

«Mind the Game!»

Die Exteriorisierung des Geistes ins Spiel gebracht

«I propose to consider the question,
«Can machines think?»»

ALAN TURING, 1950

Der aktuelle Hype um die Künstliche Intelligenz (KI) schreibt die lange und verfehlte Geschichte der KI seit ihren Anfängen in den 1940er Jahren fort. Die Verfehlung – und das hatte bereits Joseph Weizenbaum 1979 deutlich gemacht – liegt in dem eingeschränkten Begriff von «Intelligenz», der, den psychologischen Intelligenztests folgend, ausschließlich die abstrakt-kognitiven Fähigkeiten selektiert und zu isolieren versucht.¹ «Für sich allein genommen ist Intelligenz ein sinnloser Begriff. Er erfordert einen Bezugsrahmen, die Definition eines Bereichs im Denken und Handeln, um ihm einen Sinn zu geben.»² Um die historische Rekonstruktion dieses Bezugsrahmens und seine aktuelle Umgestaltung geht es uns im Folgenden, denn am Beginn wie am aktuellen Ende der KI steht das Spiel als «für alles soziale Geschehen so charakteristische[r]!» Umstand.³ Tatsächlich lässt sich die Geschichte der KI entlang der Geschichte der durch die KI in Angriff genommenen Spiele erzählen: klassisch orientiert in den 1940er und 1950er Jahren an Strategiespielen wie Schach oder Checker,⁴ pragmatisch orientiert ab den 1980er Jahren durch *machine learning*-(ML-)gestützte Lösungen für die Schach- und Go-Algorithmen eines Deep Blue von IBM oder Alpha Go von Google.⁵ Heute ist KI ein integraler Bestandteil von Video- und Computerspielen. Vor diesem Hintergrund sind unsere Thesen folgende: erstens, dass die Verlagerung kognitiver Fähigkeiten in Kulturtechniken wie mathematische, logische und informatische Systeme – dem Programm der Exteriorisierung des Geistes folgend – sich aus Aspekten des Spielens generiert, aber in den Praktiken der KI karikiert wird;⁶ zweitens, dass sich bei historischer Betrachtung zeigt, dass Intelligenz sich in der Zwischenräumlichkeit zwischen den Menschen und ihren Kulturtechniken konstituiert, eine Zwischenräumlichkeit, die wir den «sozio-intelligenten Raum» nennen; drittens, dass dieser sozio-intelligente

¹ Vgl. Joseph Weizenbaum: *Computer Power and Human Reason. From Judgment to Calculation*, New York 1976, Kapitel 9 [dt.: *Die Macht der Computer und die Ohnmacht der Vernunft*, Frankfurt/M. 1994].

² Ebd., 271.

³ John von Neumann: Zur Theorie der Gesellschaftsspiele, in: *Mathematische Annalen*, Nr. 100, H. 1, 1928, 295–320, hier 298.

⁴ Vgl. Claude E. Shannon: Programming a Computer for Playing Chess, in: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Vol. 41, 1950, 2–13. Vgl. weiterführend Serjoscha Wiemer: *Strategiespiele und das Medium-Werden des Computers. Computerschach als Faktor der Rekonzeptionalisierung des Computers als programmierbare Maschine zwischen 1945 und 1960*, in: Rolf F. Nohr, Stefan Böhme, ders. (Hg.): *Diskurse des strategischen Spiels. Medialität, Gouvernementalität, Topografie*, Münster 2014, 83–112.

⁵ 1988 wurde IBMs Deep Blue Schachweltmeister, 2017 triumphierte Googles Alpha Go über den amtierenden Go-Meister Ke Jie.

⁶ Zum Begriff der Kulturtechnik vgl. Horst Bredekamp, Sybille Krämer: *Kultur, Technik, Kulturtechnik: Wider die Diskursivierung der Kultur*, in: dies. (Hg.): *Bild, Schrift, Zahl*, München 2003, 11–22.

Raum nirgends deutlicher zum Ausdruck kommt und sich nirgends besser beobachten wie erforschen lässt als in Video- und Computerspielen. Unser Aufruf an die Medienwissenschaften in der Reflexion der KI lautet daher: *mind the game!*

I. Exteriorität des Geistes

«Kulturtechnik befördert die Leistungen der Intelligenz durch Versinnlichung und exteriorisierende Operationalisierungen des Denkens».⁷ Kulturtechniken zu entwickeln, bedeutet kognitive Aspekte des Menschen zu exteriorisieren; d.h., sie zu dissoziieren, zu explizieren und in mathematischen, logischen und informatischen Systemen formalisiert auszulagern. Diese «Exteriorisierung des Geistes» ist die epochemachende Geistesbewegung des 20. Jahrhunderts und mündet seit den 1940er Jahren mit der Entwicklung digitaler Computer in ein weltweites Forschungsprogramm.⁸ Dieses Forschungsprogramm lässt sich am treffendsten mit Douglas Engelbart als *Augmenting Human Intellect* betiteln, und beschreiben als

increasing the capability of man to approach a complex problem situation, to gain comprehension, to suit his particular needs, and to derive solutions to problems [...] [by] more rapid comprehension, better comprehension, the possibility of gaining a useful degree of comprehension in a situation that previously was too complex, speedier solutions, better solutions, and the possibility of finding solutions to problems that before seemed insoluble.⁹

Die methodische Basis dieses Forschungsprogramms entstammt, so Engelbart, der Mathematik, der Logik, der Entscheidungs- und Klassifikationstheorie, der Zeit- und Bewegungsanalyse, der dynamischen Programmierung und last, but not least der Spieltheorie.¹⁰ Woraus das zunehmende Interesse an der Theoretisierung und Formalisierung des Spiels bzw. des Spielens im Zuge der Exteriorisierung des Geistes resultiert, hatte John von Neumann bereits 1928 dargelegt. Es sind nicht die Zufallsspiele wie Roulette, «die in der sog. <Theorie des Glückspiels>» der Wahrscheinlichkeitsrechnung seit dem 18. Jahrhundert behandelt werden, die interessant seien, sondern jene Gesellschaftsspiele wie Schach oder Kartenspiele, die maßgeblich durch der «vom Willen der Spieler abhängigen Variable y_v » dominiert werden.¹¹

Es soll versucht werden, die Rückwirkungen der Spieler aufeinander zu untersuchen, die Konsequenzen des (für alles soziale Geschehen so charakteristischen!) Umstandes, daß jeder Spieler auf die Resultate aller anderen einen Einfluss hat und dabei nur am eigenen interessiert ist.¹²

Die vom Willen der «Spieler» abhängige Variable y_v , die später auch Norbert Wiens kybernetisches Programm antreiben wird,¹³ rückt den sozialen Aspekt in den Mittelpunkt formalisierter (Spiele-)Handlungen, auch wenn diese nicht unbedingt kollaborativ oder bereits als sozio-intelligenter Raum zu denken sind. Aber sie sind dessen Voraussetzung, denn nicht die automatisierte Abfolge

⁷ Bredekamp u. a.: Kultur, 18.

⁸ Die Exteriorisierung des Geistes lässt sich mit André Leroi-Gourhan zivilisationshistorisch als Medialisierung von Körperfunktionen mit Hilfe verschiedener – mit Ernst Cassirer gesprochen – symbolischer Formen beschreiben. Vgl. André Leroi-Gourhan: *Le geste et la parole*, 2 Bde., Paris 1964, 1965; Ernst Cassirer: *Philosophie der symbolischen Formen*, 3 Bde., Berlin 1923–1929; Markus Rautzenberg: *SPIELRÄUME: Artificial Agency und die Exteriorisierung des Geistes*, Vortrag 22.11.2018, Peter Behrens School of Arts, Düsseldorf 2018.

⁹ Douglas Engelbart: *Augmenting Human Intellect: A Conceptual Framework*, Menlo Park 1963, 1.

¹⁰ Vgl. ebd., 102.

¹¹ Neumann: Zur Theorie der Gesellschaftsspiele, 298.

¹² Ebd.

¹³ Norbert Wiener: *Cybernetics or Control and Communication in the Animal and the Machine*, Boston 1948.

von Regeln, sondern die absichtsvolle Auswahl der durch Regeln beschreibbaren wie gesetzten Möglichkeitsbedingungen charakterisieren Spiele als auch die Erweiterung des menschlichen Intellekts. Im Unterschied zur Theorie des Glücksspiels, die den einzelnen <Spielenden> durch Mittelung eliminiert, ist die von John von Neumann inspirierte Spieltheorie durch die individuellen Spielenden und ihre Strategiemöglichkeiten gekennzeichnet.

II. Zwei Traditionen der «artificial agency»

Wenn die Parallelführung von absichtsvoller Motivation der Auswahl und des Einsatzes von Regeln sowohl Spiele wie auch die Augmentation des menschlichen Intellekts charakterisiert, dann stellt sich als Nächstes die Frage, wie weit menschliche <Intelligenz> sich in dieser absichtsvollen Motivation begründen lässt und Maschinen diese imitieren können. Eben diese Frage diskutiert Alan Turing in seinem Essay «Computing Machinery and Intelligence» von 1950.

The Argument from Informality of Behaviour: It is not possible to produce a set of rules purporting to describe what a man should do in every conceivable set of circumstances. [...] From this it is argued that we cannot be machines. I shall try to reproduce the argument, but I fear I shall hardly do it justice. It seems to run something like this. «If each man had a definite set of rules of conduct by which he regulated his life he would be no better than a machine. But there are no such rules, so men cannot be machines».¹⁴

In anderen Worten: Die Turingtafeln menschlicher Maschinen sind unbekannt und es ist auch nicht bekannt, ob diese endlich wären. Dennoch hält Turing an der Idee fest, dass Maschinen <Intelligenz> haben können. Dazu führt er den entscheidenden, aber fatalen Umkehrschluss der KI ein: Zum einen können Maschinen Intelligenz haben, weil unabhängig von der absichtsvollen Motivation auch für Menschen gelte, «that being regulated by laws of behaviour implies being some sort of machine»;¹⁵ zum anderen, weil die absichtsvolle Motivation der Auswahl und des Einsatzes von Regeln auch als Lernen verstanden werden könne und es dann vor allem um die Exteriorisierung von Lernmechanismen ginge, wie beispielsweise die klassischen, behavioristischen Lernstrategien: negative und positive Verstärkung.

The machine has to be so constructed that events which shortly preceded the occurrence of a punishment signal are unlikely to be repeated, whereas a reward signal increased the probability of repetition of the events which led up to it. These definitions do not presuppose any feelings on the part of the machine, I have done some experiments with one such child machine, and succeeded in teaching it a few things.¹⁶

Negative und positive Verstärkung trainiert Maschinen (wie auch Hunde) anhand einer mehr oder weniger subtilen Beeinflussung der absichtlichen Motivation. Was noch fehlt, in Turings Lesart von «Computing Machinery and Intelligence», ist der Zufall.

¹⁴ Alan Turing: Computing Machinery and Intelligence, in: *Mind*, Vol. 49, 1950, 433–460, hier 452.

¹⁵ Ebd., 452.

¹⁶ Ebd., 457.

Now the learning process may be regarded as a search for a form of behaviour which will satisfy the teacher (or some other criterion). Since there is probably a very large number of satisfactory solutions the random method seems to be better than the systematic.¹⁷

Turings Essay endet mit der Hoffnung, «that machines will eventually compete with men in all purely intellectual fields», aber interessanter ist, wie es beginnt: mit Maschinen, «whose manner of operation cannot be satisfactorily described by its constructors because they have applied a method which is largely experimental».¹⁸ Heute kennen wir diese experimentellen Maschinen, die Turing lediglich andeutet, als *deep learning artificial neural networks* (DNNs).¹⁹ Deren Beginn liegt in dem Beweis von Warren S. McCulloch und Walter Pitts, dass künstliche neuronale Netze (KNNs) zu Turingmaschinen kombiniert werden können;²⁰ in Frank Rosenblatts «Perceptron» als erstem probabilistischen Modell des Lernens;²¹ in David H. Ackleys, Geoffrey E. Hinton und Terrence J. Sejnowskis Backpropagation-Algorithmus;²² in Kuniyiko Fukushimas «Neocognitron», als erstem hierarchischen KNN (*deep learning*) für die visuelle Mustererkennung;²³ und in Geoffrey E. Hinton, Simon Osinderos und Yee-Whyeh Tehs Lösung des Trainingsproblems für DNNs.²⁴

Die zweite Tradition der *artificial agency* gründet nicht in der «vom Willen der Spieler abhängigen Variable y_v »,²⁵ sondern als sogenannte symbolische KI in der semantischen Repräsentation von Wissen, das syntaktischen Regeln der natürlichen Sprache (zumeist Englisch) folgt. Analog zu von Neumann könnte man diese Tradition in der «vom Wissen der Spielenden abhängigen Variable y_w » verorten, allerdings mit dem Problem, dass Wissen sich weniger leicht formalisieren lässt als die absichtsvolle Motivation. Auch wenn sich die Anwendung der symbolischen KI in Form von Expert_innensystemen in den 1980er Jahren verorten lässt, so hat diese doch erst mit IBMs Watson an Fahrt aufgenommen. Nirgends zeigt sich die Variable y_w besser als in dem Gesellschaftsspiel Jeopardy, das Watson 2011 selbstredend gegen menschliche Mitspielende gewonnen hat. Was Watson auszeichnet, ist nicht nur seine immense Wissensbasis (Wikipedia), sondern seine Algorithmen der natürlichen Sprachverarbeitung, des *information retrieval* und der automatisierten Inferenz.²⁶

Ob man die beiden Traditionen als schwache und starke KI bezeichnen will oder ob man zwischen neuronaler und symbolischer KI streng unterscheidet, verliert zunehmend an Bedeutung, da heutzutage beide Ansätze verschmelzen.²⁷ Entscheidend für unsere Fragestellung ist jedoch, dass willentliche Handlungen als Lernen, und Lernen als Bestrafen und Belohnen in Turings Tradition der KI festgeschrieben und heute im Paradigma der KNNs umfassend ausgebeutet wird. Bestrafen und Belohnen als Optimierungsfunktion der KNN hat jedoch seinen «für alles soziale Geschehen so charakteristischen!» Ursprung verloren. Diesen Ursprung gilt es zurückzugewinnen und unsere Überlegungen gründen dabei auf zwei Beobachtungen.

¹⁷ Turing: Computing Machinery, 459.

¹⁸ Ebd., 435, 460.

¹⁹ Die Entwicklung zeigt, dass nicht alles bereits bei Turing nachzulesen ist. Vgl. Bernhard Dotzler: Kant und Turing: Zur Archäologie des Denkens der Maschine, in: Philosophisches Jahrbuch, Nr. 96, 1989, 115–131; Christoph Engemann, Andreas Sudmann (Hg.): Machine Learning. Medien, Infrastrukturen und Technologien der Künstlichen Intelligenz, Bielefeld 2018.

²⁰ Vgl. Warren S. McCulloch, Walter Pitts: A Logical Calculus of the Ideas Immanent in Nervous Activity, in: The Bulletin of Mathematical Biophysics, Vol. 5, 1943, 115–133.

²¹ Vgl. Frank Rosenblatt: The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, in: Psychological Review, Vol. 65, 1958, 386.

²² David H. Ackley, Geoffrey E. Hinton, Terrence J. Sejnowski: A Learning Algorithm for Boltzmann Machines, in: Cognitive Science, Vol. 9, 1985, 147–169.

²³ Vgl. Kuniyiko Fukushima: Neocognitron: A Hierarchical Neural Network Capable of Visual Pattern Recognition, in: Neural Networks, Vol. 1, 1988, 119–130.

²⁴ Vgl. Geoffrey E. Hinton, Simon Osindero, Yee-Whyeh Teh: A Fast Learning Algorithm for Deep Belief Nets, in: Neural Computation, Vol. 18, 2006, 1527–1554.

²⁵ Neumann: Zur Theorie der Gesellschaftsspiele, 298.

²⁶ Vgl. David Ferrucci, Eric Brown, Jennifer Chu-Carroll u. a.: Building Watson: An Overview of the DeepQA Project, in: AI Magazine, Vol. 31, 2010, 59–79.

²⁷ Vgl. Nils J. Nilsson: The Quest for Artificial Intelligence. A History of Ideas and Achievements, New York 2009.

III. Sozio-intelligente Spielräume

Die erste Beobachtung ist, dass das, was beide Traditionen der *artificial agency* gemein haben und ihre aktuelle Leistungsfähigkeit wie ihren Durchbruch erklären, die extreme Vergrößerung ihrer Datengrundlage ist. Zufall wird durch Big Data ersetzt.²⁸ Der Bias, den beide Traditionen ebenfalls gemeinsam haben, ist, dass diese Datengrundlage eine ausschließlich virtuelle ist. «Virtuell» bezeichnet hier im Sprachgebrauch der Informatik und der Ingenieurwissenschaften bereits digital aufbereitete Datengrundlagen wie Bilddaten im Internet oder inhaltliche Datengrundlagen des Text-Minings wie die von Wikipedia. Der aktuelle Erfolg der KI liegt im Virtuellen begründet und Letzteres wird durch das Internet 2.0, durch Common-Sense-Datenontologien wie Concept Net der MIT Media Lab-Ausgründung Luminoso oder die Standardisierung der Maschine-Maschine-Kommunikation des Internets der Dinge zunehmend KI-affiner gestaltet. Der Erfolg von KI wird daher durch die Virtualisierung des Sozialen und die ausgreifende «Datafizierung» befördert.²⁹ Denn es ist sehr viel schwieriger, aufwändiger und weniger erfolgreich, DNNs auf «reale Daten» wie Messdaten anzuwenden, als Watson gesprochene Konversation auswerten zu lassen, da der Kontextbezug jeweils individuell trainiert werden muss; ganz zu schweigen von der maschinenintelligenten Konzeptualisierung der Variabilität realweltlicher Kontexte. Wenn also der aktuelle Erfolg der KI im Virtuellen liegt, dann sind es wiederum die Spiele, in diesem Fall Video- und Computerspiele, die das fruchtbare Milieu für die Exteriorisierung des Geistes liefern.

Die zweite Beobachtung ist folgende: Während über einzelne KI in den Laboren der Wissenschaft und Industrie wie selbstfahrende Autos, intelligente Pflegeroboter oder intelligente *virtual humans* diskutiert wird, sind selbstfahrende Autos nicht erst seit Spielen wie Grand Theft Auto V von 2013 Alltag. Autonome *path finding*-Algorithmen, Kollisionsvermeidung oder Hinderniserkennung gehören ebenso zum Standard vieler Spiele wie Parsertechnologien in frühen Abenteuerspielen, *side kick*-KIs und zunehmend auch Aspekte des *affective computing*. Zudem ist es kein Zufall, dass der führende Grafikkartenhersteller Nvidia die entsprechende Hardware für reale selbstfahrende Autos herstellt und damit für den aktuellen Paradigmenwechsel von der Programmierung zum Maschinernen in der Automobilindustrie sorgt.³⁰ Vor dem Hintergrund, dass allein in Deutschland 34 Millionen Menschen Computerspiele spielen, formiert sich hier seit Längerem ein fruchtbares Milieu für die Exteriorisierung des Geistes, das sich als «sozio-intelligenter Spielraum» umschreiben lässt.

Bereits 2005 hat Seth Giddens darauf hingewiesen, dass Computerspiele die kulturellen Orte sind, die uns in einen «cybernetic circuit» mit jeder Menge nicht-menschlicher Charaktere – von Cyborgs über Mutant_innen bis hin zu Netzwerken und Robotern – verschalten.

²⁸ Dieser Sachverhalt zeigt sich ebenfalls im Design und in der Praxis moderner Computerspiele. So sind die *massively multiplayer online role-playing games* (MMORPG) nur spielbar, wenn eine Spielergemeinschaft, die Millionen zählt, am Spiel beteiligt ist und zur Spielfunktionalität beiträgt.

²⁹ Hier liegt einer der Ursprünge für zahlreiche «Bias-Probleme» der aktuellen KI. Vgl. Ulrike Bergemann: *biodrag. Turing-Test, KI-Kino und Testosteron*, in: Engemann u. a. (Hg.): *Machine Learning*, 338–364.

³⁰ Vgl. Benedikt Fuest, Philipp Vetter: *Diesem Chip werden Millionen autonom fahrende Autos gehorchen*, in: *Welt.de*, dort datiert 17.7.2018, www.welt.de/wirtschaft/article179447714/Zulieferer-Nvidia-dominiert-die-Technik-fuer-selbstfahrende-Autos.html, gesehen am 24.5.2019.

Yet in significant ways digital games can be seen not only as representations of a putative future technoculture – as a technological imaginary of new media – but also as actual instances of a technoculture here and now. To play a digital game is to plug oneself into a cybernetic circuit. Any particular game-event is realised through feedback between computer components, human perception, imagination and motor skills, and software elements from virtual environments to intelligent agents.³¹

Inwieweit diese nicht-menschlichen Charaktere durch KI gesteuert sind, hängt von ihrer Autonomie im Spiel ab. «In a game AI generally refers to the components of the program that respond most sensitively to the actions of the player».³² Spieledesigner Michael Mateas fasst diesen Einsatz der KI unter dem Begriff «expressive AI» zusammen:

«Expressive AI» in games «covers a diverse collection of programming and design practices including pathfinding, neural-networks, models of emotion and social situations, finite-state machines, rule systems, decision-tree learning, and many other techniques».³³

Dabei verfolgt «expressive AI» zwei Ziele: zum einen das kommerzielle Ziel, möglichst viele Interaktionen nicht prädestiniert algorithmisch, sondern KI-gesteuert zu automatisieren; zum anderen das psychologische Ziel, Spiele sozio-intelligent zu gestalten, um ihnen, mit Daniel Dennett gesprochen, Intentionalität zuschreiben zu können und damit im anthropomorphen Sinne real wirken zu lassen.³⁴ Aktuell kommt ein weiterer Einsatz der KI hinzu, der auf der Metaebene die Daten der Spielenden für Dritte auswertet, Ratings generiert oder das Verhalten der zukünftig Spielenden für neue Spiele vorhersagt: «to evaluate if people play the game as intended and investigate how gameplay and game design impact the playing experience».³⁵

³¹ Seth Giddings: Playing with Non-humans: Digital Games as Technocultural Form, in: *Proceedings of DiGRA 2005 Conference: Changing Views – Worlds in Play*, Vancouver 2005, o. S., hier 1. Seite.

³² Ebd., 10.

³³ Michael Mateas zit. n. ebd. Vgl. Michael Mateas: Expressive AI: Games and Artificial Intelligence, in: *Level Up Digital Games Research Conference proceedings & CDROM*, Utrecht 2003, o. S.

³⁴ Vgl. Daniel C. Dennett: Intentional Systems, in: *The Journal of Philosophy*, Vol. 68, 1971, 87–106.

³⁵ Tobias Mahlmann, Anders Drachen, Julian Togelius u. a.: Predicting Player Behavior in Tomb Raider: Underworld, in: *Proceedings of the 2010 IEEE Conference on Computational Intelligence and Games*, Dublin 2010, 178–185, hier 178.

³⁶ Dennett: Intentional Systems, 89.

IV. Paradox im Umgang mit «artificial agency»

Die Frage, die sich nun vor dem Hintergrund dieser beiden Beobachtungen stellt, ist die nach der Beschaffenheit des sozio-intelligenten Raums. Oder anders gewendet: Wie bekommt man die Intelligenz aus der KI und in den Zwischenraum zwischen die Menschen und ihre Kulturtechniken verlagert? In einem ersten Schritt ist dazu die Dennett zugesprochene Intentionalität von Nutzen, die Giddings für Computerspiele neu aufgerufen hat.

When one can no longer hope to beat the machine by utilizing one's knowledge of physics or programming to anticipate its responses, one may still be able to avoid defeat by treating the machine rather like an intelligent human opponent.³⁶

Das Erstaunliche an dieser von Dennett bereits 1973 formulierten Intentionalitätsthese ist, dass sie nicht nur die Akteurialität nicht-menschlicher Akteure der Akteur-Netzwerk-Theorie vorwegnimmt, wenn auch nur als Zuschreibungsoption, sondern dass sie ein Paradox im Umgang mit der *artificial agency*

offenlegt. Anders als bei Turing oder später in John Searles *chinese room*-Argument oder für Joseph Weizenbaums ELIZA,³⁷ die die Ununterscheidbarkeit zwischen Menschen und Maschine als Nachweis von KI formulieren, liegt das Gefühl, einer echten KI im Spiel gegenüberzustehen, in ihrer Unterscheidbarkeit zum Akteur «Mensch» (partielle Überlegenheit oder Andersheit) bei gleichzeitiger Ununterscheidbarkeit (autonomes Handeln, Adaptivität, Sensitivität etc. wie Menschen).

Intentionalität³⁸ ist ein von der Phänomenologie wie auch der Philosophie des Geistes aufgegriffener Begriff, der jedoch nicht unumstritten ist. Beispielsweise stellt sich die Frage, ob es auch nicht intentionale psychische Zustände wie Schmerzen gibt. Für Dennett ist dies jedoch irrelevant, denn für ihn sind intentionale Systeme solche, die durch die Modi der Intentionalität – Glauben, Fürchten, Hoffen, Wünschen etc. – auf etwas ausgerichtet sind. Dabei müssen Systeme keine intentionalen Zustände haben, man muss sie ihnen lediglich zuschreiben können als Strategien, um deren Verhalten zu erklären und vorherzusagen. Dieser Wechsel von der Ersten-Person- zur Dritten-Personen-Perspektive erlaubt es, eine intentionale Haltung nicht nur gegenüber Menschen, sondern auch gegenüber Tieren und Dingen einzunehmen. Aussagen, wie: «mein Computer nervt mich», sind solche intentionalen Zuschreibungen zu Dingen. Die spannende Frage im Kontext sozio-intelligenter Spielräume ist nun, inwieweit KI über diese bloße Zuschreibungsfunktion hinausgehen kann. Dennett selbst hat diese Frage am Beispiel des Schachcomputers untersucht. Dabei geht es ihm nicht um eine funktionale Erklärung, sondern darum, einem «intelligenten» Schachcomputer anhand der Vorhersage seines Verhaltens Rationalität zuzuschreiben. Damit autorisieren wir, so Dennett, den Schachcomputer als intentionales System. Ist es möglich, das Verhalten eines intentionalen Systems gut vorherzusagen, ist dieses ein «true believer»: «What it is to be a true believer is to be an intentional system, a system whose behavior is reliably and voluminously predictable via the intentional strategy».³⁹

V. Exteriorisierung des Geistes als epistemische Zwischenräumlichkeit

Doch Dennett gibt nur die halbe Antwort auf die obige Frage. Die andere Hälfte liefert Stanley Cavell, dessen Kritik der Intentionalität sich negativ gewendet als Kritik an Dennett lesen lässt; positiv gewendet aber als neuer Ansatz, mit dem Thema KI umzugehen.⁴⁰ Dies ist insbesondere vor dem skizzierten Paradox der «unterscheidbaren Ununterscheidbarkeit» der *artificial agency* interessant. Denn der sozio-intelligente Spielraum spannt sich gerade im Spannungsfeld zwischen Vorhersagbarkeit und unvorhersagbarer Überlegenheit bzw. Andersheit auf; also all jenen Erfahrungen mit KI, die unerwartet sind bzw. Widerständigkeit seitens der KI erzeugen. Wie ist diese Unterscheidbarkeit durch Widerständigkeit und Unvorhersagbarkeit der *artificial agency* zu deuten? Ist eventuell der Schritt vonnöten, *artificial agency* nicht als

³⁷ Vgl. Joseph Weizenbaum: ELIZA – A Computer Program for the Study of Natural Language Communication Between Man and Machine, in: *Communications of the ACM*, Vol. 6, 1966, 524–536; John Searle: *Minds, Brains and Programs*, in: *Behavioral and Brain Sciences*, Vol. 3, 1980, 417–457.

³⁸ Intentionalität als Beziehung eines Gedankens auf einen Inhalt oder ein Objekt, wie von Franz Brentano 1874 eingeführt, wird als die maßgebliche Eigenschaft des Mentalen im Unterschied zu physischen Zuständen gesehen. Vgl. Franz Brentano: *Psychologie vom empirischen Standpunkt*, Leipzig 1874.

³⁹ Daniel Dennett: *The Intentional Stance*, Cambridge 1984, 15.

⁴⁰ Vgl. Stanley Cavell: *Der Anspruch der Vernunft. Wittgenstein, Skeptizismus, Moral und Tragödie*, Berlin 2016.

rein anthropomorphe Zuschreibung aus welcher Perspektive auch immer zu verstehen, sondern als Ansatzpunkt für eine theoretische Re-Formulierung anthropozentrischer Modellierungen von Intelligenz zu nehmen?⁴¹

Die Frage nach der Überschreitung der bloßen Zuschreibungsfunktion à la Dennett wäre dann die nach der epistemischen Zwischenräumlichkeit sozio-intelligenter Spielräume. Eben hier kommt Cavell ins Spiel, dessen wichtigste philosophische Leistung die Lösung des Problems des Fremdpsychischen war. In seinem Hauptwerk *The Claim of Reason* versucht Cavell das grundlegende Rätsel der Intentionalität zu lösen.⁴² Dabei geht es um die schlichte, wenn auch weitreichende Beobachtung, dass man nie Gewissheit über die inneren Zustände anderer haben kann. Von René Descartes unter dem Namen Skeptizismus zum zentralen Problemkomplex der westlichen Philosophie erklärt, arbeitet sich die Philosophie von Descartes über Ludwig Wittgenstein bis zu Cavell daran ab; genauer gesagt an der Frage des intentionalen Status des Schmerzes. Auch wenn mir jemand seinen Schmerz noch so genau erklärt und beschreibt, gilt: Ich kann ihn nie mit Gewissheit nachempfinden, noch kann ich wissen, ob der Ausdruck der Schmerzempfindung überhaupt mit einem inneren Zustand übereinstimmt, d. h., es ist möglich, dass jemand simuliert, Schmerzen zu haben. Cavells Lösung ist dabei entwaffnend schlicht. Der Philosoph widerlegt den Skeptizismus nicht, sondern bestätigt ihn, hält ihn für unausweichlich: Ja, es gibt niemals Gewissheit über Gefühle, somatische Zustände oder Wahrnehmungen der anderen, aber das ist auch überhaupt nicht relevant, denn das was zählt, findet gar nicht *in*, sondern *zwischen* uns statt. Gleiches gilt für Intelligenz.

Anders als Dennetts Zuschreibungsperspektive von Intentionalität verortet Cavell das, was hier auf dem Spiel steht, im Begriff des «acknowledgements» und damit im epistemischen Zwischenraum. *Knowing* in diesem Sinne als Aktivität und eben nicht als Substantiv (*knowledge*) ist *a-knowledging*, also Anerkennung des Appells der anderen. Und dieses Anerkennen ist eben auch Erkennen im epistemologischen Sinne, *ein Wissen des und vom anderen*. Es ist nicht Gewissheit, sondern ein Wissen von den anderen, ein Erkennen und Anerkennen der Alterität der anderen und das bedeutet, auch wenn es auf den ersten Blick widersprüchlich klingen mag: Wissen im Sinne des Aner kennens der Unmöglichkeit einer Gewissheit über die inneren Zustände der anderen hat die Form von *Anteilnahme*. Diese ist nicht über Gewissheit, sondern über die volle Anerkennung der unverfügbaren Fremdheit der anderen überhaupt erst möglich. In Bezug auf das Beispiel der Aussage von anderen: «Ich habe Schmerzen», resümiert Cavell:

Hier könnte ich nun sagen, dass der Grund, weswegen «Ich weiß, du hast Schmerzen» kein Ausdruck von Gewissheit ist, der ist, dass es eine Reaktion auf dieses Zeigen ist; es ist ein Ausdruck von *Anteilnahme*. [...] Aber warum wird die Anteilnahme auf diese Weise ausgedrückt? Weil dein Leiden mich mit einer *Forderung* konfrontiert. Es reicht nicht, dass ich weiß (mir gewiss bin), dass du leidest – ich muss etwas tun oder offen legen (was immer getan werden kann). Kurz, ich muss es *anerkennen*, andernfalls weiß ich nicht, was «(dein oder sein) Haben von Schmerz» bedeutet.⁴³

⁴¹ Vgl. weiterführend zum Spannungsfeld von Anthropozentrismuskritik und Anthropomorphismus im Hinblick auf die anthropomediale Bedeutung von Spiel und (Computer-)Technik: Astrid Deuber-Mankowsky: Mediale Anthropologie, Spiel und Anthropozentrismuskritik, in: *Zeitschrift für Medien- und Kulturforschung*, Nr. 4, 2013, 133–148.

⁴² Vgl. Cavell: *Der Anspruch der Vernunft*.

⁴³ Stanley Cavell: Wissen und Anerkennen, in: ders.: *Die Unheimlichkeit des Gewöhnlichen und andere philosophische Essays*, Frankfurt/M. 2002, 39–75, hier 69, Herv. i. Orig.

Wichtig ist hier der Status der Anerkennung auf epistemologischer Ebene, denn es handelt sich an dieser Stelle nicht um einen rein ethischen Appell, sondern um eine Bedingung der Möglichkeit von Wissen, bei der es *gleichgültig* ist, ob der Appell gehört oder ignoriert wird: «Die vom Leiden ausgehende Forderung kann unbeantwortet bleiben. [...] Entscheidend ist jedoch, dass der Begriff in gleicher Weise durch sein Scheitern wie durch seinen Erfolg zur Geltung kommt.»⁴⁴ Die Pointe nochmal wiederholt: Gewissheit ist nicht erreichbar, auf diese kommt es aber auch gar nicht an. Von Bedeutung ist allein die im Akt des Ausdrucks in die Welt gesetzte Forderung, die als solche einen Zwischenraum eröffnet, in dem ich, du und die Welt erst sinnhaft in Beziehung treten. Es ist diese Aufforderung zum Tanz, diese Forderung nach Handlung, welche jetzt im Zentrum der Aufmerksamkeit steht. Auf sie kommt es an.

Der Schritt von der Containerillusion des Geistes, wie sie für Turing, aber auch Searle charakteristisch war, hin zur Exteriorität desselben muss natürlich auch im Feld der KI zum Tragen kommen. Jene Stimmen, die stets beteuern, dass der nächste Quantensprung in der KI nur deswegen noch auf sich warten lasse, weil die Hardware noch nicht so weit sei, verkennen die Möglichkeit, dass der ganze Ansatz vielleicht ausgereizt sein könnte und alle Gigahertz und Terabyte dieser Welt nichts daran ändern, das immer noch zu großen Teilen *inside* und eben noch nicht *out of the box* gedacht wird. Dies ist jedoch in den genannten Ansätzen buchstäblich der Fall, und das ist der Grund, warum der Begriff des Spiels hier immer wieder auftaucht. Spielraum oder Spielfeld erweisen sich als jene Zwischenräumlichkeit, in der etwas Sinn hat, etwas wissbar wird. KI befindet sich demnach genauso wenig auf einem Chip, wie Intelligenz in der Hirnschale der Einzelnen, sondern im Raum der gegenseitigen Anteilnahme der Aktant_innen untereinander.

VI. Interplay soziotechnischer Spielräume

Dass *artificial intelligence*, uminterpretiert als *artificial agency*, nicht nur als rein anthropomorphe Zuschreibung, sondern im epistemisch Zwischenräumlichen gedacht werden kann, zeigt sich auch zunehmend in der Praxis. Beispielsweise haben Lenneke Kuijer und Elisa Giaccardi eben jene Zwischenräumlichkeit mit dem Begriff der «co-performance» in die *human-computer interaction*-(HCI)-Forschung eingeführt.⁴⁵ Es geht den Autorinnen darum, den Anthropozentrismus im HCI-Design zu überwinden, denn «this human centered view on the subordinate role of artifacts is predominant in HCI-interpretations of practice theory».⁴⁶ Etwas ausführlicher:

The aim [...] is to explore the potential of co-performance in helping designers step away from «human-centered» narratives of smartness and autonomy. In this narrative the artefact is subservient to the realization of social practices «scripted» at design time. Instead, we argue, the locus of design can be shifted towards solutions that allow for a fundamentally recursive relation between design and use. In these

⁴⁴ Ebd., 70.

⁴⁵ Vgl. Lenneke Kuijer, Elisa Giaccardi: Co-performance: Conceptualizing the Role of Artificial Agency in the Design of Everyday Life, in: CHI '18 Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal 2018, Paper 125, 1–14.

⁴⁶ Ebd., 2.

solutions, human and artificial roles can then gain more room to change, according to the situated and evolving complementarity of capabilities and doings, <uniquely> human and <uniquely> artificial. This entails acknowledging computational artifacts as performers of practices in their own right.⁴⁷

Indem der Fokus auf den Spielraum zwischen Menschen und Artefakten gelegt wird – die Autorinnen selbst nutzen den Begriff <interplay> – kommen eben jene Aspekte des Paradoxes der <unterscheidbaren Ununterscheidbarkeit> der *artificial agency* in den Blick: Zum einen die Autonomie oder zumindest die Eigengesetzlichkeit des Artefakts (bzw. der KI) und zum anderen der Zwischenraum als Ort des Zusammenspiels zwischen User_innen und technischen Artefakten (bzw. KI). Wie man sich das konkret vorstellen kann, wird am einfachen Beispiel eines <therapeutischen> Objekts exemplarisch dargestellt. Die Rede ist von einer Jacke:

For example, in Rozendaal's prototype of a jacket for veterans suffering from post-traumatic stress disorder (PTSD), the veteran and the jacket collaborate towards the designed goal of reducing stress and anxiety. The jacket mirrors the stress level of the wearer through biofeedback, by actively helping the wearer to relax through deep abdominal breathing, and by inhibiting the veteran's movement when stress levels become too high to control. As such, the jacket is intended to prevent irrational or even aggressive behavior. Veterans need to rely on the jacket to calm down. However, in collaboration with the jacket, veterans may learn to internalize this competency.⁴⁸

Jacke und Patient_innen treten im Akt des Zwischenspiels miteinander in einen Austausch, der mit theoretisch begründeter Emphase als Kollaboration, als «Zusammenarbeit», bezeichnet wird; oder eben als *interplay* zwischen Exteriorisierung und Internalisierung: «veterans may learn to internalize this competency».⁴⁹ In jenem *interplay* der Exteriorisierung und Internalisierung löst sich das Paradox der <unterscheidbaren Ununterscheidbarkeit> als permanenter Akt des *acknowledgement* auf. Nicht von ungefähr werden als weitere Begriffe von den Autorinnen «entanglement» und «reciprocity» eingeführt; alles Begriffe, welche den Akzent auf die Zwischenräumlichkeit der beschriebenen Prozesse legen. Es ist den Autorinnen wichtig, hier den Begriff der <intelligence> durch den der «agency» zu ersetzen, denn nicht nur wird hiermit impliziert, dass die kollaborative Handlung zwischen Aktant_innen jenes Element ist, das eigentlich intelligent ist, auch wird damit die Zwischenräumlichkeit als performativer Raum sichtbar.

Unser Fazit ist daher: Intelligenz, verstanden als Kapazität von Einzelhirnen im Sinne der Containerillusion, wird zu Gunsten der Konzeption einer kollaborativen Exteriorität des Geistes verabschiedet. Somit ist die These von der Exteriorisierung des Geistes vollends in der KI- und Designforschung angekommen, anders gesagt: Nicht mehr Hirnschalen und Black Boxes, sondern *Spielräume* zeigen sich als eigentlicher Ort der Intelligenz, oder genauer: Sie *sind* Intelligenz. Um diese Intelligenz zu erforschen, bieten sich die sozio-intelligenten Räume der Video- und Computerspiele an. *Mind the game!*⁵⁰

⁴⁷ Kuijer u. a.: Co-performance, 2.

⁴⁸ Ebd.

⁴⁹ Ebd.

⁵⁰ Mit einem von der Volkswagen-Stiftung finanzierten Projekt möchten wir das methodische Repertoire wie Gamer-Experience-, Play- und Foren-Ethnografien entwickeln, um eben diesen sozio-intelligenten Raum im Computerspiel erforschbar zu machen.