Musikwerk zu bilden, ist durch sie nur begrenzt festgelegt. So stützt ein neurowissenschaftliches Experiment von Stefan Koelsch und Tom Fritz, welches die Universalität der an die Dur-Moll-Tonalität gebundenen Akkordfunktionen nachweisen sollte, anders als die Autoren schlussfolgern, die ihm zugrundeliegende Hypothese eben nicht. Ihr Experiment führten Koelsch und Fritz in Kamerun bei einem kleinen Volk durch, dem westliche Musik unbekannt war. Anhand allein von abgespielten Klavieraufnahmen vermochten die Versuchspersonen nicht, reguläre und irreguläre Kadenzen auseinanderzuhalten; einige der Kadenzen endigten auf der Tonika, andere etwa auf der Doppeldominante. In einem modifizierten Experiment wurden die Schlussakkorde zusätzlich unterschieden, indem die einen vom Klavier, die anderen von einer Harfe stammten. Die Aufmerksamkeit musste also erst durch eine weitere Differenz gelenkt werden, so dass die Probanden bemerkten, worauf sie ihr Hören abseits ihres gewohnten Fokus richten sollten. Erst nach dieser Modifikation, die allerdings nicht der musikalischen Praxis gerecht wird, funktionierte das Experiment im Sinne der Hypothese.[68] Ein komponierendes oder Musik analysierendes KI-System müsste, wenn es das Komponieren anhand eines Beispielkorpus von Kompositionen erlernt, wissen, welche der vielen feststellbaren Muster zur Komposition gehören und welche nicht. Um ein KI-System berechtigterweise als Musik komponierend aufzufassen, müsste es mitteilen können, welche der möglichen Konstellationen von Klängen und Klangrelationen, mehr noch, welche der verschiedenen Aspekte der Klänge selbst das resultierende Musikstück in welcher Weise als ästhetischen Gegenstand bestimmen. Es müsste angeben können, welches Werk von den Klangkonstellationen, die es erzeugt hat, gemeint ist. Einem KI-System, das regelbasiert angelegt ist, steht eine Antwortmöglichkeit nicht offen, da ihm die Regeln als Regeln nicht thematisch zur Verfügung stehen.[69] Auch fügen sich die musikalischen Regeln nicht zu widerspruchsfreier Kohärenz, das Einhalten der einen Regel kann einer anderen widersprechen. Das KI-System müsste demnach ästhetische Gründe anführen können, warum es bestimmte Entscheidungen getroffen hat. KI-Systeme, deren Erzeugnisse nicht auf regelbasiertem Vorgehen beruhen, sondern auf einem Beispielkorpus von Kompositionen, müssten angeben können, aus welchen ästhetischen Gründen sie die Aspekte an den Klängen ausgewählt haben, die sie zu ihren Kompositionen führen. Behandeln sie die Klänge rein empirisch, sind sie ohnedies nicht beim Komponieren von Musik angelangt.

Nicht zuletzt die Einlassungen Marvin Minskys zur Musik zeigen, wie weit KI-Systeme von musikalischem Denken und Hören entfernt sind.[70] Für die Musikwahr-

68 Stefan Koelsch / Tom Fritz: »Musik verstehen – Eine neurowissenschaftliche Perspektive«, in: *Musikalischer Sinn. Beiträge zu einer Philosophie der Musik*, hrsg. v. Alexander Becker / Matthias Vogel, Frankfurt a. M. 2007, S. 237–264, hier S. 242–245. **69** Zu den rein praktischen Einschränkungen von KI-basiertem Komponieren vgl. Gerhard Nierhaus: *Algorithmic Composition. Paradigms of Automated Music Generation*, Wien / New York 2009, S. 270ff. **70** Marvin Minsky: *The Emotion Machine. Commonsense Thinking, Artificial Intelligence and the Future of the Human Mind*, New York u. a. 2006, S. 196f.

nehmung geht er von einem Modell aus, das wie das Deep Learning von KI-Systemen aufgebaut ist. Einzelne Elemente und Parameter erfasst das System auf der ersten Ebene, auf einer zweiten Ebene nimmt es deren Zusammensetzung zu größeren Einheiten vor, vom Takt über längere Phrasen, schließlich synthetisiert es auf einer weiteren Ebene den ganzen Satz. Nun sind in der Musik die einzelnen Elemente bereits auf Relationen zu anderen Elementen angelegt und auf Relationen, die die Elemente, zwischen denen Relationen bestehen, übergreifend strukturieren. Musik besteht nicht aus isolierten Elementen, die erst nachträglich zu etwas Ganzem zusammengesetzt werden. Nicht einmal jemand, dem jede musikalische Erfahrung fehlt, hört voneinander isolierte Klangereignisse. Weder beginnt noch endigt das Hören mit dem jeweils einzelnen Gehörten, sondern dieses ist von vornherein auf Erwartungen bezogen, die aus vorherigen auditiven oder anderen Ereignissen resultieren, und es ist von vornherein mit Erwartungen an folgende Ereignisse aufgeladen, seien sie noch so unbestimmt.[71] Musik jedenfalls ließe sich nicht – quasi atomistisch – ausgehend von isolierten Klangelementen hören.[72]

Täuschung und Betrug

In Hoffmanns Erzählung *Die Automate* äußert der Musiker Ludwig Kritik an den menschlich aussehenden Musikautomaten als verfehlte Nachahmung von Menschen und menschlichen Handlungen. Genau genommen liegt sogar eine doppelte Täuschung vor: die Täuschung, Menschen würden musizieren, und die Täuschung, es handele sich um menschengemachte Musik. Hoffmann nimmt die Vortäuschung menschlicher, aus dem Inneren heraus vollzogener Handlungen in den Blick und zugleich die Vortäuschung eines menschlichen Individuums durch die mechanischen Künste. Seine Figur Ludwig zeigt sich enttäuscht angesichts der Nachahmung, insofern sie nur das simuliert, was Menschen ohnedies unternehmen. Von einer ähnlichen Enttäuschung berichtet der Komponist David Cope, der in den 1980er Jahren eine »Emmy« genannte komponierende KI entwickelte. Anlässlich eines Konzerts, in dem er ein von Emmy generiertes Musikstück spielen ließ, sprach ihn ein Zuhörer an, der ihm sagte, er habe seit langem kein so schönes Musikstück mehr gehört. Nachdem die Genese des Stücks in einem Vortrag im Anschluss an das Konzert erläutert wurde, änderte der Zuhörer seine Ansicht zu dem Stück radikal. Es sei oberflächlich, habe kein Gefühl und keine Seele.[73] An dem Stück hatte sich zwar nichts geändert, sehr wohl aber an den Voraussetzungen, unter denen es

71 Diesbezüglich lässt sich anschließen an Edmund Husserl: *Husserliana. Gesammelte Werke*, Bd. 33: *Die Bernauer Manuskripte über das Zeitbewusstsein (1917/18)*, hrsg. v. Rudolf Bernet / Dieter Lohmar, Dordrecht 2001, S. 43–88; ders.: *Husserliana. Gesammelte Werke*, Bd. 10: *Zur Phänomenologie des inneren Zeitbewußtseins (1893–1917)*, hrsg. v. Rudolf Boehm, Den Haag 1966, S. 29–31, 52 u. a. Vgl. Voigt: »Über das Verhältnis von Musik und Klang« (wie Anm. 66), S. 180–182. **72** Vgl. Fuchs: *Das Gehirn – ein Beziehungsorgan* (wie Anm. 46), S. 153f. **73** Sautoy: *Der Creativity Code* (wie Anm. 62), S. 203f.

wahrgenommen wurde. Der Zuhörer fühlte sich einem Betrug aufgesessen und das zu Recht. Für ein Musikstück ist nicht allein wichtig, wie seine Klangstrukturen beschaffen sind, diese können eine vollkommen andere Bedeutung erhalten durch die Zeit, zu der sie komponiert wurden, durch den Ort, an dem sie komponiert wurden, durch den institutionellen Rahmen, für den sie komponiert wurden, durch die Person des Komponisten, der sie komponiert hat und vieles andere mehr, einschließlich der vorhergehend erörterten Tatsache, dass Musik nicht aus Klängen schlechthin besteht. Diese Faktoren verändern keine Note und keinen Klang des Stücks, sondern sie verändern, als was es gehört werden sollte. Es ist auch nicht dasselbe, ob eine Amsel eine Melodie singt, ob ein Vogelfänger dieselbe Melodie pfeift, oder ob dieselbe Melodie in Richard Wagners *Siegfried-Idyll* erklingt – oder, um auf Putnam zurückzukommen, ob eine Ameise eine Spur im Sand hinterlässt, die wie eine Churchill-Karikatur aussieht, oder ein Karikaturist eine Churchill-Karikatur zeichnet. Die Problematik besteht folglich nicht allein mit Blick auf KI, aber doch besitzt sie in Bezug auf KI eine besondere Relevanz. Manipulativ vermag solche Täuschung immer genutzt zu werden. Die Täuschungen, die mit der KI verbunden sind, betreffen auf elementarer Ebene das Verhältnis des Menschen zu sich selbst und sein Verhältnis zur Realität.

Die erste Täuschung besteht in dem menschlichen Anstrich, der der Technik verliehen wird. Copes Kompositions-KI hieß »Emmy«, um sie menschenähnlicher erscheinen zu lassen. Nachdem er 2003 sein »Emmy«-Projekt einstellte, ersetzte er es durch eine interaktive Kompositions-KI, die er »Emily Howell« nannte. Aufgrund ihrer Interaktivität ließ sich ihr ein Subjektcharakter noch weit einfacher unterschieben als dem vorhergehenden KI-System »Emmy«.

Indem Cope menschliche Komponisten als Algorithmen ausführend beschreibt, unterstellt er ihnen Maschinenhaftigkeit. Seit dem 17. Jahrhundert fanden vergleichbare Strukturen im musikalischen Denken rasche Verbreitung und im selben Zuge die Idee, mittels dieses Wissens musikalisch Affekte und Emotionen effektiv steuern zu können. Unvorbereitet sind Auffassungen wie die Copes also nicht entstanden. Infolge des Staunens vor den Leistungen von KI-Systemen kommt also die Idee einer Reduktion des Denkens und Fühlens auf Algorithmen hinzu. Anscheinend besitzt es inzwischen mehr Plausibilität, die Simulation von Denken als Denken aufzufassen, als das Denken und Fühlen in der einzigen Weise zu begreifen, in der es begriffen werden kann, nämlich indem man selbst denkt und fühlt. Dieser Typus von Täuschung wird in sogenannten Wizzard-of-Oz-Experimenten unter anderem für das Trainieren von KI-Systemen für die Emotionserkennung nutzbar gemacht. Die Probanden werden dabei in dem Glauben gelassen, sie seien mit einem autonomen künstlichen System konfrontiert. Tatsächlich wird die Apparatur von den menschlichen Experimentatoren gesteuert.[74] Diese Experimente könnten mit mehr Recht Wolfgang-von-Kempelen-Experimente heißen. In ihnen wird die Simulation in der Absicht simuliert, das Handeln der Probanden auf den Umgang

74 Misselhorn: *Künstliche Intelligenz und Empathie* (wie Anm. 50), S. 59 u. 71.

mit KI-Systemen einzustellen; ein hochgradig manipulatives Verfahren, das sich nicht zuletzt dazu eignet, durch Evokation falscher Situationsannahmen und den ihnen entsprechenden Erwartungen die Erfahrbarkeit der Differenz zwischen dem Umgang von Menschen untereinander und dem Umgang des Menschen mit Maschinen herabzusetzen. In der Produktgestaltung dienen Wizard-of-Oz-Experimente nicht zuletzt dazu, die Opazität von Technik und Handlungssituation zu maximieren, um digitale Systeme geschmeidig in den Alltag integrieren zu können.[75] Maschinen simulieren Menschliches, Menschen simulieren Maschinen; ohne Täuschung und Illusion, den entscheidenden Ingredienzien von Scharlatanerie, kommt die Integration von KI in die Gesellschaft nicht aus.

Die Frage nach der Adäquatheit der Auffassung, Denken sei das Ausführen von Algorithmen einmal hintangestellt, lassen sich bei ihrer Akzeptanz menschliches Denken und maschinelle Prozesse als äquivalent, wenn nicht gar als identisch verstehen. Tatsächlich müsste angesichts dieser Reduktion des Subjekts auf die Perspektive dritter Person alles subjektiv Erlebte, alle in erster Person innegehabten Vorstellungen vom Komponieren und vom Verstehen der Kompositionen ausgeklammert werden. Das geschieht jedoch nicht, vielmehr scheint erforderlich zu sein, die Musik als ausdrucksvoll, emotional bewegend oder als tief empfunden zu hören. Zwar sind Beschreibungen von Musik mittels eines solchen Vokabulars nicht identisch mit Äußerungen über eigene Gefühle, Vorstellungen etc., jedoch nicht sinnvoll denkbar ohne die prinzipielle Möglichkeit, ihrer inne zu sein. Das allerdings setzt den Erste-Person-Standpunkt voraus; kein möglicher Gehirnzustand ist eine Vorstellung oder ein Erlebnis von irgendetwas. Während das Subjekt sich in dieser Operation zum bloßen Objekt depotenziert, stattet es virtuell die Maschine mit seiner eigenen Subjektivität aus. Im Grunde ist dies lediglich die spezifische Ausprägung des im 17. Jahrhundert schon begangenen Fehlers, den Menschen als einen Automaten aufzufassen, zu dessen Funktionen es gehöre, Bewusstsein zu erzeugen, und dessen Selbst ein Phantom sei, statt das Selbst als den gesamten Leib umfassend zu begreifen. Diesen Einwand erhob Gottfried Wilhelm Leibniz gegen Locke. Sieht sich der Mensch in der KI wie in einem Spiegel oder erblickt er in ihr staunend eine höhere Macht, vollzieht er allemal eine Angleichung ans leblos Mineralische. Nur residiert nicht mehr wie in Hoffmanns *Bergwerken zu Falun* die anziehende und schreckenerregende Bergkönigin in einem Reich aus Kristall. An ihre Stelle treten dank Sand und seltener Erden die weit profaneren Illusionen der VR-Brille.

Das Schaffen von Kunst erfordert nicht zuletzt eine relative Autonomie, insofern das Subjekt über die Art und Weise, in der es Kunst schafft, selbst entscheidet, indem es nicht nur Gestaltungsregeln nach gegebenen Vorbildern erlernt, sondern neue Regeln einzuführen vermag, indem es sich selbständig auch gegen seine Hervorbringungen erweist – indem es sie beurteilt, entsprechend seinen selbstge-

75 Vgl. Sarah Diefenbach / Marc Hassenzahl: *Psychologie in der nutzerzentrierten Produktgestaltung. Mensch – Technik – Interaktion – Erlebnis*, Berlin 2017, S. 21f. u. 179.

wählten Gestaltungszielen verändert, verwirft etc. – und schließlich den Zweck seiner Tätigkeit selbst festlegen kann. Gerade diese Autonomie fehlt der KI. Häufig ist von autonomem Fahren die Rede, wenn ausgedrückt werden soll, ein Fahrzeug bewege sich von einer KI gelenkt durch den Straßenverkehr, nicht von einem Menschen. Autonom ist das Fahrzeug darin freilich nicht, um dies zu sein, müsste es das Fahrtziel selbst wählen können, die Fahrt beenden können, wann und wo es möchte, und selbstverständlich sich entscheiden können, die Fahrt gar nicht erst anzutreten. Gerade die relative Autonomie, über die Menschen verfügen, ist bei KI nicht gewünscht, und die Bedingungen dazu werden von ihr nicht erfüllt.[76] Vergleicht man das autonome Fahrzeug mit einer Kutsche, so fällt auf, dass die Fortbewegung mit letzterer eines Kutschers bedarf, um die Selbständigkeit der vorgespannten Pferde einzugrenzen, die sonst die Kutsche – etwa weil sie den Regen draußen vor dem Stall verabscheuen – vielleicht gar nicht bewegen, sie vielleicht nur auf die nächste Weide ziehen, nicht in die nächste Stadt. Verglichen mit dem KI-gesteuerten Fahrzeug besitzen die Kutschpferde ein hohes Maß an Autonomie.

Was Kunst betrifft, liegt der Vergleich von KI-generierten Bildern mit den Bildern von Affen nahe, um die Unterschiede zwischen autonomen und heteronomen Hervorbringungen zu veranschaulichen. In den 1950er Jahren ermunterte der Zoologe Desmond Morris den Schimpansen Congo zum Malen. Etwa 400 Bilder brachte Congo hervor, ohne Algorithmen zu folgen und ohne ein umfangreiches Beispielkorpus studiert zu haben.[77] Freilich war er nicht der erste und nicht der letzte malende Affe. Die Farbflächen und -muster, die Congo malte, brachte er aus eigenem Antrieb hervor, vielleicht ein wenig ungelenk, aber doch intentional, denn auf das von ihm Gemalte vermochte er mit Variationen zu reagieren und offensichtlich war er um Symmetrie bemüht. Auch lassen Affen sich beim Malen nicht gern unterbrechen, solange sie ihr Bild nicht als fertig erkennen.[78] Vielleicht ähneln die Malereien der Schimpansen, Orang-Utans und Gorillas am ehesten den Malereien von Kindern und weniger denen von Michelangelo oder Adolph Menzel, doch sind ihre Bilder einem Kunstbegriff, der Intentionalität ebenso voraussetzt wie die eigene Motivation zum Malen, näher als jedes mittels KI generierte Bild.

Epilog

Das Komponieren von Musik ist selbstverständlich eine eher nebensächliche Anwendung von KI, trägt aber durch die Faszination, die es auslöst, zur unreflektierten Durchsetzung einer Technologie bei, die wie das Perspektiv im *Sandmann* dem menschlichen Blick die Dinge in ihr Gegenteil verkehrt. Die Simulation von etwas

76 Vgl. Spiekermann: »Zum Unterschied zwischen künstlicher und menschlicher Intelligenz« (wie Anm. 38), S. 12–15. **77** Judith Koch: *Ausstellung mit Primatenbildern. Der mit dem Affen malte*, Deutschlandfunk, 18.12.2019, https://www.deutschlandfunkkultur.de/ausstellung-mit-primatenbildern-der-mit-dem-affen-malte-100.html [30.9.2025]. **78** Eine Webseite mit Beispielen von Affenkunst, interessanten Essays und einem Interview mit ChatGTP ist https://www.affenbrut.de/positionen [30.9.2025].

wird für das gehalten, was sie simuliert, während das, was simuliert wird, allenfalls noch als Epiphänomen in Betracht kommt. Die Realität wird nach Maßstäben ihrer technischen Repräsentationen betrachtet, während die Repräsentationen den Status von Wirklichkeit zugesprochen bekommen. Welche Musik jemand hören möchte, erfährt er über die personalisierten Vorschläge auf Spotify oder Amazon. Die KI ermittelt aufgrund vorheriger Suchen und Käufe die Wahrscheinlichkeiten, mit welchen ein bestimmtes Musikrepertoire gewählt wird und die Veröffentlichungen, die optimal zum jeweiligen Kunden passen. Niemand wird gezwungen, den Empfehlungen zu folgen, anscheinend aber ist es für viele Kunden der großen Online-Anbieter selbstverständlich geworden, die auf ihre Person abgestimmten Empfehlungen fraglos anzunehmen. Ein befreundeter Berliner Lehrer, der zukünftige Erzieher musikpädagogisch unterrichtet, löst bei den Auszubildenden mit der Frage, nach welchen Kriterien sie entschieden, welche Musik sie hören, häufig verständnisloses Staunen aus: Das sage einem doch Amazon. Von den eigenen Wünschen und Interessen erfährt das Subjekt durch ehemalige Entscheidungen, die die KI mit den Mitteln der Stochastik auf die Segmente des Musikmarktes abbildet und objektiviert. Dem Subjekt spiegelt dieses Verfahren die eigene Vergangenheit als die gegenwärtigen und zukünftigen Wünsche wider. Eine Auseinandersetzung mit dem, was Musik sein kann, erspart die KI dem Subjekt. Nicht einmal in der Musik soll es offenbar mit etwas konfrontiert werden, das ihm nicht schon vorweg entspricht. Das gelingt, indem das Subjekt von seiner Subjektivität entbunden wird. Was Handlung sein könnte, nähert sich auf diese Weise bloßem Verhalten an. Neben dem Widerstand, den die Welt dem Subjekt entgegensetzt, erfordert ein sich selbst in vollem Wortsinne bewusstes, sich reflektierendes Subjekt Offenheit (zu unterscheiden von Beliebigkeit), mit der es sich den Dingen, die sich nicht seinen eigenen Schemata fügen, zuwendet und mit der es anderen Subjekten begegnet. Ohne diese weitere Bedingung wäre kein auf Gründen aufbauendes Urteilen und Handeln denkbar. Die kybernetische Schleife, in der das Subjekt in einer von smarter Technologie bestimmten Gesellschaft kreisend festgehalten wird, unterbindet beides. E.T.A. Hoffmann untersucht nicht zuletzt die möglichen Antworten auf die Inkommensurabilität von Subjekt und Welt, letztere ist diesem nicht vollkommen anzupassen, ebensowenig die verschiedenen Subjekte zueinander. Dieser Inkommensurabilität durch eine technologische Illusion entkommen zu wollen, endet für seine Figuren fatal, der Dichter Ferdinand gerät in eine tiefe seelische Krise, Nathanael stürzt sich im Wahnsinn in den Tod.

Abstract

Musizierende Maschinen mit poetischem Gemüt.
Was lässt sich von E. T. A. Hoffmanns Kritik an Automaten über Künstliche Intelligenz lernen?
E. T. A. Hoffmann war ein ausgezeichneter Analytiker der Beziehungen des Menschen zu den Maschinen, insbesondere der Projektion von Subjekteigenschaften auf Maschinen, die zugleich auf das Selbstverständnis zurückwirkt, das der Mensch von sich selbst entwickelt. Anders als häufig angenommen, kommt die romantische Ästhetik der Frage nach möglicherweise schöpferischen Leistungen technologischer Apparate entgegen, da zu ihren wichtigsten Merkmalen kombinatorische Verfahren und Zufallsproduktion gehören. Beide führt Hoffmann in seinen Kompositionen und Texten vor. Überdies sind mechanische Automaten zentraler Gegenstand einiger der Erzählungen Hoffmanns, vor allem in *Der Sandmann* und *Die Automate*. Diese Sachverhalte legen nahe, seine Überlegungen als Ausgangspunkt für eine Untersuchung des heutigen Verständnisses von Künstlicher Intelligenz (KI) zu nutzen, gerade auch, was ihre Rolle in künstlerischen oder kreativen Prozessen angeht.

Der Beitrag fragt, ob die Zuschreibung subjektähnlicher Eigenschaften an KI-Technologie angemessen ist, und weitergehend, inwieweit der einer solchen Auffassung zugrundeliegende Subjektbegriff konsistent und tragfähig ist. Komplementär dazu wird untersucht, welchen Veränderungen der Diskurs um die KI den Begriff des menschlichen Subjekts unterzieht, damit dessen Merkmale als auf Technologie übertragbar erscheinen. Diese Begriffsübertragungen erweisen sich für den Selbst- und Weltbezug des Menschen als problematisch. So nimmt es nicht Wunder, dass, um KI als Subjekt oder Akteur begreifen zu können, auch manipulative Verfahren und solche der Täuschung erforderlich sind.

Im künstlerischen und musikalischen Feld zeigen sich einige Aspekte der subjektähnlichen Auffassung von KI mit besonderer Schärfe. Dies betrifft zunächst den Außenweltbezug von Kunst. Kunst geht nicht auf in der Anwendung von Mustern nach bestimmten Verfahren; das Einbrechen von ihnen nicht immanenten Ereignissen und Dingen in sie ist von entscheidender Bedeutung für den künstlerischen Prozess. Solche integrativen Leistungen können KI-Systeme vermutlich nicht erbringen, denen ein spontaner Außenweltbezug fehlt. Eine weitere Schwierigkeit ergibt sich aus dem schwierigen Verhältnis der ästhetischen Gegenstände und ihrer materiellen Träger. Allein aus einer erzeugten Klangfolge oder aus einer erzeugten Farbverteilung auf Leinwand, Papier oder Bildschirm lässt sich nicht erschließen, um welches Kunstwerk es sich bei den jeweiligen Klangfolgen oder Farbverteilungen jeweils handelt. Weder sind Musik und Kunst in ihrer Herstellung ohne Intentionalität und Interpretation denkbar noch sind sie es als Gegenstand. Aus diesem Sachverhalt resultieren schwerwiegende Probleme für das Verständnis der Resultate von KI als Kunst und generell für ihre Interpretierbarkeit.

Music-making Machines with a Poetic Mind.
What Can We Learn about Artificial Intelligence from E. T. A. Hoffmann's Criticism of Automata?
E. T. A. Hoffmann was an astute analyst of the relationships between humans and machines – particularly of the projection of subject-like qualities onto machines, a process that in turn affects the self-understanding humans develop about themselves. Contrary to common assumptions, Romantic aesthetics is in fact quite compatible with the question of potentially creative achievements by technological apparatuses, since two of its defining features are combinatorial methods and the production of chance. Hoffmann demonstrates both in his compositions and writings. Moreover, mechanical automata play a central role in several of Hoffmann's stories, especially *The Sandman* and *The Automata*. These aspects suggest that his reflections can serve as a starting point for investigating today's concept of Artificial Intelligence (AI), especially in relation to its role in artistic or creative processes.

This paper asks whether the attribution of subject-like qualities to AI technology is appropriate, and more fundamentally, whether the notion of the subject underlying such an attribution is coherent and viable. Complementarily, it explores how the discourse surrounding AI alters the concept of the human subject to make its features appear transferable to technology. These conceptual shifts are shown to be problematic for human self-conception and world-relation. It is therefore not surprising that, in order for AI to be perceived as a subject or agent, manipulative and deceptive strategies are often necessary.

In the artistic and musical fields, certain aspects of the subject-like understanding of AI become particularly pronounced. First, this concerns art's relation to the external world. Art is not merely the application of patterns or procedures; the intrusion of events and elements that are not immanent to those patterns is crucial to the artistic process. Such integrative capacities likely lie beyond the reach of AI systems, which lack spontaneous engagement with the external world. Another difficulty stems from the complex relationship between aesthetic objects and their material carriers. One cannot determine what artwork is represented merely by analysing a sequence of generated sounds or a pattern of colours on canvas, paper, or screen. Neither music nor visual art can be conceived – either in their creation or in their reception – without intentionality and interpretation. This poses major challenges for understanding AI-generated results as artworks and for their interpretability more broadly.

Autor

Boris Voigt ist wissenschaftlicher Mitarbeiter und Privatdozent an der Universität Heidelberg. Er studierte Musikwissenschaft, Soziologie und Philosophie. 2006 erfolgte die Promotion an der Universität Hamburg. Er war wissenschaftlicher Mitarbeiter am Institut für Musikwissenschaft und Medienwissenschaft der Humboldt-Universität zu Berlin und vertrat 2014/15 ebendort die Professur für Musiksoziologie und historische Musikanthropologie. 2018 erfolgte ebenfalls an der Humboldt-Universität die Habilitation mit einer Arbeit zur Semantik der politischen Ökonomie der Musik in modernen Gesellschaften. 2020/21 vertrat er eine Professur für Historische Musikwissenschaft an der Staatlichen Hochschule für Musik und Darstellende Kunst Stuttgart. Monographien: *Richard Wagners autoritäre Inszenierungen. Versuch über die Ästhetik charismatischer Herrschaft* (Hamburg 2003) und *Memoria, Macht, Musik. Eine politische Ökonomie der Musik in vormodernen Gesellschaften* (Kassel u. a. 2008).

Boris Voigt is a research associate and Privatdozent at Heidelberg University. He studied musicology, sociology, and philosophy. In 2006, he received his doctorate from the University of Hamburg. He has worked as a research associate at the Department of Musicology and Media Studies at Humboldt University in Berlin, where he also served as interim professor of music sociology and historical music anthropology in 2014/15. In 2018, he completed his Habilitation at Humboldt University with a study on the semantics of the political economy of music in modern societies. In 2020/21, he served as interim professor of historical musicology at the State University of Music and Performing Arts in Stuttgart. Monographs: *Richard Wagners autoritäre Inszenierungen. Versuch über die Ästhetik charismatischer Herrschaft* (Hamburg, 2003) and *Memoria, Macht, Musik. Eine politische Ökonomie der Musik in vormodernen Gesellschaften* (Kassel et al., 2008).